The Handbook of Advanced Proficiency in Second Language Acquisition
This outstanding multi-volume series covers all the major subdisciplines within linguistics today to offer a comprehensive survey of linguistics as a whole.
To see the full list of titles available in the series please visit www.wiley.com/go/linguistics-handbooks
The Handbook of Advanced Proficiency in Second Language Acquisition

Edited by

Paul A. Malovrh and Alessandro G. Benati

WILEY Blackwell
In loving memory of
Andrew Francis Malovrh
(1946–2018)
Contents

Notes on Contributors ix
Acknowledgments xvii

1 Introduction
Paul A. Malovrh and Alessandro G. Benati

Part I Advanced L2 Capacity: Orientations on Acquisition 7

2 Systemic Functional Linguistics and Advanced Second Language
Proficiency
Marianna Ryshina-Pankova

3 Psycholinguistic Approaches and Advanced Proficiency
Leah Roberts

4 What Does Critical Period Research Reveal about Advanced L2 Proficiency?
Michael H. Long, Gisela Granena, and Fátima Montero

5 Generative Approaches to Second Language (L2) Acquisition
and Advanced L2 Proficiency
Jason Rothman, Fatih Bayram, Tanja Kupisch, Terje Lohndal,
and Marit Westergaard

6 Interaction-Driven L2 Learning: Advanced Learners
Nicole Ziegler and Lara Bryfonski

7 Sociocultural Theory: Mediating Learners toward Advanced Proficiency
Matthew E. Poehner

Part II Advanced Proficiency and Performance: Multiple Dimensions
and Contexts 131

8 Advanced-Level Grammatical Development in Instructed SLA
Heidi Byrnes

9 Individual Differences in Advanced Proficiency
Paula Winke and Susan M. Gass

10 The Prior Language Experience of Heritage Bilinguals
Cristina Sanz and Julio Torres
## Contents

11  Meeting the Demands of Globalization: One Goal of ISLA Research  
    **Paul A. Malovrh and Nina Moreno**  
    199

12  Task Condition Effects on Advanced-Level Foreign Language Performance  
    **Gavin Bui, Peter Skehan, and Zhan Wang**  
    219

### Part III  Advanced Phonology  
239

13  Advanced-Level L2 Phonology  
    **John Archibald**  
    241

14  Markedness and Advanced Development  
    **Fred R. Eckman**  
    264

15  Advanced Second Language Segmental and Suprasegmental Acquisition  
    **Kazuya Saito**  
    282

16  Connected Speech in Advanced-Level Phonology  
    **Burcu Gokgoz-Kurt and D. Eric Holt**  
    304

17  Voice Onset Time in Advanced SLA  
    **Alfonso Morales-Front**  
    323

### Part IV  Advanced Grammar  
341

18  Advanced-Level Mood Distinction  
    **Aarnes Gudmestad**  
    343

19  Advanced Conceptualizations of Tense and Aspect in L2 Acquisition  
    **M. Rafael Salaberry**  
    361

20  Inflectional Morphology  
    **Roumyana Slabakova**  
    381

21  Advanced Lexical Development  
    **Stuart A. Webb**  
    401

22  Word Order and Information Structure in Advanced SLA  
    **Cristóbal Lozano and Marcus Callies**  
    419

23  Advanced-Level Semantics  
    **Tania Ionin**  
    442

### Part V  Advanced-Level Pragmatics, Discourse, and Sociocultural Literacy  
461

24  Advanced-Level Pragmatics in Instructed SLA  
    **Feng Xiao**  
    463

25  Advanced Reading Proficiency in Collegiate Foreign Language Learners  
    **Keiko Koda and Sihui Echo Ke**  
    483

26  Advanced Second Language Pragmatic Competence  
    **Naoko Taguchi**  
    505

27  Advanced Rhetoric and Socially Situated Writing  
    **Gregg Fields and Paul Kei Matsuda**  
    527

28  Variable Structures and Sociolinguistic Variation  
    **Kimberly L. Geeslin**  
    547

Index  566
John Archibald (PhD, Toronto, Canada) has been Professor of Linguistics at the University of Victoria (Canada) since 2010, following 19 years at the University of Calgary (Canada). He specializes in second language phonology, and is author or editor of seven books and approximately 30 journal articles and book chapters.

Fatih Bayram is a researcher at the University of Portsmouth, UK. His research primarily focuses on heritage language bilinguals. Bayram investigates the acquisition process of home and societal language in childhood in immigrant contexts, outcomes of language development in (young) adulthood, and on-line processing of the grammatical system during language production and comprehension.

Alessandro G. Benati is Head of School of Languages and Area Studies and Professor of Second Language Acquisition at the University of Portsmouth (UK). He is internationally known for his research in second language learning and teaching, with special emphases on processing instruction. He has a strong publications record with over 15 established monographs and articles in international journals. He has coordinated national and international research projects and he is honorary visiting Professor at York St. John University (UK).

Lara Bryfonski is a PhD candidate in applied linguistics at Georgetown University. Her doctoral research focuses primarily on interaction and corrective feedback in second language acquisition as well as task-based language teaching and learning. Lara is also a licensed English as a second language (ESL) teacher and has taught ESL in a variety of contexts in the United States and abroad.

Gavin Bui is Associate Professor at the English Department of Hang Seng Management College in Hong Kong. His research interests lie in task-based language teaching, development of second language (L2) fluency and lexis, and L2 motivation. His recent publications appeared in Language Teaching Research (Sage, 2016) and TESOL Encyclopedia of English Language Teaching (Wiley, 2017).
Heidi Byrnes is George M. Roth Distinguished Professor of German Emerita at Georgetown University. Her research focuses on adult second language literacy acquisition, particularly at the advanced level, with articles, edited and co-edited books, and special journal issues addressing the development of advanced literacy, particularly in writing. She is a past president of the American Association for Applied Linguistics and is the recipient of numerous professional association awards. She currently serves as editor-in-chief of the Modern Language Journal.

Marcus Callies received a PhD in English linguistics from the University of Marburg (Germany). Since 2014 he has been Professor of English Linguistics at the University of Bremen (Germany). His main research interests are learner corpus research with a focus on lexicogrammatical variation, discourse-functional and pragmatic aspects of advanced learner varieties, and English for academic purposes. He is serving as editor of the International Journal of Learner Corpus Research and vice-president of the Learner Corpus Association.

Fred R. Eckman holds the rank of University Distinguished Professor in the Department of Linguistics at the University of Wisconsin-Milwaukee. Eckman’s major research interest is in second language (L2) acquisition, and has centered mainly on L2 phonology, with a secondary focus on L2 syntax, and L2 acquisition theory. His writings have appeared in a number of anthologies and professional handbooks, as well in several journals, including Applied Linguistics, Language Learning, Second Language Research, and Studies in Second Language Acquisition.

Gregg Fields is a doctoral student in Writing, Rhetorics, and Literacies at Arizona State University-Tempe. His research interests include second language writing, language identity, metacognition, writing pedagogies, and writing program administration.

Susan M. Gass is University Distinguished Professor at Michigan State University. She has published more than 30 books and more than 150 articles in the field of second language acquisition, with works translated into Russian, Korean, and Chinese. She is the co-author (with Alison Mackey) of Second Language Acquisition: An Introductory Course of Second Language Research: Methodology and Design. She is the winner of local, national, and international awards. She has served as the president of the American Association for Applied Linguistics and the Association Internationale de Linguistique Appliquée and is currently co-editor of Studies in Second Language Acquisition.

Burcu Gokgoz-Kurt (PhD, University of South Carolina, 2016) is currently a faculty member at Dumlupınar University (Turkey). Her research focuses on the acquisition of second language (L2) phonology, pronunciation instruction, L2 speech perception, and cognitive aspects of L2 processing. Her doctoral work investigated the effects of training in learning word-boundary palatalization as a connected speech phenomenon in L2 English in relation to attention control. She has extensive experience teaching English as a foreign language and English as a second language in the United States and Turkey.

Gisela Granena is an Assistant Professor in the School of Languages of the Universitat Oberta de Catalunya (Spain). She has published research on the role of cognitive aptitudes in both instructed and naturalistic contexts; aptitude-treatment interactions; task-based language teaching; measures of implicit and explicit language knowledge; and the effects of early and late bilingualism on long-term second language achievement. Recent publications include Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment (John Benjamins, 2013) and Cognitive Individual Differences in Second Language Processing and Acquisition (John Benjamins, 2016).

Aarnes Gudmestad is Associate Professor in the Department of Modern and Classical Languages and Literatures at Virginia Polytechnic Institute and State University. Her research focuses on the intersection between second language acquisition and sociolinguistics. Current projects address issues pertaining to morphosyntactic structures in native and second language Spanish and French. She has published articles in journals such as Language Learning, Studies in Second Language Acquisition, Studies in Hispanic and Lusophone Linguistics, and Canadian Journal of Linguistics.

D. Eric Holt (PhD, Georgetown University, 1997) is Associate Professor of Spanish and Linguistics at the University of South Carolina. Scholarly work treats phonological theory, language variation and change, Hispanic sociolinguistics, and language acquisition, particularly of connected speech phenomena in Spanish by English-speaking learners.

Tania Ionin is an Associate Professor of Linguistics at the University of Illinois at Urbana-Champaign. Her main areas of research are second language acquisition of semantics, with a focus on the nominal domain (definiteness, specificity, genericity, and the mass/count distinction), and experimental semantic investigations of quantifier scope and the interpretation of indefinites in Russian and English.

Sihui Echo Ke is an Assistant Professor at the University of Kentucky. Her recent research has focused on the contributions of morphological awareness to adult and child second language reading.

Keiko Koda is a Professor in the Department of Modern Languages at Carnegie Mellon University. Her research interests include second language reading, biliteracy development, and foreign language instruction and assessment.
Tanja Kupisch is Professor of Linguistics at the University of Konstanz and Professor II in the Language Acquisition, Variation and Attrition (LAVA) research group at UiT The Arctic University of Norway. Her research focuses on first, second, and third language acquisition, early bilingualism, and bilectal language acquisition. She has worked on various aspects of DP syntax, such as article omission, genericity, nominal gender, and adjective placement, as well as on pronunciation.

Terje Lohndal is Professor of English Linguistics at NTNU Norwegian University of Science and Technology, and Adjunct Professor at UiT The Arctic University of Norway. He also co-directs the AcqVA research group, which focuses on language acquisition, variation, and attrition. Lohndal attempts to identify the basic building blocks of language based on evidence from acquisition, linguistic variation, and attrition, focusing in particular on syntax and its interfaces with morphology and semantics.


Cristóbal Lozano (PhD, Essex University, UK) has worked on second language acquisition as a teaching fellow (Essex University), a postdoctoral researcher (Universidad Autónoma de Madrid, Spain), and a Fulbright visiting researcher (Pennsylvania State University). He has been Associate Professor of Applied Linguistics at the Universidad de Granada (Spain) since 2010. His research focuses on the second language Spanish and English acquisition of phenomena at the syntax–discourse interface by combining corpus and experimental methods. He directs the CEDEL2 corpus (Corpus Escrito del Español L2).

Paul A. Malovrh is Associate Professor of Hispanic Linguistics and Second Language Acquisition at the University of South Carolina. His research focuses on the potential and limitations of foreign and second language (L2) instruction in interlanguage development, with emphasis on L2 processing strategies across different instructional contexts. His recent work emphasizes advanced-level language proficiency, in particular, and explores its relationship with globalization and curricular design.

Paul Kei Matsuda is Professor of English and Director of Second Language Writing at Arizona State University. Co-founding chair of the Symposium on Second Language Writing and editor of the Parlor Press Series on Second Language Writing, Paul has published widely in various edited collections as well as journals
such as *College Composition and Communication*, *College English*, *English for Specific Purposes*, *International Journal of Applied Linguistics*, *TESOL Quarterly*, and *Written Communication*.

**Fátima Montero** is a PhD student in the Second Language Acquisition (SLA) program at the University of Maryland. Her research interests include English and Spanish instructed SLA, maturational constraints, implicit and explicit second language learning, negative feedback, individual differences, and Spanish applied linguistics.

**Alfonso Morales-Front** is Associate Professor in the Department of Spanish and Portuguese at Georgetown University. His research interests focus on prosodic aspects of Spanish phonology (syllable, stress, and intonation), the acquisition of first and second language phonology, and study abroad.

**Nina Moreno** is Associate Professor of Spanish at the University of South Carolina. Her areas of expertise include teacher education, Spanish applied linguistics, cognitive processes in language learning, and computer-assisted language learning. She has published articles and software reviews in journals such as *Language Learning*, *Foreign Language Annals*, and *CALICO*, and co-authored *Introducción a la lingüística hispánica actual: teoría y práctica* (Routledge, 2017).

**Matthew E. Poehner** is Associate Professor of World Languages Education and Applied Linguistics at The Pennsylvania State University. His research examines Dynamic Assessment, Mediated Development, and Systemic Theoretical Instruction, derived from Vygotsky’s Sociocultural Theory, as a basis for second language educational practice. His most recent book (co-authored with J. P. Lantolf) is *Sociocultural Theory and the Pedagogical Imperative in L2 Education: Vygotskian Praxis and the Research/Practice Divide* (Routledge, 2014).

**Leah Roberts** is Professor of Psycholinguistics in the Department of Education, and Leader at the Centre for Research in Language Learning and Use, at University of York (UK). Her research focuses on second language learning and processing at the word, sentence and discourse levels. Her recent research topics include what adult and child learners can acquire after only limited exposure to a new language, as well as the ways in which two languages interact in the mind of functionally bilingual speakers during real-time language comprehension.

**Jason Rothman** is Professor of Multilingualism and Language Development in the School of Psychology at the University of Reading (UK) and Professor II of Linguistics in the Language Acquisition, Variation and Attrition (LAVA) research group at UiT The Arctic University of Norway. His work is grounded within the generative conceptualization of mental linguistic representation and computation, investigating language acquisition and processing in children and adults’ native and non-native language. Recent work appears in *Studies in Second Language*
Marianna Ryshina-Pankova is Associate Professor of German at Georgetown University. As Director of Curriculum she is actively involved in the maintenance, evaluation, and renewal of the undergraduate curriculum and in mentoring graduate students teaching in the program. Her research interests include application of systemic functional theory to the study of second language (L2) writing development and advanced literacy assessment, as well as L2 pedagogy and content- and language-integrated curriculum design.

Kazuya Saito is a lecturer in the Department of Applied Linguistics and Communication at Birkbeck, University of London (UK). His research interests include how second language learners develop various dimensions of their speech in naturalistic settings, and how instruction can help optimize such learning processes in classroom contexts.

M. Rafael Salaberry (PhD, Cornell University, 1997) holds an appointment as Mary Gibbs Jones Professor of Humanities at Rice University. He is a Professor of Spanish Linguistics and Second Language Acquisition in the Department of Spanish and Portuguese and he is also the Director of the Center for Languages and Intercultural Communication (CLIC) at Rice University.

Cristina Sanz (PhD, University of Illinois at Urbana-Champaign) is Professor of Spanish Linguistics and Director of Spanish Programs. An expert on multilingual development and education, she is interested in the interaction between context, including study abroad and computer-assisted learning, and individual differences, especially the role of prior experience and sociocognitive variables in second language development. She has published over 80 books, articles, and chapters; her volume *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice* (Georgetown University Press, 2005) received the Modern Language Association’s Mildenberger Prize.

Peter Skehan is Professorial Research Fellow at St. Mary’s University (UK), having previously worked at the University of Auckland (New Zealand), the Chinese University of Hong Kong, and King’s College, London, amongst others. His research areas are foreign language aptitude and second language acquisition, especially task-based learning and performance. His recent publications include a chapter in the *Annual Review of Applied Linguistics* (2016) and an article on aptitude and grammar in *Applied Linguistics* (2015).

Roumyana Slabakova is Professor and Chair of Applied Linguistics at the University of Southampton (UK) and is Professor at UiT The Arctic University of Norway. Her research interests focus on the second language acquisition of meaning—more specifically, phrasal-semantic, discourse, and pragmatic meanings. Her monographs include *Telicity in the Second Language* (John Benjamins,
2001) and *Meaning in the Second Language* (Mouton de Gruyter, 2008). She co-edits the journal *Second Language Research* and is the co-author (with Naoko Taguchi) of *Second Language Pragmatics* (Oxford University Press, 2017).

**Naoko Taguchi** is a Professor in the Department of Modern Languages at Carnegie Mellon University where she teaches courses in second language acquisition and Japanese language/culture. She is the author of *Context, Individual Differences, and Pragmatic Development* (Multilingual Matters, 2012) and *Developing Interactional Competence in a Japanese Study Abroad Program* (Multilingual Matters, 2015), and the co-author (with Roumyana Slabakova) of *Second Language Pragmatics* (Oxford University Press, 2017).

**Julio Torres** (PhD, Georgetown University) is an Assistant Professor of Applied Linguistics in the Departments of Spanish and Portuguese and Linguistics at the University of California, Irvine. He is also the director of the Spanish language program, and is affiliated with the National Heritage Language Resource Center at UCLA. His research interests include heritage/second language acquisition, bilingualism, cognition, and task-based language learning.

**Zhan Wang** is Assistant Professor at the Center for Language Education at the South University of Science and Technology in China. She received a PhD in Applied Linguistics from The Chinese University of Hong Kong and completed postdoctoral training at the University of Pittsburgh and Pittsburgh Science of Learning Center. Her research interests include task-based language teaching, psycholinguistics, and computer-supported collaborative learning.

**Stuart A. Webb** is a Professor of Applied Linguistics at the University of Western Ontario (Canada). Before teaching applied linguistics, he taught English as a foreign language in Japan and China for many years. His research interests include vocabulary, second language acquisition, and extensive reading, listening, and viewing. His latest book (co-authored with Paul Nation), *How Vocabulary is Learned*, was published by Oxford University Press in 2017.

**Marit Westergaard** is Professor of English Linguistics at UiT The Arctic University of Norway and Adjunct Professor at NTNU Norwegian University of Science and Technology. She is also head of the LAVA/AcqVA research groups (Language Acquisition, Variation and Attrition). She has published widely in first as well as second/third language acquisition, multilingualism in children and adults, and diachronic change, mainly focusing on linguistic properties such as word order and grammatical gender.

**Paula Winke** is an Associate Professor at Michigan State University. She teaches language teaching methods and language assessment in the Second Language Studies, TESOL, and Applied Linguistics Program within the School of Language Sciences & Literary and Cultural Studies. She researches language assessment
issues and how individual differences affect second language acquisition. She is a former president of the Midwest Association of Language Testers (MwALT) and in 2012 received the TESOL Award for Distinguished Research.

**Feng Xiao** is Assistant Professor of Asian Languages and Literatures at Pomona College, Claremont, California. His research interests include pragmatics, second language acquisition, bilingualism, and Chinese linguistics.

**Nicole Ziegler** is Assistant Professor of Second Language Studies at the University of Hawai‘i at Mānoa. Her research program focuses on instructed second language acquisition (ISLA), including mixed method and interdisciplinary research in second language conversational interaction, task-based language teaching, and computer-assisted language learning, specifically synchronous computer-mediated communication.
The inception of the present volume goes back to the summer of 2015, when we had a conversation regarding the growing need for advanced-level users of foreign language in the private and public sectors in an increasingly globalized world. Our conversation quickly evolved into a list of questions regarding what we really know (and do not know) about advanced proficiency in the field of second language acquisition (SLA). While extensive research had been conducted on the topic, we understood that more questions than answers remained. Previous research had already established various frameworks for pursuing the investigation of advanced proficiency and had posited specific questions to be addressed; this volume merely attempts to bring together some of the ways in which such questions are being explored by many of the field’s leading scholars, with the aim of drawing further attention to the need for a greater understanding of higher levels of language use. A search of key terms using any search engine will continuously yield a list of names of scholars who have dedicated their attention to pursuing the topic—names such as Byrnes, Maxim, and Ortega, to name but a few—and we acknowledge that, without their seminal work, the present volume could never have been realized. Their work provided us with the basic structure of the volume, as well as the specific themes to be explored, and we are grateful for their contributions to the field of SLA.

When we first presented our idea for the edited volume to Wiley, the initial response was extremely encouraging, and our proposal was quickly polished and submitted for review. The reviewers’ responses were unanimous: the need for a comprehensive collection addressing advancedness in SLA, in terms of varying theoretical approaches, multiple perspectives and orientations to context, tasks, and learner profiles, phonological and grammatical development, genre, and socially situated use, was long overdue. We were then fortunate enough to bring on board an impressive list of scholars to contribute their expertise to the volume. We, therefore, wish to thank first and foremost, the editorial and production teams at Wiley. Tanya McMullin, our acquisitions editor, provided us with continued guidance and feedback to keep the project on schedule. Manish Luthra, our production editor, and Giles Flitney, our copy-editor, were continuously available to provide advice regarding formatting issues throughout the editing process.
In addition, we thank the outstanding and wonderful contributors who shared their expertise and research in each chapter of the collection. In all cases, we found the editorial process to be enjoyable, as we received nothing short of complete professionalism and cooperation from those whose work comprise the present volume. In short, the project is indebted to the wonderful support of the Wiley team and to the scholars who contributed to it.

As is the case with any volume of this breadth and depth, a considerable amount of time was devoted to editing and peer review. To that end, we wish to thank a number of individuals who selflessly provided their insight and thoughtful feedback in order to maintain a high quality of work in the final draft. They are:

<table>
<thead>
<tr>
<th>John Archibald</th>
<th>Lara Bryfonski</th>
<th>Heidi Byrnes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fred R. Eckman</td>
<td>Kimberly L. Geeslin</td>
<td>Burcu Gokgoz-Kurt</td>
</tr>
<tr>
<td>Aarnes Gudmestad</td>
<td>Lara Gurzynski Weiss</td>
<td>D. Eric Holt</td>
</tr>
<tr>
<td>Tania Ionin</td>
<td>Keiko Koda</td>
<td>Michael H. Long</td>
</tr>
<tr>
<td>Alison Mackey</td>
<td>Alfonso Morales-Front</td>
<td>Nina Moreno</td>
</tr>
<tr>
<td>Matthew E. Poehner</td>
<td>Leah Roberts</td>
<td>Jason Rothman</td>
</tr>
<tr>
<td>Marianna Ryshina-Pankova</td>
<td>Kazuya Saito</td>
<td>Peter Skehan</td>
</tr>
<tr>
<td>Roumyana Slabakova</td>
<td>Naoko Taguchi</td>
<td>Julio Torres</td>
</tr>
<tr>
<td>Paula Winke</td>
<td>Feng Xiao</td>
<td>Nicole Ziegler</td>
</tr>
</tbody>
</table>

We close our acknowledgement with mention of some of those who have been instrumental in the successful completion of the present volume, through their support and sacrifices, and through their understanding of the importance of our work. They include department chairs, such as Dr. Nicholas Vazsonyi at the University of South Carolina, who acknowledges the important work of his SLA professors through his support for and enthusiasm toward projects such as this. They also include colleagues, such as Professor Matthew Weait, and all members of CARILSE (Centre for Applied Research and Innovation in Language Sciences and Education) at the University of Portsmouth (UK), who understood the value of this project and provided their support and excellent advice throughout. And they include our loving families, our spouses in particular, who have come to know far more about SLA than they probably ever expected to, through their ongoing support of (and interest in) our work. And, finally, they include our graduate students, whose curiosity and thoughtful questions regarding acquisition and advanced proficiency provide a constant motivation for learning more about such topics. We dedicate this volume to them, and to all scholars who strive to understand more about second language acquisition.

Paul A. Malovrh
University of South Carolina, USA

Alessandro G. Benati
University of Portsmouth, UK

December 2017
1 Introduction

PAUL A. MALOVRH1 AND ALESSANDRO G. BENATI2

1 University of South Carolina
2 University of Portsmouth, UK

Globalization over the past 20 years has engendered a renewed interest in language learning among researchers and employers alike. Increasingly, studies in second language acquisition (SLA) not only address the capacity and/or potential of advanced language learning and use, but also consider the social demands for multilingual actors, as the traditionally popular descriptors of an individual’s language ability—those such as bilingual or fluent—fail to sufficiently describe proficiency in everyday life. What does it mean to be advanced in a second (L2) or foreign (FL) language?

The need for a more visible and understandable profile of advanced-level language use—one that is understood by theorists and practitioners alike—serves as the main impetus for the present volume. The volume is of particular interest to theorists and scholars of SLA who seek out a better understanding of advanced-level proficiency as well as a comprehensive update of contemporary research. The need for such an update can be seen in a variety of recent works in the field, such as The Routledge Handbook of Second Language Acquisition (Routledge, 2012), which contains a chapter devoted specifically to advanced-level proficiency. In it, Heidi Byrnes (the author) calls for a more comprehensive theoretical basis for understanding advanced-level language use and capacity, one that combines cognitive, social, semantic, and textual orientations to acquisition. The present volume responds to such an invitation by bringing together a state-of-the-science review of literature addressing various orientations to advanced-level proficiency, including different approaches to exploring the L2 learner’s capacity for advanced-level language use, the complexities of defining advanced proficiency across various genres and socially situated contexts, and a linguistic profile of the performance of advanced users across phonological, morphosyntactic, and pragmatic domains. From a theoretical perspective, the volume addresses a growing need (and interest)
in SLA to better understand advanced proficiency. From a practical one, it provides a profile and description of advanced-level performance according to a variety of contexts, across specific genres, through different modes of communication, and in terms of grammatical structure.

The volume is motivated by a growing need to understand advanced proficiency, within academic and professional circles, as well as the potential and limitations of classroom instruction. In addition to developing the skills necessary for students to analyze and critique literary and cultural texts, language programs must also produce advanced- to superior-level users of foreign language in order to meet the growing societal and global demands for multilingualism. Program assessment metrics often use descriptors of student performance such as advanced, near-native, professional, and fluent, to name just a few, but without any shared understanding among directors, department chairs, professors, advisors, or SLA experts, of the meaning of such terms. What does an advanced-level learner look like in terms of linguistic knowledge and ability to communicate? Does enrollment in an upper-division university course constitute an advanced level of linguistic knowledge? Is a fluent language user also an accurate one? What level of language proficiency can realistically be achieved by a group of students in four years of study? Is studying abroad the only way for a student to become an advanced speaker? All too regularly, such questions lead to confusion, frustration, and discord among practitioners. This volume addresses such questions, balancing empirical research and literature reviews to describe and account for advanced-level development, psycholinguistic phenomena, variable performance, and the potential of fostering development through classroom intervention.

Previous research addressing advanced proficiency has tended to do so by following a specific theoretical approach, ranging from psycholinguistic processing strategies, to cognition, to the critical period, and to ultimate attainment, to name a just a few. Other works have focused on variable performance according to either individual differences or contextual and situational constraints. And others have taken a pedagogical interventionist approach as a means to explore the limits of instruction at advanced levels, usually arguing for reform in curricular design. The current volume aims to synthesize various dimensions of advanced proficiency with different orientations of, approaches to, and capacity for advancedness. It consists of five main parts organized in such a way that various fundamentally distinct themes may be addressed. The first part addresses the capacity for advanced proficiency by reviewing literature within distinct theoretical frameworks and approaches that address advanced-level development. Marianna Ryshina-Pankova provides an overview of systemic functional linguistics in Chapter 2, Leah Roberts addresses psycholinguistic approaches in Chapter 3, and Michael H. Long, Gisela Granena, and Fátima Montero ask what research regarding the critical period may tell us about advanced proficiency in Chapter 4. Jason Rothman, Fatih Bayram, Tanja Kupisch, Terje Lohndal, and Marit Westergaard discuss advanced proficiency according to generative grammar in Chapter 5. Interaction-driven approaches to advancedness are discussed in Chapter 6 by
Nicole Ziegler and Lara Bryfonski, and Matthew E. Poehner provides an overview of advanced proficiency according to Sociocultural Theory in Chapter 7.

Part II of the volume explores the complexity of defining advanced-level language use, given the multitude of internal and external factors constraining or obscuring our ability to assess proficiency. It focuses on contextually constrained language use and individual differences in advanced performance, as well as different methods of intervening with development in instructed settings. In Chapter 8, Heidi Byrnes provides a thorough call for future research regarding advanced-level grammatical development in instructed contexts from the perspective of systemic functional linguistics, whereas Paul A. Malovrh and Nina Moreno, in Chapter 11, examine the need for structural reform in basic language programs from the perspective of language processing, arguing that development toward advanced proficiency in instructed settings needs to start from the ground up. In Chapter 9, Paula Winke and Susan M. Gass offer profiles of advanced learners in the form of clusters based on empirical data measuring individual differences, and Cristina Sanz and Julio Torres discuss prior language experience of heritage bilinguals in Chapter 10. Concluding Part II, Gavin Bui, Peter Skehan, and Zhan Wang provide an overview of the effects of task conditions on advanced-level performance in Chapter 12.

The volume then proceeds to examine advanced proficiency from a linguistic perspective. Part III begins a review of literature yielding a profile of advanced-level language users’ capacity for phonological development and performance according to several phonological constructs. John Archibald begins the section by providing an overview of advanced-level phonology in Chapter 13, followed by Fred R. Eckman’s chapter on markedness in the context of advancedness in Chapter 14. Kazuya Saito explores segmental and suprasegmental advanced acquisition in Chapter 15, Burcu Gokgoz-Kurt and D. Eric Holt investigate connected speech in Chapter 16, and Alfonso Morales-Front discusses voice onset time (VOT) in Chapter 17.

Another unique feature of the present volume is its compilation of research and literature reviews with regard to the production and grammatical development. Part IV explores the notion of linguistic profiles, and yields a state-of-the-science update regarding various structures and features. In Chapter 18, for example, Aarnes Gudmestad explores four possible ways of classifying advanced proficiency in terms of mood distinction, and in Chapter 19, M. Rafael Salaberry describes previous theoretical definitions of aspect and the possible effects of explicit instruction on advanced knowledge. Roumyana Slabakova examines L2 learners’ variability in their use of inflectional morphology in Chapter 20, and Stuart A. Webb, in Chapter 21, explores lexical development and how we can determine if an L2 learner has reached (or is at) an advanced level. In Chapter 22, Cristóbal Lozano and Marcus Callies argue that one hurdle for advanced L2 learners is the acquisition of lexical and morphosyntactic alternations, when constrained by information-structure factors, and Tania Ionin provides a review of current research evidence on advanced learners’ knowledge of semantics in Chapter 23.
Introduction

Part V completes the volume with a profile of advanced-level performance in terms of socially situated language use by exploring topics such as cultural literacy, interlanguage pragmatics, and advanced rhetoric and writing. Feng Xiao begins the section, in Chapter 24, by reviewing the results of empirical research measuring pragmatic competence of high proficiency levels in instructed contexts. Keiko Koda and Sihui Echo Ke, in Chapter 25, define reading as a three-phased process among tertiary-level FL learners, consisting of text meaning building, personal meaning construction, and knowledge refinement. In Chapter 26, Naoko Taguchi offers a profile for advanced pragmatic competence by synthesizing findings from cross-sectional studies comparing L2 learners’ pragmatic performance, and in Chapter 27, Gregg Fields and Paul Kei Matsuda explore rhetoric and writing as two important factors describing and defining advanced second language acquisition, and they discuss how these two factors influence the development of socially situated language acquisition in general. Finally, in Chapter 28, Kimberly L. Geeslin identifies the abilities used by second language learners to vary the language they produce. She discusses how linguistic variation among native speakers functions as a mechanism to express and interpret information about individual characteristics of the speaker and the context in which the interaction takes place. The chapter provides an analysis of how these abilities might be developed in second language acquisition.

Each chapter in the volume concludes with a description or definition of advancedness according to its respective content, and provides a call for future research. Among them, the reader will note various consistent threads interwoven throughout the volume, such as the need for more longitudinal data, the need for more direct analysis of internal mechanisms using on-line measures, and the need for the triangulation of research data elicited (or collected) by different means. In short, the volume makes a strong argument for the need for more sophisticated research design, as well as more analyses of multiple advanced-level skills. It also sets the stage for meeting a number of goals established in previous SLA research. Consider, for example, the list of directions, posited in Byrnes’s (2012, pp. 516–517) conclusion, which need to be pursued in order to better understand advancedness in SLA:

1. A push toward theorizing advancedness through functionally and textually oriented approaches to language analysis;
2. Explicit linking of advancedness to multilingualism;
3. Expansion of research contexts into additional domains;
4. Expansion of the domains of inquiry regarding advancedness;
5. Longitudinal studies investigating diverse aspects of instructed language learning;
6. Research on learning multiple languages to advanced levels;
7. Development of corpora that use to greatest advantage the capacities of corpus-based analysis; and,
8. Further specification of research methodologies suited for investigating the development of L2 abilities.
The present volume provides an overview of the research that has been conducted in the six years since, and takes note of the areas in which more still needs to be done. For example, Malovrh and Moreno (Chapter 11, this volume) note the need for more longitudinal data examining the effects of instructed learning, and Lozano and Callies (Chapter 22, this volume) note the need for more corpus-based analyses. The reader will find chapters that touch on at least one of the above-cited themes cited in Byrnes’s conclusion. What becomes evident to us is that, despite the pioneering work of scholars such as Maxim, Norris, and Ortega, to name just a few, the study of advancedness in SLA is still in its infancy, and to re-emphasize Byrnes’s (2012) assertion, the field of SLA needs, “a sufficiently comprehensive theoretical basis for understanding advanced language … and use” (p. 506). The present volume contributes to the field by compiling (and adding) research to serve as a foundation upon which such an understanding may be developed.

To conclude our introductory chapter, we call on our readers to make note of topics that were excluded from our table of contents but which deserve attention by future scholars if we are to continue to theorize on advancedness in SLA. For example, what can we learn about advancedness through the study of balanced bilingualism? And, what can the study of conversation and interaction tell us about advancedness in situated language use? Furthermore, what other grammatical and morphological features deserve attention and should be studied in future research to add to the profiles of advancedness yielded in the present volume? It is our belief that the importance of such research will only continue to grow, and its demand will increase in proportion to the changing geopolitical landscape of globalization. As the world continues to demand “advanced” multilingual actors, researchers in SLA will continue to be called upon to help us understand precisely what that means, and to more appropriately situate the concept according to theoretical, academic, professional, social, practical, and pedagogical contexts.

REFERENCE

Part I  Advanced L2 Capacity: Orientations on Acquisition
Introduction

The research of L2 proficiency, specifically at the advanced levels of acquisition on which this chapter focuses, finds its basis in the systemic functional linguistics (SFL) theory of language (Halliday & Matthiessen, 2004; Martin & Rose, 2003), as well as its applications in L1 learning and educational settings (e.g. Christie & Derewianka, 2008; Coffin, 2006; Rose & Martin, 2012). In the L2 acquisition field in general and in the English for academic/specific purposes (EAP, ESP), English as a second language (ESL), heritage language (HL), and foreign language (FL) contexts in particular, the interest in SFL as a theoretical approach to advanced proficiency and a basis for pedagogical practice has emerged in response to the relative paucity of knowledge about what advanced language capabilities exactly are (Byrnes, 2012), the lack of longitudinal research that considers advanced levels of acquisition (Ortega & Byrnes, 2008), and the dearth of materials and pedagogical frameworks for fostering and assessing advanced language use.

The chapter addresses how, from the theoretical standpoint, the SFL approach has the potential to illuminate our understanding of advanced L2 proficiency. The next two sections are dedicated to an overview of the major research studies that provide the field with a rich description of advanced L2 use and a review of the scholarship that centers on the SFL-enhanced curriculum construction, pedagogy, and assessment. The chapter ends with suggestions for an SFL-informed research agenda for the future.
SFL-based assumptions about language and language development: Implications for the construct of L2 advancedness

This section briefly reviews five major characteristics of SFL—(i) a functional theory of language, (ii) complementarity of context and text, (iii) complementarity of system and instance, (iv) congruent versus incongruent meaning-making, and (v) a systemic model of language—to reveal the relevance of these features to our understanding of advanced L2 proficiency and development.

Functional theory of language

One of the most foundational assumptions made within SFL has to do with a view of language as a functional tool necessary for making sense of experience or for construing experience into meaning. This assumption has implications with regard to language development from two interrelated perspectives: historically in human society and over a lifespan in individual users. In the SFL view, historically, language has evolved to fulfill fundamental human needs, also referred to as metafunctions (Halliday, 1990): the ideational to reflect on reality, the interpersonal to act on it and relate to others, and the textual to organize these meanings in a coherent and comprehensible way. This functional nature of language is manifested not only in the fact that language is used to achieve certain goals, but most importantly, in its very system in which particular lexicogrammatical resources help realize particular metafunctions (i.e. the grammatical system of Mood realizes the interpersonal metafunction: imperatives as part of Mood are used to make others do what we want).

Within individuals, language develops under the same pressures as it does phylogenetically: the various lexicogrammatical systems are learned not as “a domain of human knowledge” (Halliday, 1993, p. 94), but rather, emerge in the process and under the need for construing experience into meaning and acting on the environment into which one is born. Empirical research by Halliday and colleagues (e.g. Christie & Derewianka, 2008; Derewianka, 1995; Halliday, 1993; Painter, 1999) traces the various stages of development language users go through, as they expand their ability to mean ideationally, interpersonally, and textually. For example, with regard to the ideational metafunction, the trajectory is from learning to construe reality in terms of concrete, then general, and finally, at the advanced stage of schooling, abstract entities.

Just like in the L1, L2 development can be defined first and foremost in terms of the three functional needs mentioned above—as meeting the challenge of knowledge construction, sharing and negotiating with others, and organizing this communication in a comprehensible way, and at the same time grappling with the necessity to use the relevant lexicogrammatical resources to achieve these goals. This overarching definition includes and goes beyond the traditional notions of complexity, accuracy, and fluency as common technical measures for evaluating
L2 ability in that it presents this ability as meaning-making intimately connected with and dependent upon the aspects of context.

**Complementarity of context and text**

A special contribution of SFL to our understanding and assessment of advanced language use is in illuminating precisely what these salient contextual aspects are and how they are construed through wordings. Capturing this *complementarity of context and text*, Figure 2.1 first specifies context in terms of two strata. The context of *culture or genre* refers to culturally valid communication purposes; the context of *situation or register* is constituted by audience and discourse participants or tenor, content of communication or field, and role of language in communication or mode. Second, the figure depicts the relationship between the strata: The contextual layers of genre and register activate the ideational, interpersonal, and textual meanings in the semantic layer that gets realized by specific lexicogrammatical resources through which the context is ultimately construed. The diagram should be imagined as a prism: if we look from above, we see how language changes

![Diagram of complementarity of context and text in the SFL model of language](https://via.placeholder.com/150)

**Figure 2.1** Complementarity of context and text in the SFL model of language, based on Halliday & Matthiessen (2004).
under the influence of contextual features; looking from below, we note how language construes various strands of context (and Systemic Functional Grammar, SFG, further explicates these connections); and looking from within we appreciate the constituents of a given layer (for example, the lexicogrammar as the systems of transitivity, mood, modality, and theme).

Laying out the relationship of stratification between the aspects of context, semantics, and text, this model of language allows one to theorize advanced language capacity not in terms of stable, fixed, and decontextualized grammatical systems (a view from within), but rather as an ability to use dynamic sets of semantic and linguistic patterns to construe culture in particular context-bound instances of verbal action and interaction.

**Complementarity of system and instance**

But how would one connect these distinct instances with learning the system of culture and the system of language? SFL does so by doing away with the strict differentiation typical of traditional approaches to language and language development between system and instance, semantics and pragmatics, and competence and performance, postulating their dialectic complementarity instead, with system and instance on a continuum.

While it would be practically impossible to describe the myriad of instances language users interact in, the notion of instance type as a subpotential that lies between the instance and the system proposed in SFL helps combine the instances into larger categories that the SFL theorist Martin refers to as genres. Originally suggested by Bakhtin (1986), but further developed by Martin and his colleagues (Macken & Rothery, 1991; Martin, 1985; Rothery, 1990), this notion helps describe culture in terms of recurrent goal-oriented social processes construed as oral or written texts with a relatively stable staged generic structure and typical lexicogrammatical resources used to realize it.

The notion of genre helps specify advanced language development as a movement along the line of instantiation toward functional abilities in constructing a variety of text types in a variety of culture-specific registers (Matthiessen, 2006). This view of language development is also supported by the usage-based theories (Bybee, 2008) and research of multicompetence (Hall, Cheng, & Carlson, 2006) that link the frequency and type of language use in various interactive contexts as communities of practice to cognitive restructuring, stabilization, and expansion of communicative repertoires.

**Congruent vs. incongruent meaning-making**

The discussion above has emphasized the ability to participate through language use in a variety of contexts as a defining characteristic of advancedness. SFL helps define advanced proficiency further by clarifying what particular types of contexts and genres that represent them enable and at the same time foster advanced language use. In particular, SFL linguists differentiate among contexts by postulating
a trajectory from those that are characterized by the absence of ideational and interpersonal distance to those that are characterized by its presence (Eggins, 2004), a factor that contributes to conceptual complexity and as a result to complexity in language use. Ideational distance is understood as the distance between language and the social activity one is involved in:

At one end of the continuum, language accompanies social activity and itself is a kind of action, like, for example, in cooking or playing a game of cards. At the other end, there is no other activity but the one in which language constitutes the social process, that of reflection, and this occurs, for example, when one is writing an article. While in the first case language depends largely on the context of situation, it creates this context through text in the second one (Ryshina-Pankova, 2015a, p. 56).

The distance in tenor has to do with the spatial distance between the participants, as well as the distance in their relationships:

On the one end, interaction occurs among several participants, with the possibility of immediate feedback and with interactants knowing each other well. On the other end, monologic communication is addressed at potential non-intimates who are not in close and frequent contact and cannot provide immediate feedback (Ryshina-Pankova, 2015a, p. 56).

While situations of lesser distance are the ones construed by primary discourses that constitute communication with friends and family in the context of mundane and often oral or oral-like everyday interaction, language users experience and construe greater distance as they engage in written or written-like secondary discourses with non-intimates about domain-specific subject matter in institutional settings (Gee, 1998). It is this shift of language use from primary to institutional settings that necessitates development of advanced language capabilities. In other words, advanced L2 proficiency is not only about the socioculturally appropriate language use in context. It can be further described through the concept of literacy or literacies in various secondary discourse contexts. Consequently, the definition of L2 advanced proficiency in connection to the successful functioning in institutional settings converges with that of L1 literacy with its requirements of academic or professional language use (Bhatia, 1993; Cummins & Man Yee-Fun, 2007). In this sense, advanced proficiency is inextricable from writing development in literacy contexts.

The change toward greater interpersonal and ideational distance in secondary discourse interactional contexts is most clearly conceptualized in SFL through the notion of incongruency and the linguistic resource of grammatical metaphor (GM) as its major manifestation. On the one hand, primary discourse communication is construed congruently, with a typical correspondence between language form and language function: propositions are realized by clauses (1: two propositions, two clauses), dynamic actions by verbs (1: got sick, didn’t go), static participants through nouns and pronouns (1: I, the student, he), commands through imperatives (2: close), and modality through modal verbs (3: must) and mood adjuncts (3: certainly).
EXCERPT 1: Ideationally and logically congruent
a. Because I got sick, I didn’t go to school.
b. Because the student prepared for the examination, he wrote an excellent paper.

EXCERPT 2: Interpersonally congruent
Close the window!

EXCERPT 3: Interpersonally congruent
a. It will certainly rain today.
b. It must rain today.

On the other hand, reflecting the greater ideational and interpersonal distance of secondary discourses, incongruent construal disrupts the typical correspondence between grammatical form and grammatical function and represents reality metaphorically, as in Excerpts 4, 5, and 6.

EXCERPT 4: Ideationally and logically incongruent (experiential and logical GM)
a. The student’s excellent preparation resulted in a well-written paper.
b. My sickness prevented me from going to school.

EXCERPT 5: Interpersonally incongruent (interpersonal GM of Mood)
Could you close the window?

EXCERPT 6: Interpersonally incongruent (interpersonal GM of modality)
a. It is clear that it will rain today.
b. I think it will rain today.

Excerpt 4 contains experiential GM: dynamic processes (go to school, prepare for examination) are no longer construed typically through the grammatical class of verbs but through gerund forms that are between verbs and nouns (going to school) and nominalizations (preparation); furthermore, qualities are no longer construed in a typical way through the grammatical class of adjectives but as nominalizations (sickness) that are formed as a result of juxtaposing two meanings, that of a quality (sick) and that of a thing (sickness). The use of ‘prevented’ and ‘resulted’ are instances of logical GM, where causality is realized not by a conjunction that connects two propositions interclausally, but intraclausally by the grammatical class of verbs that is typically used to denote processes, not logical relations. And finally, Excerpts 5 and 6 provide examples of interpersonal GM. (5) presents a metaphor of mood where an interrogative is used to realize the command function typically construed through imperatives. (6) is a metaphor of modality where a degree of certainty typically construed through modal verbs or mood adjuncts within the clause is expressed through an additional projecting clause that is either explicitly objective (It is clear) or explicitly subjective (I think) (Halliday & Matthiessen, 2004).
Such realizations that constitute the metaphorical style of meaning-making that objectifies experience for reflection and presents as explicitly objective or explicitly subjective one’s authorial stance for particular interpersonal effects are considered to be crucial features of advanced language use.

**Systemic model of language**

While traditional approaches to language stress syntagmatic relations that students learn as traditional syntax rules (e.g. Chomsky’s generative grammar), a major defining feature of the SFL theory that has “systemic” in its title is its emphasis on paradigmatic relations, i.e., different options that are available at various strata of the language system (Figure 2.1). For example, in the grammatical system of Mood in the domain of interpersonal meanings, the possible options for relating to one’s interlocutor are: declarative, interrogative, and imperative clauses, with or without a tag. Our choice within this grammatical system depends on the system network or available options at other strata in the model of language, the semantic and the contextual ones. At the discourse-semantic level, do we mean a command or a request? If it is a command, within the situational stratum is it directed toward an equal, a superior, or an inferior, to someone we know well or less well? In each case, a new system network opens, simultaneously enabling and limiting our choices. Within the context of US middle-class culture: Would the use of a direct imperative be more or less typical to encode a command directed toward a child? Is the interrogative how are you? a demand for information that necessitates the use of a declarative in response?

As in the pragmatics-inspired approaches to language development (Bardovi-Harlig, 2013), SFL helps envision the advanced language user as someone who has at her disposal a variety of semiotic resources for making ideational, interpersonal, and textual meanings and has developed a meta-awareness about the significance and communicative effects of these choices within the contexts of the target culture. In this regard, SFL points out the importance of contextually and textually appropriate choice for describing advanced linguistic ability, also noted by other applied linguists, especially those studying acquisition of vocabulary and lexis. For example, already Pawley and Syder (1983) in their seminal article on “native-like selection” and “native-like fluency” stress the importance of selection as a capacity to choose, among a variety of resources and grammatically possible combinations, the typical or unmarked “form–meaning pairings.” In SFL terms, this is the ability to go up and down the system networks at various strata of the context–text continuum.

To conclude, through a functional and textually oriented model of language, SFL offers a theoretical framework that allows one to capture advanced language proficiency and development from multiple intersecting perspectives, as called for, for example, in the work by Cumming (2013): from the perspectives of contexts and purposes of use, intended meanings as particular stances on reality, expression of identity and dialogic relationship with others, and ultimately their construal through language forms. The next section presents some of the specific descriptions of L2 advancedness enabled by the SFL approach.
SFL-based descriptions: Specifying L2 advancedness

This section gives an overview of the findings reported in the SFL-based studies that investigate advanced L2 proficiency as language use within a variety of academic genres and suggests developmental trajectories within advanced literacy contexts where learners have to respond to the challenges of summarizing, reflecting on, arguing about, and interpreting specialized knowledge. Because interaction in secondary discourses is frequently conducted in the written mode, many, but not all, studies (see, for example, the references to oral interviews and presentations in Achugar & Colombi, 2008, and Colombi, 2006) refer to written data. The findings are organized with regard to specific linguistic resources within three major meaning-making functions: ideational, interpersonal, and textual.

Advanced ideational resources: Construing knowledge through experiential and logical GM

A resource crucial for meaning-making in secondary discourse contexts identified in SFG is that of ideational GM (see Excerpt 4). Its role in construing advanced literacy genres has been described as contributing to their abstractness, technicality, density, objectivity, authoritative voice, cohesive and coherent structuring, and academic reasoning in English and other languages (e.g. Drury, 1991; Halliday & Martin, 1993; Magnusson, 2013; Ravelli, 2003; Schleppegrell, 2004a; Yang, 2011).

In the research on advanced L2 development, two types of studies of GM can be distinguished: those that offer a developmental outlook and trace the emergence of this feature longitudinally or cross-sectionally and those that investigate and contrast the deployment of GM across a sample of high-versus low-rated texts.

Overall, similar to the findings in research on L1 academic literacy development (Derewianka, 1995; Magnusson, 2013; Painter, 2003), longitudinal studies from a variety of L2 instructional contexts report an increase from lower to higher levels of instruction in the frequency and variation of experiential GM. Colombi (2002, 2006) and Achugar and Colombi (2008) analyzed written and oral expository texts by heritage speakers of Spanish created over a period of nine months. Presenting the results of three different case studies, they point to an increase in lexical density due to the use of nominalizations as the most prominent manifestation of GM, as well as logical metaphor (as in Excerpt 4). What is significant about these studies is that they juxtapose syntactic complexity (for example, as the amount of subordination), a frequent measure of assessing L2 development (Norris & Ortega, 2009), with lexical density enabled by the use of GM and propose the latter to be a distinguishing feature of advancedness. Similar findings of the increase in lexical density and decrease in syntactic complexity are also reported in the quasi-longitudinal study of German-learner writing across four curricular levels by Byrnes, Maxim, and Norris (2010). The surprising decrease in syntactic complexity at the highest levels of the curriculum is attributed to a different mechanism of connecting meanings within the clause (e.g. Excerpt 4) rather than across clauses (e.g. Excerpt 1) enabled through the use of GM.
Focusing solely on nominalizations, Byrnes (2009) traced the development of GM in the texts of three genres (narrative, journalistic article, and public speech) by 14 learners of German that were written over three consecutive levels (intermediate, early advanced, and advanced) of a well-articulated curriculum. Byrnes demonstrated that the use of GM more than tripled between early advanced and advanced levels, even if GM deployment was not always felicitous. Using the data from the same program, Ryshina-Pankova’s (2010) quasi-longitudinal study of book reviews across early advanced, advanced, and high advanced curricular levels revealed not only a gradual increase in the use of GM (as nominalizations) with each level but also the difference in their type. Higher levels were characterized by a decrease in the use of faded metaphors as expressions that lost their metaphorical meaning or their connection to the process (e.g. relationship) and an increase in “fresh” metaphors through compound nominalizations (e.g. reconciliation attempts). An increase in the use of GM as nominalizations is also reported within a shorter longitudinal span of one semester by Yasuda (2015), who investigated the development of linguistic resources by Japanese adult learners of English for the genre of summary writing in biology.

While the longitudinal L2 studies record an overall increase in the use of GM from intermediate to advanced levels of instruction or within a course, L2 investigations that focus on samples of L2 advanced-level texts demonstrate that once students reach a particular threshold it is not so much the frequency of GM as its particular use that helps capture the quality of texts and differentiate between the high- and low-rated ones. Noting the inconclusive results of the quantitative analysis of GM in German-learner journalistic articles for distinguishing between the high- and low-rated texts, Ryshina-Pankova and Byrnes (2013) propose that it is the deployment of GM (nominalizations) as a tool for conceptual refiguration and textual configuration of content that impacts the quality of texts. They define conceptual refiguration as the use of GM to reinterpret, evaluate, and theorize about experience by generalizing discrete occurrences into abstract notions, and textual configuration as a way to relate GMs both to the congruent realizations and to other GMs by positioning them in strategic places at the level of the text, paragraph, and clause.

A similar conclusion is reached by Liardet (2013, 2016) in two separate studies of expository texts by university Chinese learners of English. Expanding the focus on GM from nominalizations to other GM types (e.g. process as quality: infect → infectious), she argues that it is not so much the frequency in the use of GM by students who are already highly proficient in L2, but the GM’s textual impact that distinguishes success. Liardet proposes to evaluate the textual impact of GM by analyzing its role in anaphoric reconstrual, nominal group elaboration, and cause and effect metaphorical networks. For example, the use of GM for anaphoric reconstrual, a notion similar to that of textual configuration in Ryshina-Pankova and Byrnes (2013), helps learners create cohesion in discourse and move between different levels of meanings in text, from specific to general, concrete to abstract, congruent to incongruent and vice versa (cf. Ryshina-Pankova, 2010).
Advanced interpersonal resources: Expressing stance and constructing relationship between discourse participants through the use of interpersonal GM and Engagement

One of the key requirements in constructing secondary discourse genres has to do with an ability to build an intersubjective relationship with the audience and project a subtle stance that is both authoritative and inclusive of other alternative opinions. In SFL, the realization of these interpersonal meanings has been studied through the concept of interpersonal GM and a discourse semantic system of Appraisal, and particularly Engagement (Martin & White, 2005). While a variety of studies adopt this perspective to explore the interpersonal aspect of advanced literacy in L1 (e.g. Derewianka, 2009; Hood, 2010; Lancaster 2014; Wu, 2006), there are relatively few investigations in the L2 field reviewed below.

With regard to interpersonal GM, Schleppegrell (2004b), in her contrastive study of L1 and L2 laboratory reports, discusses the appropriate use of the GM of modality in academic writing (as in Excerpt 6). In particular, she points out the importance of the objective interpersonal GM (Halliday & Matthiessen, 2004) where the degree of probability or obligation is realized objectively through a clause complex, as in Excerpt 7.

EXCERPT 7: Explicitly objective interpersonal GM
a. It is obvious that these results are in error vs. These results must be in error.
b. It is not possible to draw any firm conclusions about this from the data vs. I cannot draw any firm conclusions about this from the data.

She comments that the objective interpersonal GM helps conceal the writer as the source of the belief and “construe the evaluation as fact, rather than opinion” (p. 183) contributing to the objective presentation of an argument. In Schleppegrell’s analysis, it is the proficient L1 writers who construe their evaluative stance by drawing on this resource. The development of objective presentations is also noted by Colombi (2006) in the expository texts and interviews by the heritage learners of Spanish.

At the same time, the use of the subjective interpersonal GM, also realized through a clause (as in Excerpts 6b and 8) is identified by Schleppegrell (2004b) as a less preferred option in academic writing that is, however, overused in L2 writer texts contributing to their lack of authoritativeness.

EXCERPT 8
I believe that these results are in error.

This observation is further supported by Liardet’s (2015) study that points to the reliance of the university Chinese learners of English on subjective interpersonal GM in their written essays and the resulting sense of tentativeness and subjective reasoning in their texts.
While the use of the subjective interpersonal GM might not be an appropriate resource in academic written discourse, it is reported as a marker of development in oral interviews on bilingualism with advanced learners of Spanish by Colombi (2006) and Achugar and Colombi (2008). These researchers present the results for one of the participants in the study who with time started using the subjective interpersonal GM as a resource to “present a subjective evaluation as an authorized speaker” (Achugar & Colombi, 2008, p. 53). Colombi (2006) concludes that it is a clear awareness about the effect of the objective and subjective interpersonal GM, as well as the ability to realize these GMs through a variety of syntactic and lexical structures, like projecting mental clauses (e.g. *I think, I believe*) or a predicated theme (*it is possible to argue, it is clear*), that define an advanced L2 user.

Another perspective on the construal of interactive meanings in advanced literacy contexts is enabled by the analyses that use the system of Engagement (Martin & White, 2005) that describes a variety of semantic choices and their possible linguistic realizations deployed to express stance and negotiate with others in academic argumentation. For example, the system differentiates between *bare assertions* or *monoglossic* statements (Excerpt 9) that do not explicitly refer to the voices of others and *heteroglossic* statements (Excerpt 10) that explicitly engage in dialogue with alternative views.

**EXCERPT 9**: Bare assertion, monoglossic
This film is an excellent choice for moviegoers of all ages.

**EXCERPT 10**: Heteroglossic
It seems like the film is an excellent choice for moviegoers of all ages. (One possibility among others.)

Within the heteroglossic option, language users can deploy expanding strategies that open the dialogic space to acknowledge other perspectives, as in *entertain: attribute* (Excerpt 11), as well as contracting strategies that close the dialogic space by rejecting the opinions of others, as in *disclaim: counter* (Excerpt 12), or by proclaiming one’s own, as in *proclaim: pronounce* (Excerpt 13).

**EXCERPT 11**: Entertain: attribute
According to the *New York Times* film critic, the film is an excellent choice for moviegoers of all ages. (Acknowledges one expert opinion among others.)

**EXCERPT 12**: Disclaim: counter
Although the storyline in the film is somewhat stereotypic, it is still an excellent choice for moviegoers of all ages. (Invokes a contrary position that is not rejected directly.)

**EXCERPT 13**: Proclaim: pronounce
I am convinced the film is an excellent choice for moviegoers of all ages. (Strong level of writer commitment.)
While negotiation with the audience has been studied in expert and advanced learner texts by non-SFL researchers through a focus on particular discrete linguistic markers of stance (Chafe & Nichols, 1986; Hyland, 2005; Palmer, 1986), the SFL-based Engagement stands out as offering a comprehensive meaning-oriented analytical approach to identifying strategies of argumentation and systematically connecting them to a variety of linguistic resources.

While there is a lack of longitudinal L2 studies focusing on the development of Engagement resources, those that do use this framework present a contrastive view of effective and less effective strategies. For example, Wu’s (2007) comparison of geography essays by L2 university students reveals a greater use of heteroglossic and expanding strategies (like supporting one’s claims by reference to other researchers through entertain: attribute) in high-rated essays versus higher use of bare assertions and contractive options (like proclaim: pronounce) in the low-rated texts. This over reliance on proclaim resources that may render the texts overly assertive or imposing and thus less persuasive is also found in Ho’s (2011) contrastive investigation of argumentative texts by Vietnamese learners of English. While more successful writers do make use of contractive options, Wu notes that they differ from the ones in the low-rated texts. For example, rather than proclaim: pronounce, they employ proclaim: endorse (i.e. this study demonstrates) to show alignment with other sources and disclaim: counter to explicitly include an alternative point of view.

Further enhancing our understanding of successful management of dialogic negotiation in writing by advanced L2 users, Lancaster’s (2011) study of economic policy papers by high-rated L1 and lower-rated L2 upper level undergraduate students reveals additional patterns of successful versus less successful use of Engagement. While high-rated papers were characterized by the alternation of contracting and expanding strategies at the paragraph level and the use of rhetorical pairs (concede–counter: it is true ... but) that helped “establish solidarity with a disagreeing reader” and problematize an issue, low-rated essays by L2 writers were either too assertive by employing more contracting moves or didn’t structure the alternation between expansive and contractive moves in a textually clear way. Lancaster concludes that it is both the dialogic balance and the dialogic control as the appropriate texturing of Engagement devices that are two major factors contributing to successful construal of interpersonal meanings in advanced literacy texts.

**Advanced textual resources: Organizing meanings in text through thematic choices**

The third area that defines advanced L2 capacity from the SFL-perspective concerns the ability for control over rhetorical structuring of ideational and interpersonal meanings in secondary genres. In SFL, a major resource that helps organize texts rhetorically is that of theme. Defined as the “element in a clause or larger unit of text which comes first” (Coffin & Hewings, 2004, p. 156), theme creates internal cohesion by operating as a textual connector between the previous and the following discourse (e.g. Mauranen, 1996; Schneider & Connor, 1990). At the same time, through particular patterns of theme selection and progression, theme...
foregrounds certain ideational and interpersonal meanings, so as to achieve the communicative goals of certain genres, thus acting as an instrument for contextual generic coherence (e.g. Francis, 1989; Ghadessy, 1995).

In the analysis of a second language, theme has been studied both developmentally and contrastively in a sample of L2 learner texts. Developmentally, Ryshina-Pankova’s cross-sectional studies of theme in L2 book reviews (2006, 2010, 2011) demonstrate distinct changes in the use of theme across three increasingly advanced consecutive curricular levels. For example, Ryshina-Pankova (2006) reports an increase toward higher curricular levels in the use of complex theme (defined as a theme with more than three lexical elements: e.g. *the most important book of this genre*) and its structural variety (from using adjectives and participles to prepositional phrases, appositions, and embeddings through relative clauses). Significant in this finding is not simply the increase in the use of lexically complex themes, but rather the way they are deployed: they help fulfill the goals of the content summary and motivation and evaluation moves of the book review genre. For example, by using lexically complex themes in the content summary move writers are able to include a large amount of detail about the story line, the book itself, or its author, in a way that makes their summaries informative but does not detract from keeping the focus on the main occurrences in the rendering of the plot (e.g. *the narrator who at the beginning of the story was hit by a car and lies dying in the street*).

A shift in the selection of what gets to be the theme toward advanced instructional levels is also observed in Ryshina-Pankova’s 2010 study. The use of experiential GM in the textually strategic position of theme is more characteristic of book reviews by more advanced and native speaker writers. In their texts, GM in theme not only anaphorically condenses the previous discourse, but also pushes the discourse forward in line with the communicative goals of the book review genre by presenting an interpretation, comment, and evaluation of the book within the GM itself (*criticism*) or cataphorically enabling these evaluations in the subsequent structure of the clause (*the tragedy of history will affect every reader*). Overall, thematized experiential GM constructs a discourse of logical reasoning and argumentation in the genre where persuasion is an important communicative goal.

Finally, Ryshina-Pankova’s (2011) study examines the thematization of interpersonal elements in book reviews across the three curricular levels. The study demonstrates how more advanced writers “move away from direct expression of authorial opinion manifested in thematization of the writer of the review and employ a more intersubjective and thus more persuasive reader orientation in their texts” (p. 243). The foregrounding through theme of subjective interpersonal meanings as a characteristic of beginner writer discourse is also noted in Coffin and Hewings’s (2004) investigation of L2 writers’ IELTS (International English Language Testing System) exams. These researchers conclude that the argumentative genre of the exam is construed by the L2 writers through thematization of proclaim: pronounce resources (*I think, I believe, I hope*), rather than the argumentative process itself. This high level of commitment to one’s argument leads to the sense of inappropriate assertiveness and interferes with the communicative goal of persuasive genres.
To conclude this section on the SFL-based description of the features of L2 advanced proficiency, one can note the following: The L2 studies reviewed above define advanced proficiency in terms of the ability to engage in incongruent meaning-making and produce secondary discourse genres in academic and institutional settings. In particular, SFL provides researchers with tools for identifying major linguistic features that are crucial for this production. In contrast to other approaches, identification of these linguistic elements prioritizes a semantic and contextual approach. This enables one to capture a greater variety of linguistic forms essential for construing the academic genres (e.g. GM is realized not only through nominalizations but also through causative verbs) and understand their functions in discourse.

Thus, with regard to the ideational meanings, the L2 studies demonstrate the crucial importance of ideational GM in its various manifestations as a powerful resource for meaning-making in secondary discourses. On the one hand, these studies report a gradual increase in the use of ideational GM toward higher instructional levels; on the other, they elucidate precisely how GM helps construe ideational and textual meanings valued in academic and professional registers. Furthermore, exploring the construal of interpersonal meanings in learner discourse, the studies trace the path toward more intersubjective and objective relationships with the audience manifested in the use of objective interpersonal GM and specific Engagement resources that help create a balance between expanding and contracting dialogic strategies and contribute to the persuasiveness of arguments. Finally, drawing on the significance of theme in organizing ideational and interpersonal meanings in advanced literacy texts, the studies of theme in L2 discourse reveal an increase in selection of particular types of themes toward higher proficiency levels. In particular, the developmental trajectory seems to be toward the lexically complex, lexically dense (through ideational GM), and interpersonally intersubjective elements that contribute to the creation of new textual worlds and the ‘crystalline’ organization of written discourse.

Overall, the most significant contribution of these studies of L2 development and use appears to be in revealing how the identified advanced literacy resources are not only linguistic and textual features but also crucial tools for knowledge construction and interpretation within secondary discourse settings. Corroborating this idea, Colombi states: “From an SFL perspective, developing knowledge and understanding of the content area and developing control of the linguistic resources that construct and communicate that knowledge and understanding are essentially the same thing” (2006, p. 160).

SFL-inspired curriculum, pedagogy, and assessment: Fostering advanced L2 proficiency

The SFL-based multidimensional approach to the analysis of text in context, as well as the applications of the theory in L1 education, have informed course and curriculum construction, pedagogy, and assessment that aim to develop advanced
proficiency in L2s in general and advanced L2 writing abilities in particular. An extraordinarily useful concept that has been adopted and adapted for the purposes of L2 education, especially for developing materials and tasks across various instructional contexts and thematic areas at the advanced levels, is the SFL notion of genre and generic stages (e.g. Crane, Liamkina, Maxim, & Ryshina-Pankova, 2005; Donohue, 2012; Pang, 2002; Weigert, 2004).

Within advanced L2 instruction explicitly oriented toward teaching academic writing, successful utilization of the SFL is reported, for example, in Devrim (2013), a study that focuses on the importance of GM as a tool for online tutoring; in Yasuda (2015), who describes an instructional sequence for teaching the generic stages as well as the resource of GM in the genre of summary in biology; in Shum, Tai, and Shi’s (2016) account of the effectiveness of the SFL-based Reading to Learn pedagogy (Rose & Martin, 2012) in teaching the genre of academic discussion to advanced L2 writers; and in Chang’s (2010) investigation that revealed the effectiveness of introducing L2 writers to the meta-language of the Engagement framework for raising their awareness about the appropriate realization of authorial voice in academic articles.

With regard to curricular scaffolding as establishment of a sequence of courses oriented toward advanced L2 proficiency in all modalities (speaking, reading, writing, and listening), the most extensive application of the SFL has been achieved by the German program at Georgetown University (but also see Maxim, Höyng, Lancaster, Schaumann, & Aue, 2013; and Yasuda, 2011). Drawing on the work in L1 literacy contexts (Coffin, 2006), this content- and language-integrated text-based curriculum is built around a sequence of culturally relevant genres to enable content learning, foster the gradual development of control over a variety of text types as reflecting different cultural contexts, and address the development of lexicogrammatical resources that make up these genres. In order to enable the shift toward L2 advancedness, the curriculum is based around four trajectories (Byrnes et al., 2010; Crane, 2006; Ryshina-Pankova, 2013): (i) from overt to covert dialogicality; (ii) from congruent to incongruent language use; (iii) from oral-like to written-like language use; and (iv) from personal stories to historical recounting and accounting to explanations, comparisons, and contrasts to argumentation about public issues in the German society. Significantly, these trajectories are translated into the statements about specific lexicogrammatical features that are the focus at a particular point in the curriculum (for example, the shift from interclausal to intraclausal logical connections at the advanced level). Extensive research on the program demonstrates various positive outcomes with regard to advanced L2 development in terms of syntactic ability (Byrnes et al., 2010, chapter 9), coherence and cohesion (Ryshina-Pankova, 2006), or reaching the humanities content learning goals (Ryshina-Pankova, 2015b).

Pedagogically, SFL-based approaches place value on explicit instruction through the Teaching and Learning Cycle (Rothery & Stenglin, 1995) that draws on the Sociocultural Theory of learning (Vygotsky, 1978). In this cycle, genres typical of a particular content or discipline area are taught through deconstruction as exploration of the context and textual modeling, joint construction as textual scaffolding, and finally, independent construction. While there is a dearth of studies that look at
the specifics of teacher discourse in SFL-based instruction, the modeling stage as a minutely detailed analysis of teacher talk is presented, for example, in the Reading to Learn, Learning to Write framework (Rose & Martin, 2012). For the scaffolding stage, Mohan and Becket (2001) describe an approach to teacher feedback as in recasts that focuses not on formal error correction but rather on reformulation of student contributions through the use of GM as a way to expand learner meaning-making potential within an ESL content-based class. The pedagogical sequence over an entire Teaching and Learning Cycle is also presented for the genre of political appeal in Byrnes et al. (2010).

With regard to the independent production stage, one of the central concepts that is instrumental in implementing it is that of genre-based tasks. This notion helps connect in a systematic way model texts with the student-produced ones (e.g. a cooking talk show is based on a recipe) and context and content with language features (Byrnes, Crane, Maxim, & Sprang, 2006). Genre-based tasks serve as a link not only between pedagogy and assessment of learning outcomes, in terms of both content and language learning, within a particular instructional unit or course, but importantly in the program as a whole: the so-called prototypical genre-based tasks that follow the aforementioned trajectories (p. 24) across the curriculum serve as instruments for overarching performance assessment at its various focal points (for example, at the end of the language requirement sequence).

The scholarship on the application of SFL in L2 educational settings demonstrates how this theoretical framework has the potential to contribute to advanced L2 instruction by overcoming at least these three challenges: (i) through the construct of genre and genre-based tasks, to circumvent the limitations of the L2 classroom and create an environment where learners are motivated to use language to achieve real communicative purposes and thus imagine themselves as part of the target culture community and where this use is supported through modeling and scaffolding; (ii) with its focus on academic or professional literacy as a crucial aspect of advanced language use, to provide guidance as to precisely how to see language learning as content learning and language production as knowledge construction, thus helping address the pressing need to redefine L2 education in terms of its major contribution to achieving the goals of (higher) education; and finally, (iii) through its well-defined conceptualization of changes in language use from primary to secondary discourse contexts, to envision developmental trajectories and the corresponding selection and sequencing of the pedagogical materials and tasks that would help learners to indeed progress.

**Outlook: An agenda for the SFL-informed research on L2 advanced proficiency**

The overview of research presented in the chapter demonstrates the great contribution SFL can make to the conceptualization, description, and fostering of advanced L2 proficiency. The following directions for future research may further enhance our theoretical and practical insights in this area:
• Research of advanced L2 performance in a variety of languages beyond English, a variety of genres and modes, including oral and multimodal genres, and a variety of contexts, including, for example, distance learning;
• Construction of curricular frameworks that utilize SFL-based insights, support long-term development toward L2 advancedness, and enable ecologically valid longitudinal and cross-sectional research within such programs;
• Longitudinal studies that further specify major developmental stages with regard to learning how to construe ideational, interpersonal, and textual meanings;
• Development of advanced learner corpora and methodologies that enable one to capture the distribution and significance of SFL-identified features in advanced language use;
• Detailed analyses of L2 classroom discourse that utilize SFL tools and focus on the semantic-linguistic links in teacher-student interaction;
• Assessment research of the SFL-informed programs that demonstrates the impact of the approach on learner outcomes.

NOTES

1 In SFL, any language production, independent of its mode, spoken or written, that is meaningful and has a social purpose is a text (cf. Christie, 1999).
2 Lexical density is defined as the number of content-carrying lexical items (nouns, adjectives, adverbs, and verbs) as a proportion of the total number of words in the text.

REFERENCES


Liardét, C. L. (2015). “As we all know”: Decoding Chinese EFL learners’ use of interpersonal grammatical metaphor. Presentation at the International Systemic Functional Congress (ISFC), Aachen, Germany.


3 Psycholinguistic Approaches and Advanced Proficiency

LEAH ROBERTS
University of York, UK

Introduction

Second language research from a psycholinguistic perspective is often concerned with the extent to which learners’ processing is native-like. This is particularly the case for researchers interested in grammatical processing, because it is L2 grammatical knowledge that lies at the heart of much theoretical debate in the literature. In fact, in contrast to most L2 acquisition studies (Plonsky, 2013), sentence processing studies have focused almost exclusively on learners characterized as ‘advanced late-learners’ (i.e. not balanced bilinguals). This is because the aim of such research often has been to examine whether or not L2 processing is fundamentally similar to or different from L1 processing (Clahsen & Felser, 2006; Cunnings, 2017; Hopp, 2010). This mirrors the debate in the L2 acquisition field surrounding the question of the extent to which L2 learners who began learning the language post-childhood can acquire the grammar to native-like levels (see Slabakova, 2009). In sum, theories of L2 sentence processing attempt to account for a great number of findings that there are L1/L2 differences, even at highly advanced levels of proficiency.

Most L2 sentence processing studies can be divided into two types, those whose focus is on processing/parsing routines (e.g. the moment-by-moment interpretation of ambiguous or structurally complex sentences, and use of grammatical information during parsing), and those assessing learners’ on-line sensitivity to (un)grammaticalities in the target language, often in the area of agreement processing. Psycholinguists working in the former area (e.g. Clahsen & Felser, 2006; Grüter, Rohde, & Schafer, 2014; Jackson & Roberts, 2010; Papadopoulou, 2006; Roberts & Liszka, 2013) have developed models of L2 sentence processing to account for observed L1/L2 differences at advanced proficiency (e.g. the Shallow
Psycholinguistic Approaches and Advanced Proficiency

Structure Hypothesis, SSH, Clahsen & Felser, 2006; Reduced Ability to Generate Expectations, RAGE, Grüter et al., 2014). Those focusing on learners’ sensitivity to (un)grammaticalities employ psycholinguistic methods, and contrast on-line data with off-line grammaticality judgments (e.g. Hopp, 2010; Juffs & Harrington, 1995, 1996; White, 1987, 1989, 2003). Unlike in the L2 parsing literature, these latter studies tend to examine the processing of L2 learners of different proficiency levels (as well as other potential factors such as L1 background), data from which can inform formal theories of SLA (e.g. Full Exposure/Full Access, Schwartz & Sprouse, 1996).

All psycholinguistic approaches rely on time-sensitive measures, for instance a button-push in self-paced reading/listening, or the length of time spent reading a word in one condition versus another during eye-tracking. The basic assumption underlying this approach is that slower processing speed reflects some kind of (comparative) computational effort and thus difficulty in interpreting the input. In the next sections, I set out research into investigating L2 learners’ parsing strategies (including incremental parsing and re-analysis) and their application of grammatical knowledge and sensitivity to agreement violations during real-time processing, and I offer a picture of the nature of L2 sentence processing at advanced levels of proficiency.

L2 parsing and advanced proficiency

As noted above, given the theoretical interest in end-state attainment, many L2 parsing researchers focus on the extent to which L2 learners’ parsing is native-like, and have therefore often investigated only highly advanced L2 learners. These learners have often been independently matched for proficiency and, therefore, the focus is on the question of whether once L2 grammatical knowledge is established off-line, it can be applied in real-time processing (Felser, Roberts, Marinis, & Gross, 2003; Frenck-Mestre & Pynte, 1997; Papadopoulou & Clahsen, 2003; Rah & Adone, 2010; Roberts & Felser, 2011; Witzel, Witzel, & Nicol, 2012). Taking their cue from monolingual psycholinguistic research, researchers ask whether L2 learners incrementally process the input (attempting the analysis of each word as it enters the current parse), and where misanalyses take place, whether learners can revise their interpretations in real time. The phenomena under investigation include ambiguous constructions (garden-paths), syntactically complex constructions (dependencies and referential processing), pronouns and structural attachment preferences (e.g. relative clause attachment), the processing of which have provided data informing various models of L1 parsing (e.g. The Garden Path theory, Frazier, 1987).

Incremental processing

It is uncontroversial that readers and listeners incrementally process the input in real time. That is, they do not wait until the end of a sentence or utterance before they begin to attempt an interpretation. A major question in L2 parsing research is
whether L2 learners incrementally process the input to the same extent as native speakers, putting to use grammatical and lexical-semantic knowledge in real time, immediately each word enters the parse. Evidence of incremental processing can be found when comprehenders are shown to misanalyze the input, for instance during the processing of temporarily ambiguous, or ‘garden-path’ sentences (1). There is much evidence of L2 incremental parsing using this type of construction. For instance, in a word-by-word grammaticality judgment task, Juffs and Harrington’s (1995, 1996) advanced Chinese-English L2 learners slowed down when reading the disambiguating verbs *(proved/cheated)* in pre-posed adjunct (1a) and complement (1b) clause garden-path sentences, in comparison to non-garden-path constructions with intransitive verbs (e.g. *arrive*) (1c).

(1) a. After Bill drank the water proved to be poisoned.
   b. Sam warned the student cheated on the exam.
   c. After Sam arrived the guests began to eat.

Just like native speakers, the learners therefore showed evidence of initially misinterpreting the ambiguous NPs (*the water, the student*) as direct objects in the garden-path sentences (see also French-Mestre & Pynte, 1997; Juffs, 2004), and given that there were no slow-down effects in the non-garden-path items, were clearly sensitive on-line to the subcategorization properties of the verbs used. This type of finding has been observed in a number of studies with learners ranging from intermediate to advanced levels of proficiency (Dussias & Cramer Scaltz, 2008; French-Mestre & Pynte, 1997; Hopp, 2015; Jackson & Roberts, 2010; Jegerski, 2012; Juffs, 2004, 2005, 2006; Juffs & Harrington, 1995, 1996; Rah & Adone, 2010; Roberts & Felser, 2011). Evidence that learners can rapidly make use of subcategorization information and biases of L2 verbs to constrain initial parsing decisions was also found in Dussias and Cramer Scaltz (2008). Spanish-English L2 learners’ processing was slower following disambiguation when the input went against the verb’s bias for either a direct object (2a) or a sentence complement (2b), thus processing the items like native speakers. This was the case except where they encountered verbs which differed in biases between Spanish and English, in which transfer from the L1 bias was observed.

(2) a. The CIA director confirmed the rumor could mean a security leak.
   b. The ticket agent admitted the mistake when he got caught.

Although L2 learners can incrementally process sentences in real time, akin to native speakers and as long as they have the requisite knowledge, some research has shown that even learners of very advanced levels of proficiency may have trouble recovering from their misanalyses on-line, particularly for more syntactically complex sentences (Jacob & Felser, 2016; Roberts & Felser, 2011). These studies have used semantic information such as pragmatic plausibility to investigate on-line recovery from misanalysis (or re-analysis, or re-ranking of favored analyses) in such temporary ambiguities (cf. Pickering & Traxler, 1998). For instance, Roberts
Psycholinguistic Approaches and Advanced Proficiency

and Felser (2011) reported that during self-paced reading, advanced Greek L2 learners of English were garden-pathed in both complement clause sentences (3a) and those containing pre-posed adjuncts (3b), with plausible ambiguous NPs (the book/the song) leading to higher reading times in comparison to the implausible ambiguous NPs (the girl/the beer). Thus as with earlier garden-path studies, the subcategorization properties of the preceding verbs influenced the extent to which the learners analyzed the ambiguous NPs as direct objects (read the book/ played the song).

(3) a. The journalist wrote the book (the girl) had amazed all the judges.
   b. While the band played the song (the beer) pleased all the customers.

However, it was only with the less complex complement clause constructions that there was evidence that the readers had recovered from their initial misanalysis. Specifically, recovery was in evidence in the reading times because on and beyond the disambiguating complement clause verb (had amazed in 3a), the reading time patterns reversed, with processing difficulty observed for the plausible condition. Therefore, like native speakers, this is evidence that the learners’ commitment to their original erroneous plausible analysis (wrote the book) is stronger than their commitment to the erroneous implausible one (wrote the girl) and was therefore more difficult to abandon in the face of new evidence. Furthermore, both plausible and implausible complement clause conditions elicited equally high scores on the comprehension questions that appeared at the end of each item (over 90%). Thus taken together, the learners appeared to have fully recovered from their misanalysis of these temporarily ambiguous sentences. This was not the case for the more structurally complex pre-posed adjuncts. No such reversal in reading time pattern was observed. That is, although they were again very sensitive to the plausibility of the ambiguous NP (song/beer) as potential direct object of the preceding verb (played), following disambiguation the implausible condition (played the beer) continued elicit a higher processing cost, right up until the end of the sentence. The authors argued that recovery from the initial misanalysis was achieved in the pre-posed adjunct conditions, but this occurred not during on-line processing as was the case for the structurally less complex complement clause sentences, but at a later stage, given that accuracy was significantly higher following the implausible (95%) compared to the plausible pre-posed adjunct conditions (85%).

Similar evidence that learners may have trouble recovering from initial misanalyses also comes from studies on the processing of filler-gap constructions like wh-questions. Native speakers initially attempt to integrate a fronted wh-constituent (the filler) with the first potential subcategorizer (or to posit a gap at the first available site) (e.g. Boland, Tanenhaus, Garnsey, & Carlson, 1995), but again, the plausibility of the resulting sentence fragment affects the relative cost of recovery (e.g. Stowe, 1989), and this too has been observed for L2 learners, even when wh-constructions are formed differently in the learners’ L1 (e.g. Williams, Möbius, & Kim, 2001). However, Jackson and van Hell (2011) found that their intermediate learners did not show evidence of fully recovering from misanalysis
in processing subject extractions. Similar results have been reported for lower working memory L2 learners (Dussias & Piñar, 2010; Williams, 2006).

The findings of the above studies show that learners can incrementally process the input and that their on-line parsing decisions can be seen to be constrained in ways akin to native speakers, but that L1–L2 differences at the semantic level can influence this process (see also Frenck-Mestre & Pynte, 1997). However, even for highly advanced L2 learners, recovery from misanalysis may be problematic when this involves structurally more complex input (see also Jacob & Felser, 2016).

Investigating incremental processing in verb final languages such as German, allows for an examination of the extent to which learners can make use of morphosyntactic information such as case-marking to build up the structure of a sentence in real time before encountering lexical information provided by the verb. L2 learners appear able to do so, but seemingly only at advanced levels of proficiency. Such proficiency differences were reported in Jackson (2008), who investigated whether English-German L2 learners would show the same processing preference for subject over object extractions in German single-clause _wh_-questions (4) as has been reported for native speakers (Bader & Meng, 1999; Schlesewsky, Fanselow, Kliegl, & Krems, 2000), or whether their L1 (English) object-extraction preference would be observed. Jackson tested items where the lexical verb appeared early on (past simple, _traf_, 5a) or later, at the end of the sentence (present perfect, _hat…getroffen_, 5b).

(4) a. Past simple
   Welche Ingenieurin traf den/der Chemiker gestern Nachmittag im Café?
   Which-NOM/ACC engineer met the-ACC chemist/the-ACC chemist yesterday afternoon in the café?
   ‘Which engineer met the chemist/did the chemist meet yesterday in the café?’

b. Present perfect
   Welche Ingenieurin hat den/der Chemiker gestern Nachmittag getroffen?
   Which-NOM/ACC engineer has the-ACC chemist/the-ACC chemist yesterday afternoon in the café met?
   ‘Which engineer met the chemist/did the chemist meet yesterday in the café?’

All items were fully ambiguous until the second NP was encountered, which disambiguated the constructions via case-marking toward either a subject-first ( _den Chemiker_ ) or object-first ( _der Chemiker_ ) _wh_-sentence. The advanced proficiency group performed like the native speakers of German. That is, they showed a processing cost for the critical NP when it forced an object-first interpretation via nominative case ( _der Chemiker_ ). Thus the highly proficient group could make use of morphosyntactic information in building up the _wh_-items on-line, just like native speakers. Interestingly, this was also the case for the less proficient group, but only for the past simple conditions (5a). Specifically, a processing cost was found for dispreferred object-first items if a thematic verb had been processed
prior to disambiguation (Welche Ingenieurin traf den/der Chemiker). This processing cost was not visible for the present perfect conditions (5b) until the end of the sentence, at the point at which they processed the thematic verb (see also Hopp, 2015). In sum, these findings suggest that in some cases, at lower levels of proficiency, L2 processing may in fact not be fully incremental, even in simple, single-clause constructions, in the absence of lexical-thematic information at the verb.

As well as proficiency affecting the extent to which native-like incremental processing is observed in on-line studies, the specific tasks that are undertaken may also influence on-line processing. That is, native-like parsing may be more likely when participants’ attention is directed toward the experimental manipulation via secondary truth-value or judgment tasks which are undertaken at the same time as learners are reading in real time. This can be seen in the comparison of results of two studies, one in which learners judged the acceptability of the experimental items, during their on-line reading (Jackson & Dussias, 2009), and the other in which identical materials were used, but with no supplementary task (Jackson & Bobb, 2009). In their study, Jackson and Dussias (2009) reported that only their highly proficient English learners of German processed unambiguous who-extractions (5) in a native-like way with increased reading times on the matrix verb (denkst) and the following NP (du) for the subject versus the object extraction items, demonstrating a German native-like subject- versus object-first preference (e.g. Felser, Clahsen, & Münte, 2003; Fiebach, Schlesewsky, & Friederici, 2002).

(5) Wer (Wen) denkst du, bewunderte den Sportler nach dem Spiel?
Who-NOM (Who-ACC) think you, admired the athlete after the game
‘Who do you think missed the teacher/whom do you think the teacher missed after the game?’

In the parallel study whose participants were required only to read for comprehension (Jackson & Bobb, 2009), no processing time difference was observed between the two conditions, and also, there were no effects of individual differences of either proficiency or working memory. It is possible, given the differences between these two studies, that the lower proficient L2 learners in the earlier Jackson (2008) study may have incrementally processed the experimental items if they had been required to perform a secondary metalinguistic task, rather than the comprehension task, which did not focus their attention on the case-marking information.

Applying grammatical knowledge during on-line processing
The above research findings demonstrate that even L2 learners of highly advanced proficiency, and who demonstrate the requisite knowledge off-line, may encounter problems with syntactic processing. That is, there are more observed L1/L2 differences with incremental processing in complex sentences, recovery from misanalysis, and linking a fronted element with its subcategorizer, whereas learners have little trouble employing lexical-semantic information in real-time parsing.
Advanced L2 Capacity: Orientations on Acquisition

(e.g. Dussias & Cramer Scaltz, 2008; Jackson & Bobb, 2009; Jacob & Felser, 2016; Juffs & Harrington, 1996; Roberts & Felser, 2011). The findings of persistent difficulties with grammatical processing in advanced learners has led to a range of models being constructed to explain such L1/L2 differences (e.g. the SSH, Clahsen & Felser, 2006; RAGE, Grüter et al., 2014). The theoretical debate surrounds the question of whether parsing differences reflect qualitative or fundamental differences between L2 learners and native speakers (as proposed in the SSH, Clahsen & Felser, 2006), or whether such differences are better attributed to factors such as differences in proficiency, cognitive capacity, processing speed, or the use of memory retrieval mechanisms (e.g. Cunnings, 2017; Dekydtspotter, Schwartz, & Sprouse, 2006; Grüter et al., 2014; Hopp, 2010). The question of fundamental L1/L2 differences is parallel to the debate that is found in the L2 acquisition literature, as to whether late (post-puberty) L2 learners have access to Universal Grammar after the Critical Period (cf. Lenneberg, 1967).

Most researchers involved in this debate assume (implicitly or explicitly) an informationally encapsulated (modular) cognitive system, and a separate grammar and parser, and ask whether, once grammatical knowledge has been independently established, it can be accessed and put to use during parsing. For example, in a self-paced reading study with highly advanced L2 learners from either a (+wh) (Greek, German) or (-wh) L1 background (Chinese, Japanese), Marinis, Roberts, Felser, and Clahsen (2005) found that none of the L2 learners performed like native speakers in their on-line processing of long-distance dependencies (6), despite demonstrating grammatical knowledge of the constructions in a separate, off-line judgment task. Specifically, their native English control group showed a facilitation effect when integrating a fronted NP object the nurse with an embedded verb had angered in conditions in which there was a purported intermediate gap-site at the clause boundary (ei'), as in (6a), in comparison to object-extraction sentences with no syntactic gap (6b), with shorter reading times for the former (8a) at the subcategorizing verb had angered.

(6) a. The nurse [ whoi the doctor argued [ ei' that the rude patient had angered ei ] is refusing to work late.
   b. The nurse [ whoi the doctor’s argument about the rude patient had angered ei ] is refusing to work late.

The L2 learners showed no effects of facilitation, with no difference in processing speed between the two conditions, although they had no difficulty in establishing wh-dependencies in general, as evidenced by the fact that they were slower to read the critical verb had angered in the extraction constructions in comparison to matched-for-length non-extraction conditions (e.g. The nurse believed that the doctor argued that the rude patient had angered the manager in the hospital). The authors argue that these results support a fundamental difference account for L2 parsing, specifically the SSH, which states that L2 learners are less able than native speakers to assign grammatical detail to parses during on-line processing, making more use of lexical-semantic and general heuristics in parsing, even those of very advanced
proficiency. That is, in this study, although the learners could comprehend the experimental sentences, they did not parse the sentences like the native speakers, who were argued to have utilized abstract syntactic information by reactivating the filler at the grammatically-determined intermediate position, rendering it more active in working memory and thus more accessible for integration at its base position after had angered (see also Felser & Roberts, 2007, for similar findings in a cross-modal priming study).

As noted above, the L2 learners in the Marinis et al. (2005) and Felser and Roberts (2007) studies were highly advanced. They were all living in the United Kingdom, and were studying at UK universities, and this fact can be argued to lend support to fundamental difference accounts. However, it may be that the learners were not highly advanced enough to demonstrate native-like use of structural information in on-line parsing, given the fact that a more recent study has found that L2 learners with prolonged naturalistic exposure to English could indeed make use of such detailed abstract information in their processing of comparable long-distance wh-dependencies (Pliatsikas & Marinis, 2013).

Other research findings that go against fundamental difference accounts come from studies which demonstrate that like native speakers, advanced L2 learners are indeed sensitive to abstract syntactic information during on-line processing (e.g. Cunnings, Batterham, Felser, & Clahsen, 2010; Omaki & Schulz, 2011). This was found to be the case in Cunnings et al.’s eye-tracking study which showed that L2 learners (whether from a typologically close L1 or not, German vs. Chinese) respected relative clause islands (7a), which prohibit the establishment of filler-gap dependencies within them (cf. the Subjacency Condition, Chomsky, 1973).

(7) a. Everyone liked the book (city) that the author who wrote continuously and with exceptionally great skill saw whilst waiting for a contract. Island condition

   b. Everyone liked the book (city) that the author wrote continuously and with exceptionally great skill about whilst waiting for a contract. Non-island condition

Specifically, like native speakers (e.g. Traxler & Pickering, 1996), only in non-island conditions (7b) did they attempt to link the filler (book/city) with the first encountered subcategorizing verb (wrote) and when this led to an implausible dependency (wrote the city vs. wrote the book), higher total fixation times were observed. However, and crucially, there was no effect for either natives or L2 learners of the plausibility of the filler (see also Juffs, 2005; Juffs & Harrington, 1995; Omaki & Schulz, 2011).

If one assumes that such island effects are grammatical constraints, which are applied on hierarchical representations, restricting how far the parser searches for a gap (e.g. Phillips, 2006; Stowe, 1986; Traxler & Pickering, 1996; Wagers, Lau, & Phillips, 2009), then the results go against the predictions of the SSH.¹ Despite the fact that the L2 learners appeared to respect the island violations in this study, they
Nevertheless differed from native speakers in that they showed no evidence of on-line recovery. This is similar to what has been observed in other L2 studies (e.g. Roberts & Felser, 2011; Jackson & Bobb, 2009).

**Referential processing and advanced proficiency**

Referential processing has also been a topic of psycholinguistic investigation in highly competent L2 learners (Ellert, 2013; Felser & Cunnings, 2012; Felser, Sato, and Bertenshaw, 2009; Pan & Felser, 2011; Patterson, Trompelt, & Felser, 2014; Roberts, Gullberg, & Indefrey, 2008). The interest lies in the extent to which L2 learners establish dependencies between pronominals and their antecedents in a native-like way, and whether there are influences of a learner’s mother tongue on this process. In general, it has been observed that although learners resolve pronouns during on-line parsing, the process may be disrupted in the presence of other potential antecedents in the discourse, even if they are not grammatically available (Felser & Cunnings, 2011; Roberts et al., 2008). This has been taken to support a number of theories. For instance, that learners are more sensitive to discourse than to syntactic information (cf. the SSH, Clahsen & Felser, 2006), that they may have more trouble integrating information on-line from multiple sources (cf. the Interface Hypothesis, Sorace, 2011), and that their processing is more susceptible to interference in memory retrieval (cf. Cunnings, 2017).

An example comes from eye-tracking studies of reflexive pronoun processing with advanced Japanese and German learners of English (Felser et al., 2009; Felser & Cunnings, 2011). The L2 learners demonstrated knowledge of the locality constraints governing reflexive binding in English, as evidenced by the results of an off-line task, even though German patterns like English (and unlike Japanese) with no long-distance binding. Despite this, the two L2 groups patterned in the same way as each other, and differently to native speaker controls, during their on-line reading of the experimental items (8).

(8) John (Jane) and Richard were very worried in the kitchen of the expensive restaurant. John (Jane) noticed that Richard had cut himself with a very sharp knife.

First pass fixation times in the native speakers showed that they immediately linked the reflexive anaphor with the binding-accessible NP (Richard), but not the inappropriate referent that appeared earlier in the discourse (John/Jane). Although the L2 learners were able to link the reflexive pronoun with the grammatically correct antecedent off-line, on-line they differed from the natives in that there was no evidence of immediate application of such locality constraints. They were slower to read the conditions in which the matrix subject (John) matched in gender with the anaphor, rather than the local, grammatically appropriate antecedent (Jane).

Similar on-line results have been observed with subject pronoun resolution. For instance, Roberts et al. (2008) found that advanced L2 learners of Dutch showed
processing disruption in comparison to the processing of native speakers when a potential competitor was available for resolution in the earlier discourse. Furthermore, and akin to the results of the Felser et al. (2009) Felser and Cunnings (2011) studies above, this was the case, irrespective of the first language of the L2 groups (German vs. Turkish). However, unlike the above studies, there was an effect of the learners’ L1 in their final interpretations of the pronouns. Specifically, in ambiguous discourse context ((9a) *Peter/Hans*), the German learners resolved the pronoun toward the local antecedent (*Peter*), as did the native Dutch, whereas the Turkish learners also sometimes linked the antecedent with the non-local referent (*Hans*), arguably under the influence of their null subject L1, in which an overt pronoun often indicates topic shift and a disjoint, rather than local co-reference, is appropriate.

(9) a. *Peter* en Hans zitten in het kantoor. Terwijl *Peter* aan het werk is, eet *hij* een boterham.
   ‘Peter and Hans are in the office. While *Peter* is working, *he* is eating a sandwich.’

b. De Werknemers zitten in het kantoor. Terwijl *Peter* aan het werk is, eet *hij* een boterham.
   ‘The workers are in the office. While *Peter* is working, *he* is eating a sandwich.’

The on-line data revealed patterns akin to those in Felser et al. (2009) and Felser and Cunnings (2011). That is, the L1 effect was not observed in the on-line reading time data. Both groups of L2 learners slowed down on their reading of the pronoun (*hij* ‘he’) in the contexts in which a competitor antecedent (*Hans*) was available (9a) in contrast to the condition where only one referent was grammatically available (*Peter* [9b]). This was in contrast to the native Dutch speakers, who read the pronoun more quickly in the ambiguous condition (9a), because the discourse context (re-introducing a referent) aligns fully with the grammatical information in the verb and pronoun facilitated resolution.

Taken together, the results of the above studies suggest that even at highly advanced levels, L2 parsing may differ from that of native speakers, even in cases where the L1 and the L2 are closely related and irrespective of the learners’ final interpretations for the pronoun (see also Ellert, 2013).

**Processing relative clause attachment ambiguities**

Another construction that has been the focus of L2 parsing research is the so-called relative clause (RC) attachment ambiguity (10), which has been of interest because there are reported cross-linguistic differences in the preferred attachment site of the RC, and thus researchers have investigated L1 transfer effects on L2 parsing preferences (see Papadopoulou, 2006, for an overview). Specifically, some languages attest the first NP as the preferred host for an ambiguous RC (*the servant*...
was on the balcony, e.g. German, Greek, French, Spanish), whereas in others, NP2 is preferred (the actress was on the balcony, English, Norwegian).

(10) Someone shot the servant of the actress who was on the balcony.

A number of studies have examined the effects of L2 proficiency on L2 learners’ processing preferences with these constructions (e.g. Dussias, 2003, 2004; Dussias & Sagarrà, 2007; Fernández, 1999; Miyao & Omaki, 2006), but in others, only advanced L2 learners have been tested (Felser et al., 2003; Papadopoulou & Clahsen, 2003). An example comes from Felser et al. (2003) who manipulated agreement in a self-paced reading study to force either an NP1 or an NP2 attachment for the ambiguous RC (11). The Greek and German learners in this study were all advanced, and had been studying at a UK university for approximately two years, and the authors investigated whether they would transfer the NP1 preference from their mother tongue to the on-line processing of the English items.

(11) The dean called to the students of (with) the professor who was (were) in the hall.

The native English controls slowed down when they were forced via number agreement to attach the RC to NP1, in comparison to NP2, but the L2 learners showed no preference for either NP1 or NP2 on-line when the complex NP antecedent contained a genitive (of). In contrast, they showed a native-like NP2 preference with a thematic preposition (with). Parallel results were reported in Papadopoulou and Clahsen (2003), even though the target language (Greek) exhibits the same NP1 parsing preference as the learners’ native languages (Russian, Spanish, German). In sum, neither a target-like preference, nor transfer of the L1 preference was observed in these studies and the authors argue that the results support a fundamental difference account, such as the SSH. This is because the L2 learners’ parsing appeared not to be guided by the structural principles that apply in L1 parsing with complex genitive NP antecedents: Predicate Proximity underlining the NP1 preference, stating that speakers should prefer attachment to the host closest to the main predicate (Gibson, Pearlmutter, Canseco-González, & Hickok, 1996), and a locality constraint (Late Closure, Frazier, 1987) guiding NP2 attachment in languages like English. Yet the learners were sensitive to lexical-thematic information in the conditions containing with, which is assumed to restrict the attachment domain to NP2 (cf. Frazier & Clifton, 1996).

However, as noted above, other studies have found both L1 transfer of attachment preferences and a move toward the target language preference when learners were more proficient, determined by the amount of exposure they had to the target language (e.g. Dussias, 2003, 2004; Fernández, 1999; Miyao & Omaki, 2006). For instance, in an eye-tracking study, Dussias and Sagarrà (2007) reported that Spanish-English L2 learners with target language exposure of approximately 5 years exhibited an NP2 attachment preference, like native English speakers, whereas a group who had only been exposed for a short time (9 months) appeared
to have transferred their L1 preference to English. This effect of exposure was also the case for Frenck-Mestre’s (2002) English- and Spanish-French L2 learners who had been resident in France for 5 years and who demonstrated a native-like NP1 attachment preference. Furthermore, learners of relatively low proficiency have been shown to exhibit target-like attachment preferences with constructions with a biasing preceding discourse (cf. English classroom-learners of French, Dekydtspotter, Donaldson, Edmonds, Fultz, & Petrusch, 2008).

One of the reasons that may underlie the mixed results reported in relative clause attachment studies in comparison to the more consistent findings for research using constructions such as subject-object ambiguities and wh-filler-gap constructions, is that the RC adjuncts are non-obligatory constituents and thus open to a more flexible interpretation than obligatory arguments (cf., e.g. the Construal Hypothesis, Frazier & Clifton, 1996). Data from electroencephalogram (EEG) studies on L2 learners’ processing of syntactic versus semantic violations support this idea. That is, syntax-related components (LAN, P600) are less consistently in evidence overall in L2 processing studies that semantically related components (N400), but they are more consistently elicited by structural violations involving obligatory constituents versus optional ones (e.g. *Das Eis wurde im __ gegessen, The ice-cream was in-the eaten, Hahne & Friederici, 2001 vs. *Le chauffeur qui est dans la __ dort, The driver who is in the __ is sleeping, Isel, 2007). The underlying reason is arguably that there are stronger parsing expectations in the case of obligatory in contrast to optional constituents (Isel, 2007). In the latter, factors (discourse-pragmatic, frequency, etc.) may contribute to the interpretation, and this in turn can make on-line interpretation more problematic for L2 learners and therefore more susceptible to proficiency and L1 effects because they are not obligatory arguments.

**Agreement processing and advanced proficiency**

As opposed to studies focusing on L1/L2 parsing strategies reported above, in which the majority of participants are highly advanced learners, research into L2 learners’ agreement processing has much more often specifically focused on questions of proficiency, examining groups at different levels. These studies apply psycholinguistic techniques such as self-paced reading and eye-tracking, but they differ from the parsing research in that they make use of the ‘violation paradigm,’ testing whether L2 learners are sensitive to ungrammaticalities in their real-time sentence processing. As such, although they may speak less to the question of the application of parsing strategies in L2 processing, they are arguably important for formal theories of L2 acquisition (e.g. Full Exposure/Full Access, Schwartz & Sprouse, 1996). This is because such research can provide time-sensitively measured performance data to supplement traditionally elicited, untimed judgment or production data (e.g. Hopp, 2010; Juffs & Harrington, 1995; 1996; White, 1987, 1989, 2003). In other words, such studies examine learners’ knowledge and processing of the ‘limits of grammaticality’ (their understanding of what is and
what is not possible in a language), akin to studies on ultimate attainment in the L2 acquisition literature.

Various kinds of agreement processing have been examined, including subject-verb, case, and gender concord (Alemán-Bañón, Fiorentino, & Gabriele, 2014; Chen, Shu, Liu, Zhao, & Li, 2007; Coughlin & Tremblay, 2013; Foote, 2011; Foucart & French-Mestre, 2012; Jiang, 2004, 2007; Keating, 2009, 2010; Sagarra & Herschensohn, 2010, 2011, 2013), and in a number of cases, highly advanced L2 learners may perform like native speakers; however, proficiency may well interact with other individual factors, such as L1 background, working memory, or processing speed.

As with the parsing research reported above, there are studies reporting that although learners can judge violations correctly as ungrammatical during off-line metalinguistic tasks, they may not be sensitive to this information on-line, particularly if the phenomenon under investigation is not instantiated in their L1, even for highly advanced L2 learners (e.g. for subject–verb agreement violations with Chinese L2 learners of English see Jiang, 2004; and for tense–aspect agreement violations with French and German L2 learners of English, see Roberts & Liszka, 2013). Such findings support fundamental difference accounts of L2 acquisition, in that late L2 learners may not ultimately be able to acquire L2 grammatical features if they have not been learned via L1 acquisition (cf. Hawkins & Chan, 1997), or at least, it may not be possible to make use of such information in real-time processing (see Keating, 2009, for eye-tracking data).

L1 influence and proficiency effects were found in Hopp (2010), who presents the results of a series of experiments which employed both traditional grammaticality and speeded grammaticality judgment tasks (GJTs), as well as reading time studies in which the participants read the sentences for comprehension. He tested advanced and near-native L2 learners of German on a range of morphosyntactic agreement phenomena: number (12), gender (13) and case (14). The aim was to investigate effects of proficiency and L1 background on the learners’ ability to detect number (12) and gender violations (13) during on-line processing, and to assess whether processing slowdowns would be observed with (German) dispreferred accusative-nominative word order constructions (14).

(12) *Er glaubt, dass der Förster im vorigen Jahr den Angler umgebracht haben.
He believes that the forester (SG) in previous year the fisherman (SG)
 killed have.

(13) *Er glaubt, dass die seit langem vermisste Hund im Garten gefunden wurde.
He believes that the (fem) since long missed dog (masc) in the garden
 found was.

(14) Er denkt, dass den Hotelier im August der Gastwirt angezeigt hat.
He thinks that the (ACC) hotelier in August the (NOM) landlord sued has.

Of particular note are the results from the off-line judgment task of the near-native L2 groups. All the learners, including the English, in whose L1 gender is not
instantiated, patterned in a native-like way, which argues against a fundamental difference account, and rather supports a continuity account of L2 acquisition (cf., e.g. Schwartz & Sprouse, 1996). On-line, all the near-native L2 learners showed a processing disadvantage for the dispreferred object-subject constructions (14), but there was an L1 influence in their ability to detect the morphosyntactic errors, with only the Russian learners performing similarly to the native speakers. As noted earlier, L1 influence in the ability to employ grammatical knowledge on-line for very highly advanced learners has also been observed in other studies (e.g. Foucart & Frenck-Mestre, 2011; Jiang, 2004, 2007; Ojima, Nakata, & Kakigi, 2005; Rossi, Gugler, Friederici, & Hahne, 2006; Saggarra & Herschensohn, 2010, 2011, 2013; Tokowicz & MacWhinney, 2005) and could be argued as evidence that a grammatical feature must be available in the L1 if it is to be fully acquired. However, Hopp (2010) reports an interesting finding in his speeded grammaticality judgment task with native German speakers: when they were put under processing pressure, their ability to detect gender (but not number) violations dropped to the level that was observed in his near-native Dutch and English learners. This suggests that it may be processing difficulties at the heart of L2 learners’ non-native-like performance in some circumstances (Goad & White, 2006).

In an on-line agreement processing study in which the learners read sentences for meaning, rather than while performing an judgment task in parallel, Foote (2011) found that her advanced English learners of Spanish performed similarly to native speakers in their ability to detect violations of both number (15) and gender (16), even when the agreeing elements were non-adjacent ([15, 16b]), unlike the results of a comparable study by Keating (2009), whose L2 learners were arguably of lower proficiency.

(15) a. Veo que tu padre es/*son de Texas.
   I see that your father is from Texas.
   b. El reloj del hombre es /*son de Suiza.
      The watch of the man is from Switzerland.

(16) a. Dicen que el libro blanco/*blanca está en esa mesa.
    They say that the book white is on that table.
   b. El pollo del taco está rico/*rica pero picante.
      The chicken of the taco is tasty but spicy.

Similar findings to those in Foote (2011) were reported in Coughlin and Tremblay (2013), whose results also showed that working memory capacity may interact with proficiency, since their highly advanced English L2 learners of Spanish of higher working memory capacity were native-like in their on-line grammaticality detection. In sum, the results of studies into agreement processing show that at advanced levels of proficiency, L2 learners appear to be able to access their grammatical knowledge to compute grammatical relations on-line. However, the ability to do so is influenced by an interaction between proficiency in the target language and other factors, very often L1/L2 similarities in grammar, as well as differences in speed and cognitive capacity.
Conclusion

In this chapter, an overview of L2 psycholinguistic research was provided, zooming in on real-time sentence processing studies, discussed in relation to advanced proficiency. The majority of L2 parsing studies investigate only advanced learners (in contrast to L2 acquisition studies, Plonsky, 2013), in an attempt to address the question of whether or not L2 learners (beyond the critical period for language acquisition) can process the target language like native speakers. This focus on ‘end-state’ learners is unsurprising, given that many researchers wish to use online data to inform theoretical debates in L2 acquisition and processing.

As to whether or not there are fundamental differences between late L2 learners and native speakers (cf. SSH, Clahsen & Felser, 2006) or not (cf., e.g. Cunnings, 2017; Hopp, 2010), there are mixed results. For instance, some results show that on-line sensitivity to grammatical violations becomes more native-like, with fewer effects observed of the L1 and of individual differences with developing proficiency (e.g. Hopp, 2010; Indefrey, 2006; Jackson, 2012). On the other hand, a number of studies have found that even highly advanced L2 learners may not compute grammatical relations like native speakers during on-line sentence comprehension, even though their performance in off-line judgment tasks may be native-like (e.g. Marinis et al., 2005; Roberts, Gullberg & Indefrey, 2008; Roberts & Liszka, 2013). Such psycholinguistic research findings have been used to support or discredit models of L2 acquisition (see Slabakova, 2009, for an overview). However, psycholinguistic approaches to L2 acquisition are relatively rare in comparison to experimental studies of L2 acquisition in general, and therefore it may be that it is the types of populations studied, specific grammatical phenomena that are being tested, and/or the task that is being undertaken that influence the extent to which proficiency impacts on learners’ real-time sentence processing (see also Roberts, 2013, for discussion). Furthermore, proficiency is often correlated with other individual difference factors such as (L2) working memory and speed of processing (Berquist, 1997; Harrington, 1992; Harrington & Sawyer, 1992; Havik et al., 2009; Juffs & Harrington, 2011; Miyake & Friedman, 1998; Osaka & Osaka, 1992; Service, Simola, Metsanheimo, & Maury, 2002), and in many studies it is not always possible to tease them apart (Juffs & Rodríguez, 2014). This would be a fruitful avenue for future research.

Results show that there is evidence that L2 learners of even advanced proficiency have trouble with grammatical processing, particularly with revision and handling non-local dependencies (e.g. Jackson & Bobb, 2009; Jackson & van Hell, 2011), and these results have been taken to support fundamental difference accounts (e.g. Marinis et al., 2005). However, ‘proficiency’ in such studies is often measured by standardized grammatical tests, and it may be that the learners classed as ‘advanced’ were not advanced enough. Counter-evidence from studies in which learners have had prolonged exposure to the target language (e.g. Pliatsikas & Marinis, 2013; French-Mestre, 2002) suggests that taking a snapshot of proficiency from such tests may not allow for the identification of highly advanced, let alone ‘end-state,’ L2 learners. Related to this point, very few L2 parsing studies
investigate bilinguals (in stark contrast to the lexical processing field), and almost none examine the development of L2 processing over time, via longitudinal studies. In sum, it would push forward the field greatly if researchers were to examine a far greater range of participants, and measure their proficiency in far broader terms. In this way, we may better be able to address fundamental questions about the nature of advanced L2 competence.

NOTES

1 There is debate in the literature as to whether island effects are in fact processing effects in nature, reflecting parsing difficulties encountered at the clause boundary of the second relative clause (who wrote) when an attempt is made to link a filler NP (e.g. the author) with the relative pronoun while at the same time having to keep the filler (the book) active in working memory (e.g. Kleunder, 1998, 2004).
2 See van Hell and Tokowicz (2010) for an overview of EEG findings in L2 sentence processing.
3 Other factors may affect L2 learners’ agreement processing, for instance, Kaan, Ballantyne, and Wijnen’s (2015) L1 and L2 participants who were matched for reading speed showed equal sensitivity to number agreement violations.

REFERENCES


Coughlin, C., & Tremblay, A. (2013). Proficiency and working memory based explanations for non-native speakers’ sensitivity to agreement in sentence


What Does Critical Period Research Reveal about Advanced L2 Proficiency?

MICHAEL H. LONG,1 GISELA GRANENA,2 AND FÁTIMA MONTERO1

1 University of Maryland
2 Universitat Oberta de Catalunya

The relationship between Critical Period research and advanced L2 proficiency

One of the most robust and widely accepted empirical findings in research on second language acquisition (SLA) is the inverse relationship between age of first sustained exposure to a new language and eventual attainment in that language. Age of first exposure is usually referred to in the literature as age of onset (AO). The correlation between AO and ultimate second language (L2) attainment is typically around -0.60. Put another way, assuming sufficient time and opportunity to learn, AO can usually account for approximately 36% of the variance in the long-term achievement of a group of L2 acquirers, making it the strongest single predictor of L2 proficiency.

What is less widely accepted is the claim that age effects are the result of biological constraints on the human capacity for language learning. The claim has many variants. The most traditional involves the existence of a single critical period of peak sensitivity to linguistic input lasting from birth to puberty, after which acquisition in all linguistic domains is impaired. However, there is growing empirical support for the idea of a series of sensitive periods (‘sensitive,’ rather than ‘critical,’ to recognize the greater variability in period closures, or offsets, in humans than in lower animals, and the fact that some learning can certainly take place after the offsets), each affecting a different linguistic domain (phonology,
lexis and collocations, morphology and syntax) and/or different types of features within domains. Significant declines in ability have been observed from as early as age 2 for phonology, between 9 and 12 for lexis and collocations, and in the mid-teens for morphology and syntax (Granena & Long, 2013; and for review, Long, 2013). This chapter will focus on advanced learners’ problems in morphology, syntax, and lexis and collocations. For phonology, see Archibald (Chapter 13, Advanced-Level L2 Phonology).

Explanations for maturational constraints vary, too. One involves a change in the mechanisms involved in language learning at different ages, e.g. a loss of implicit learning ability, and a corresponding shift to reliance on explicit learning, in the mid-teens (DeKeyser, 2000). Another points to a decline in the capacity for instance learning by age 12 (Hoyer & Lincourt, 1998), of particular significance for the acquisition of arbitrary, non-rule-governed, linguistic domains, like lexis and collocations (Janacsek, Fiser, & Nemeth, 2012; Long, 2013; Nemeth, Janacsek, & Fiser, 2013). There is also some support for non-mechanism-related accounts (Rastelli, 2014; Thiessen, Girard, & Erickson, 2016). Rather than a discontinuity in the cognitive processes involved, the same processes, e.g. statistical learning, are held to operate uninterruptedly across the lifespan, with the well documented variation in outcomes due to such factors as differences in input and changes in the underlying general cognitive architecture.

One prediction made by all versions of the maturational constraints claim is that attainment of native-like abilities (native-like, not near-native) is impossible if AO is later than X or Y or Z, depending on linguistic domain or categories of features within a domain. These predictions could be falsified by demonstrating that one or more learners has in fact achieved native-like abilities in an L2 in those domains despite first L2 exposure having occurred after the specified ages. Many researchers have set out to do exactly that, but none has succeeded yet (DeKeyser, 2006; Long, 2005, 2007). In fact, the evidence for sensitive periods has become stronger the closer researchers have looked (Abrahamsson & Hyltenstam, 2009; Granena, 2012, 2016; Granena & Long, 2013). In their 2009 study, Abrahamsson and Hyltenstam, for example, found that of 41 highly proficient participants (a subset of an initial sample of 195 screened into a study through being “passed” as native speakers (NSs) by NS raters on the basis of short samples of their spontaneous speech), only three scored within the NS range on all 10 measures; their AOs were 3, 7, and 8. In a study of 65 Chinese immigrants to Spain with lengthy periods of residence in the target-language environment (Granena & Long, 2013), the latest AOs at which there was evidence of native-like attainment were 5 in phonology, 9 in lexis and collocation, and 12 in morphosyntax. If maturational constraints were really a myth, as skeptics maintain (e.g. Marinova-Todd, Marshall, & Snow, 2000), the “exceptions” would be ten a penny—millions of late-starters, long-term residents in the target-language environment, who have attained native-like L2 abilities across the board—and not exceptional at all. The failure to find counter-examples is something the skeptics have yet to explain.

The connection between research on age effects and advanced L2 proficiency is that the search for late-starting native-like attainment has logically involved
studying advanced learners, for it is among them that such individuals would exist if they existed at all. Consequently, so-called Critical Period (CP) research offers fertile ground for examining just what it is that learners who have achieved advanced proficiency can, and still cannot, do. Whatever one’s position on maturational constraints, the findings should provide valuable empirical evidence as to the linguistic characteristics of advanced L2 proficiency.

Early results

By no means all CP research has involved very advanced learners. Many early studies focused on identifying age effects more generally, which meant that samples were often drawn with a wide range of AO and L2 proficiency. Many employed ex post facto correlational designs, where variability on both measures is needed to test for relationships between scores on each of them. As a result, the tests of L2 abilities were frequently easy enough to be usable with lower-proficiency learners, and nowhere near difficult enough to reveal weaknesses among very advanced learners. For example, grammaticality judgment tests would often include items testing knowledge of the need for a word-final -s on plural nouns and third-person singular present in English—hardly the stuff of very advanced proficiency.

Results consistently showed a strong inverse relationship between starting age (AO) and long-term attainment, usually with correlations of around -.60. Several factors appeared to constitute persistent sources of difficulty, most notably (lack of) salience. Salience is a cover term for a variety of often inter-correlated factors (for discussion, see DeKeyser, 2005), some involving properties of the stimulus itself, such as phonological salience (syllabicity), morphological regularity, and semantic complexity, whereas others concern criteria that could alter perceived salience from the learner’s perspective, such as input frequency, whether or not a feature has communicative value (e.g. English -ed vs. redundant third-person singular -s), whether it is processable, therefore learnable, at the time it is encountered, and whether familiarity with similar linguistic phenomena in previously learned languages, including the learner’s first language (L1) and foreign languages, add salience to a feature that might otherwise go unnoticed. For example, word-final inflectional morphology in an L2 Romance language can be expected to be more perceptually salient for an NS of Russian, with its rich inflectional morphology, than for an NS of Chinese, with no such morphology (for relevant empirical work, see Cintrón-Valentín & Ellis, 2015). Similarly, the transparency of a morpheme can be a crucial determinant of its perceptual salience to learners. Transparency is lacking when, say, English L2 articles express meanings that are/are not recognized and encoded in a learner’s L1, or a past time distinction is made in one language but not in another (e.g. subtle T–V distinctions, or the existence of three past tenses in French vs. two in English). It is lacking when an L2 feature, e.g. subject pronouns in a pro-drop L2, is optional, and the conditions for provision and omission are tricky.
Goldschneider and DeKeyser (2001) showed that the first group of (broadly linguistic) components of salience accounted for much of the variance in the accuracy order of English grammatical morphemes. This was a valuable finding, but the morphemes are all items mastered in the early stages of interlanguage development. In a grammaticality judgment test (GJT) study of the acquisition of Hebrew morphology by adult immigrants to Israel (DeKeyser, Alfi-Shabtay & Ravid, 2010), salience became more important with increasing AO. A study suggesting the greater importance of salience than markedness was reported by Bardovi-Harlig (1987), who showed that preposition-stranding in L2 English (Who did he give the book to? Where had the refugees come from?), which leaves the preposition in salient utterance-final position, was acquired before pied-piping (From where had the boat set out? To which university did he donate the money?). The generalizability of that finding is obviously limited, however, given the vastly different frequencies of the two constructions in spoken English. Saliency’s role at very advanced levels remains an open question. The fact is, most studies that have involved very advanced learners to test some version of the Critical Period Hypothesis (CPH) have used global proficiency measures, e.g. whether or not an accent was native-like, or whether transcripts of speech could be identified as those of natives or non-natives on the basis of grammar and word choice. Few examinations of very advanced learners have looked at the degree of mastery of particular structures or lexical items and collocations. There has been some such research, however.

**Morphology and syntax**

Highly proficient L2 learners tend to have had an early AO, as the result either of migration or of the commencement of a formal language program. In a major study that “scrutinized” apparent linguistic native-likeness (Abrahamsson & Hyltenstam, 2009), out of an initial sample of 195 Spanish-speaking learners of Swedish who had either self-selected or been nominated by acquaintances as having native-like Swedish, 154 were identified as non-native speakers (NNSs) by NS judges simply on the basis of listening to brief recorded samples of their speech. Of the 41 who the judges believed were native Swedes, 31 (i.e. 75%) were childhood learners with AO ≤ 11, and 10 were adult learners with AO ≥ 12. Of the 10 adult learners, none scored within the NS range across the whole set of tasks employed in the second phase of the study. In fact the latest AO for a participant who scored within the native range on all tests was 7. Scrutinizing performance in this way indicated that even those L2 learners who can “pass” as natives in everyday life differ from true natives in a variety of linguistic subdomains.

Disparities are not random and are not only found in phonology and lexis, but also in morphology and syntax, where the window of learning opportunity is considered to close later. In each case, differences do not affect a whole domain, but are found at the level of particular linguistic features. A re-analysis of datasets from two studies (Granena, 2012, and Granena & Long, 2013) was performed with
the aim of identifying similarities and disparities at the level of linguistic features within domains between early L2 learners, late L2 learners, and NSs.

In the first study (Granena, 2012), participants had been born to Chinese-speaking parents who had immigrated to Spain as adults. Their L1 was Chinese, and they had first been exposed to the L2 at or after age 3. They were, therefore, sequential bilinguals (i.e. L2 learners) rather than simultaneous bilinguals (Gathercole, 2007). The study included 100 L2 learners in Madrid and 20 Spanish NS controls (N = 120), all of whom were at least 18 years of age at time of testing. Half the Chinese-Spanish bilinguals (n = 50) were early-childhood L2 learners (42% males and 58% females) with AOs ranging from 3 to 6. The other half (n = 50) were late L2 learners (34% males and 66% females) with AOs 16 and older. Six grammatical structures were investigated: (i) noun-adjective gender agreement, (ii) subject-verb (person) agreement, (iii) noun-adjective number agreement, (iv) subjunctive mood, (v) perfective/imperfective aspect contrasts, and (vi) passives with ser/estar. Some of the structures (those involving grammatical agreement) are formal features involving uninterpretable grammatical relations typically acquired by L1 speakers of Spanish by age 3, and with few errors (Meisel, 1990; Slobin, 1985). The others (subjunctive, aspect, and passives) involve interpretable features (i.e. grammatical choice requires access to semantic interpretation) and are not mastered with 100% accuracy by L1 Spanish speakers until age 7 or later (Montrul, 2004).

Participants in the original study were administered a battery of tasks that included four GJTs, crossing modality (visual/auditory) and speed (timed/untimed), a word-monitoring task, and an error-correction task. In this re-analysis, percentage scores were computed for each of the target structures on the four GJTs. First, the scores for the three structures that involved grammatical agreement were compared with scores for the three structures that did not. Multiple comparisons with Bonferroni correction showed that early-childhood learners scored significantly lower than NSs on agreement structures (p < .001), but comparably on structures that did not involve agreement (p = .065). There were no significant differences between early-childhood learners and late learners on structures involving agreement (p = .399), but early-childhood learners scored significantly higher than late L2 learners on structures that did not involve agreement. Follow-up analyses revealed that NS and early-childhood learner scores differed on each of the three structures involving agreement: noun-adjective gender agreement, noun-adjective number agreement, and subject-verb (person) agreement. In all three cases, early childhood learners scored significantly lower than NSs (p < .001). They scored significantly higher than late L2 learners on number agreement (p = .040), but comparably on gender and person agreement (p = .110 and p = .853). In the case of structures that did not involve agreement, early-childhood learners scored like NSs on the subjunctive (p = .322) and aspect contrasts (p = .628), but significantly lower on the passive (p = .018).

Tables 4.1, 4.2, and 4.3 show the different patterns identified in group comparisons for individual structures. A statistically significant difference is indicated visually by an additional row. As can be seen, there were three different patterns
in the data. Early childhood L2 learners scored comparably to NSs and significantly higher than late L2 learners for the subjunctive and aspect contrasts (Table 4.1), significantly lower than NSs but higher than late L2 learners for number agreement and passive (Table 4.2), or significantly lower than NSs and comparably to late learners for gender and person agreement (Table 4.3). Scores on gender and person agreement were the lowest in both groups of L2 learners, whereas scores on the subjunctive were the highest.

**Table 4.1**  NSs = AO 3–6 > AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjunctive</td>
<td>96.5 (6.7)</td>
<td>88.5 (7.4)</td>
<td>91.0 (10.9)</td>
</tr>
<tr>
<td>Aspect</td>
<td>88.5 (7.4)</td>
<td>85.0 (10.4)</td>
<td>67.0 (17.6)</td>
</tr>
</tbody>
</table>

*Note:* Standard deviations are provided in parentheses.

**Table 4.2**  NSs > AO 3–6 > AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>89.5 (7.3)</td>
<td>74.2 (14.1)</td>
<td>66.4 (17.1)</td>
</tr>
<tr>
<td>Passive</td>
<td>83.0 (9.8)</td>
<td>73.8 (13.4)</td>
<td>54.6 (11.5)</td>
</tr>
</tbody>
</table>

*Note:* Standard deviations are provided in parentheses.

**Table 4.3**  NSs > AO 3–6 = AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>93.5 (5.9)</td>
<td>68.8 (10.4)</td>
<td>62.8 (10.1)</td>
</tr>
<tr>
<td>Person</td>
<td>87.5 (7.4)</td>
<td>70.3 (14.0)</td>
<td>65.4 (11.8)</td>
</tr>
</tbody>
</table>

*Note:* Standard deviations are provided in parentheses.
The second re-analysis involved data from Granena and Long (2013). Participants had the same profile as in Granena (2012), Chinese L1-Spanish L2 speakers, but in this study AOs were split into three groups: 3–6 (early-childhood), 7–15 (late-childhood), and 16–29 (late). Three of the structures in Granena and Long (2013) were the same as in Granena (2012): the subjunctive, aspect contrasts, and gender agreement. Tables 4.4, 4.5, and 4.6 show the patterns identified for each of these structures. As can be seen, the patterns are very similar to those identified in Granena’s (2012) data. For the subjunctive and aspect contrasts, there were no significant differences between early-childhood L2 learners (AOs 3–6) and NSs ($p = .999$ and $p = .583$, respectively). Late-childhood L2 learners (AOs 7–15), on the other hand, scored significantly lower than NSs on the subjunctive ($p < .001$), marginally lower on aspect contrasts ($p = .050$), and lower than early-childhood learners on the subjunctive, but comparably to early-childhood learners in the case of aspect contrasts ($p = .531$). Finally, gender agreement was again the structure where L2 learners scored the lowest, as well as the structure showing no significant differences between early-childhood, late-childhood, and late L2 learners.

On the basis of these results, it seems that some grammatical distinctions and constructions are acquired later than others, as shown by the existence or absence of significant differences between L2 learner and NS performance. A statistically significant difference between NSs and early-childhood learners suggests that native-like performance for the grammatical distinction in question can only be achieved if the L2 learning process starts before, or at, age 3 or 4. On the other hand, the lack of a significant difference between NSs and L2 learners suggests a later window of learning opportunity. In the two datasets analyzed, grammatical agreement (i.e. morphological covariation) seemed to be an early acquired grammatical distinction, particularly in the case of gender and person agreement. These two agreement types differ from noun–adjective number agreement in that, in number agreement, number information usually stems from the conceptual level: the singular form is used to refer to one entity, whereas the plural form is used to refer to two or more entities. On the other hand, gender and person agreement can be argued to operate over a syntactic, rather than a semantic, representation (Bock, Eberhard, Cutting, Meyer, & Schriefers, 2001; Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008; Nicol & Greth, 2003).

While grammatical agreement seems to be an early acquired feature likely to pose a problem even for very advanced, near-native L2 speakers, it seems possible for more interpretable features that contribute to grammatical meaning and that are not purely formal in nature to be acquired later. This is the case with the subjunctive and aspect contrasts, for which there were no significant differences between NSs and early-childhood L2 learners or between NSs, early-childhood learners, and late-childhood learners. In the case of the passive construction, however, there were significant differences between NSs and early-childhood learners, despite the passive also being an interpretable feature. This indicates that factors other than interpretability are at work in how early or late grammatical distinctions are acquired. One of these factors could be that some grammatical distinctions lie at the interface between two domains or subdomains (Sorace, 2004), e.g. syntax...
Table 4.4  NSs = AO 3–6 > AO 7–15 > AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 7–15 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjunctive</td>
<td>87.5</td>
<td>86.7</td>
<td>60.7</td>
<td>54.6</td>
</tr>
<tr>
<td></td>
<td>(5.2)</td>
<td>(6.1)</td>
<td>(14.6)</td>
<td>(18.1)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are provided in parentheses.
What Does Critical Period Research Reveal about Advanced L2 Proficiency?

and semantics in the case of copula choice in eventive/stative passives in Spanish (Bruhn de Garavito & Valenzuela, 2008).

Note, however, that regardless of the grammatical distinction, L2 learners with AOs 16+ always performed significantly lower than NSs and, in many cases, significantly lower than those with AOs in early-childhood and/or late-childhood. From this perspective, all the grammatical distinctions investigated in Granena (2012) and Granena and Long (2013) are acquired relatively early by L2 learners (i.e. before age 15) and pose learning problems for late L2 learners, with AOs of 16 or more. Based on the findings reported here, the severity of learning problems can vary at higher levels of proficiency, with uninterpretable features or features involving the intersection of grammatical subdomains posing a more severe learning problem for advanced L2 learners whose L1 lacks those particular features.

An area where L2 performance was found to be indistinguishable from that of NSs, regardless of AO, was unmarked word order (syntax). Word order is clearly a matter of linguistic competence and plays a special role in discourse situations, where speakers who do not conform to the norms are quickly judged to be non-native (Johnston, 1995). In Spanish, word order is considered to be free, but this freedom is more apparent than real if the distribution of information is taken into

### Table 4.5 NSs = AO 3–6 = AO 7–15 > AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 7–15 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect</td>
<td>79.6 (6.8)</td>
<td>75.2 (9.9)</td>
<td>71.3 (10.4)</td>
<td>62.5 (10.1)</td>
</tr>
</tbody>
</table>

_Note_: Standard deviations are provided in parentheses.

### Table 4.6 NSs > AO 3–6 = AO 7–15 = AO 16+.

<table>
<thead>
<tr>
<th></th>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 7–15 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>86.6 (5.9)</td>
<td>61.5 (15.4)</td>
<td>59.2 (13.6)</td>
<td>51.0 (14.9)</td>
</tr>
</tbody>
</table>

_Note_: Standard deviations are provided in parentheses.
account. As part of the Granena and Long (2013) study, two word order preference tests were administered to participants. One assessed unmarked word order and included inversion in questions (e.g. ¿Quién Juan quiere que le escriba? vs. ¿Quién quiere Juan que le escriba?), verb raising over short adverbs (e.g. Antonio prepara rápido las maletas vs. Antonio rápido prepara las maletas), verb raising over -mente (manner) adverbs (e.g. El niño abrió tranquilamente la puerta sin mirar vs. El niño tranquilamente abrió la puerta sin mirar), adverb ordering (e.g. Sofia insistentemente llama a veces al teléfono vs. Sofia a veces llama insistentemente al teléfono), and raising over two adverbs (e.g. Antonia escribe siempre lentamente los deberes vs. Antonia siempre lentamente escribe los deberes). The second preference test focused on marked (or discourse-based) word order, particularly word order with unaccusative verbs (e.g. Desapareció un cuadro de Velázquez y no logran encontrarlo), SV word order with unergative verbs (e.g. Sonia nadó muy bien y ganó el primer premio), VS word order with unergative verbs (e.g. Bailó Laura pero porque ya se habían ido todos), and object clitics in topicalized constructions (e.g. La noticia la leí por la mañana antes de ir a trabajar vs. Un frasco de perfume me va a regalar Pedro).

Table 4.7 shows the means and standard deviations for the different AO groups on the unmarked word order preference test. As can be seen, there were no significant differences between the groups ($p > .05$). Table 4.8 shows the results for marked word order. In this case, there were significant differences between NSs

<table>
<thead>
<tr>
<th>Table 4.7</th>
<th>NSs = AO 3–6 = AO 7–15 = AO 16+ .</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSs (%)</td>
<td>AO 3–6 (%)</td>
</tr>
<tr>
<td>Unmarked order</td>
<td>90.3 (5.2)</td>
</tr>
</tbody>
</table>

*Note: Standard deviations are provided in parentheses.*

<table>
<thead>
<tr>
<th>Table 4.8</th>
<th>NSs = AO 3–6 &gt; AO 7–15 = AO 16+ .</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSs (%)</td>
<td>AO 3–6 (%)</td>
</tr>
<tr>
<td>Marked order</td>
<td>73.7 (11.7)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Standard deviations are provided in parentheses.*
and both late-childhood and late L2 learners ($p < .001$) and also between early-childhood learners, on the one hand, and late-childhood and late learners, on the other ($p < .001$), but not between NSs and early-childhood learners ($p = 1.000$), or between late-childhood and late L2 learners ($p = .199$). The confluence between word order and discourse, an interface between two domains (i.e. using word order in different pragmatic contexts), could explain the differences observed. As argued by Bruhn de Garavito and Valenzuela (2008), the fact that the word order/discourse interface is external and, therefore, that relevant information is only supplied by context, makes it a particularly challenging area for L2 learners.

**Lexis and collocations**

While most work on advanced L2 proficiency (and on SLA in general) has focused on morphology and syntax, it is well known that as proficiency increases, grammar is less of an issue. Vocabulary and collocations, on the other hand, become an increasingly significant part of the learning task and remain lifelong areas of weakness for all but (by definition) the tiny minority of near-natives. The passive vocabulary size of the average English NS ranges from 20,000 to 35,000 words, and NSs know many times more collocations than that. The average size of the L2 lexicon varies, but it is usually much smaller. Skillful speakers and writers can often circumvent lexical voids through such strategies as topic avoidance, circumlocution, and use of generic items when specific terms in the L2 are unknown (‘glass’ for ‘goblet,’ ‘river’ for ‘stream,’ ‘dog’ for ‘poodle,’ ‘metal’ for ‘iron,’ ‘laugh’ for ‘giggle,’ ‘walk’ for ‘stumble,’ and so on), but the fact remains: vocabulary and collocations, not grammar, are typically the areas in need of most work at advanced levels. In fact, in a study of the written work of Hebrew-speaking learners of English, Laufer and Waldman (2011) found (no doubt in part because advanced learners attempt more, and more complex, language in all domains, not just vocabulary and collocations) that intermediate and advanced learners produced significantly more deviant noun-verb collocations than elementary learners. There was an inverse relationship between proficiency and correctness (2011, p. 663).

The problems of even very advanced learners with vocabulary and collocations have been documented by researchers testing some version of the CPH. In one of the first CPH studies to focus exclusively on lexis and collocations, Spadaro (1996, 2013) compared the performance of 38 very advanced, highly educated, adult NNSs of Australian English with that of 10 NS controls of comparable gender and educational background. Motivated by previous findings concerning AO effects on the acquisition of morphology and syntax, she divided the NNSs into three groups: those with AOs of 0–6 ($n = 13$), 7–12 ($n = 15$), and 13 and older ($n = 10$). Participants completed the Kent–Rosanoff word-association task, an oral production task (retelling the story in a video showing the destructive behavior of a dog left alone in a house), and a battery of seven written tasks probing use of core vocabulary and multi-word units. Transcripts of the oral narratives were coded for frequency of participants’ use of core and
non-core vocabulary and multi-word units, with NS judges also asked if they
could distinguish the NS and NNS narratives.

The seven written tasks, with sample items, were as follows. Results for each
task appear in square brackets.

1. **Core preference** \((k = 10)\): learners’ preference for core vocabulary over more appropriate non-core items, e.g.
   
   She agreed to _____ a kidney to save somebody’s life (give > donate)
   The drunk ____ over to the bus-stop (walk > stumble)
   
   [No significant differences among the groups.]

2. **Word discrimination** \((k = 24)\): identifying 12 real two-syllable English words (sunshine, moonlight) from a list of 24 that included 12 morphologically plausible counterfeits (nightbreak, kindwill).
   
   [NS > Group 2 (AO 7–12).]

3. **Core overextension** \((k = 15)\): picking one sentence in four that incorrectly extended the use of a core lexical item, such as ‘pass,’ e.g.
   
   Bessie passed a relaxing month in Provence
   Twenty years have passed since they last saw each other
   You’ve got the money you demanded – now pass the hostage
   The whole class passed the weekly test
   
   [NS > Group 2 (AO 7–12) and Group 3 (AO 13+).]

4. **Multi-word unit completion** \((k = 18)\): providing the final word to complete three-word phrases, e.g. aches and ____ , labor of ____ , tie the ____ .
   
   [NS > Group 2 (AO 7–12) and Group 3 (AO 13+). Group 1 (AO 0–6) > Group 3 (AO 13+).]

5. **Multi-word unit supply** \((k = 10)\): provide unit using single words supplied, e.g. gab (gift of the gab) spick (spick and span), kilter (off kilter), and beck (beck and call).
   
   [NS > Group 2 (AO 7–12) and Group 3 (AO 13+). Group 1 (AO 0–6) > Group 2 (AO 7–12) and Group 3 (AO 13+).]

6. **Multi-word unit correction** \((k = 25)\): identifying and correcting the incorrect word in idiomatic uses in sentences, e.g.
   
   I’m afraid you’re growling up the wrong tree
   The boss likes to throw his size around
   I didn’t want to do it but he bent my arm
   
   [NS > Group 2 (AO 7–12).]

7. **Multi-word unit transformation** \((k = 18)\): identifying 10 sentences (among 18) in which the “frozenness” of units had been violated through transformations of various kinds (active/passive, declarative/interrogative, positive/negative, word order changes), e.g.
   
   Whose eye is she the apple of?
   The bush was thoroughly beaten about by her
   Dolores is the party’s life and soul
   
   [NS > Group 2 (AO 7–12) and Group 3 (AO 13+).]
The reliability of each of the seven tests was acceptable, as was the battery’s overall reliability (Cronbach’s alpha = .895). The uniformly high NS scores indicated that they were measuring the targeted construct, native-like lexical competence, which only the NS group as a whole possessed.

When results across all seven written tasks were compared, Spadaro found that NS scores were on average 10 percentage points higher than those of Group 1 (AO: 0–6), which in turn were over 10 points higher than those for groups 2 and 3 (AO: 7–12, and 13+). Groups 2 and 3 differed very little from one another. Results for Task 3, core overextension, probing use of core vocabulary items, revealed that even at the very advanced proficiency levels represented in Spadaro’s study, there was a tendency for non-native participants not to know the outer limits on use of some core vocabulary items, such as break, set, and lead. The same weakness was echoed in results on tasks 6 and 7, where errors were sometimes due to inappropriate use of core items in multi-word units. In Task 6, for example, asked to correct ‘Vincent’s father hauled him on the coals for denting the car,’ where on should be changed to over, NNSs in groups 2 and 3 offered such solutions as ‘Vincent’s father dragged, raked, held, put, and the rather alarming grilled, him on the coals for denting the car.’ As Spadaro notes, knowledge of how core words participate in collocations and larger phrases is one of the criteria commonly proposed for knowing a word. The NNSs were less sure of the semantic boundaries and collocational possibilities of the core items.

All the NNS groups (but not the NSs) found the phrase-level tasks much harder than the word-level tasks, with Task 4 the best discriminator among the four groups. Means for the NS and NNS groups 1, 2, and 3 were 96.1, 88.4, 69.0, and 65.1, respectively. The task yielded the highest F of the seven ($F = 7.8269$) and the largest effect size (1.7). With the exception of Task 1, where no statistically significant differences were observed, the general pattern was for the NS group to score statistically significantly higher than at least one of groups 2 and 3, and usually higher than both of them. Conversely, the NS group never scored statistically significantly higher than the youngest NNS group, Group 1 (AO 0–6). Group 1 scored more like the NS group than the two groups with AOs of 7 or later, and statistically significantly better than one or both of them on tasks 4 and 5.

Results on the Kent–Rosanoff word-association task were not significantly different across the four groups, suggesting that very advanced NNSs come to resemble NSs with regard to the patterns of association between items in their mental lexicon. Where the oral narratives were concerned, the NSs produced longer texts than the NNS groups, but not statistically significantly so. There were no differences, either, in NS and NNS proportions of use of core and non-core words. Finally, when four NS raters were asked to assign native or non-native status to the transcripts, each NNS participant was misidentified as a NS by at least one rater, an indication of the very high overall proficiency of the participants. In general, however, numbers of ‘native’ ratings assigned to participants mirrored results on the seven written tasks, with NSs receiving the highest number, followed by Group 2 (AO 0–6), with groups 3 and 4 considerably lower and not differing very much from one another.
As noted previously, the likely cause of the lexical deficit is the weaker capacity for instance learning. It is not simply a matter of reduced time on task for late starters. Length of residence (LoR) was only a factor on Task 4 in Spadaro's study, multi-word unit completion; however, it served not to increase scores, but to reveal greater differences among the AO-defined groups when it was controlled for using ANCOVA (analysis of covariance). Lower overall quantity of input can certainly be a predictor of higher proficiency (although usually only for the first 1–3 years at most), but the development of lexis and collocations shows the same discontinuity, followed by a decline, as is witnessed for phonology, morphology, and syntax, just after a different AO. The deterioration in Spadaro's sample seemed to occur soon after an AO of 7.

A second study to examine the lexical and collocational abilities of very advanced learners was that of Granena and Long (2013), already discussed with reference to results for morphology and syntax. The development of lexis and collocations by the same learners was examined using a set of Spanish tasks based closely on those developed by Spadaro for English. As shown in Table 4.9, consistent with Spadaro's results (see also Hyltenstam, 1988, 1992; Munnich & Landau, 2010), Granena and Long found a noticeable (and statistically significant) age-related discontinuity in the development of Spanish vocabulary and collocations, with a clear drop-off in long-term attainment by learners with an AO of 7 or older. Like Spadaro, they interpreted this as indicating an offset between 7 and 9 of a hypothesized sensitive period for lexis and collocations, distinct from the earlier one for phonology and the later one for morphology and syntax. In other words, just as with Spadaro’s participants, the same (late) AO meant that prospects for the acquisition of grammar were still positive well after lexis and collocations had become a long-term problem.

Several tasks differentiated the groups in Granena and Long’s study, none more so than idiom correction (see Table 4.9). The 12 NSs performed statistically significantly better than the latest AO group on all the tasks and also better than the AO 7–15 group on all the tasks except receptive core vocabulary and verbal collocations. NSs even performed significantly better than the AO 3–6 group (n=20) on idiom correction. The AO 3–6 group did statistically significantly better than the AO 7–15 group (n=27) on that task, and on idiom unit completion, prepositional verbs, both productive and receptive core vocabulary, and word/non-word. The AO 7–15 group did statistically significantly better than the AO 16+ group (n=18) on idiom correction, idiom unit completion, prepositional verbs, and oral compound completion. The AO 16+ group failed to outperform any other group on any task. Prepositional verbs and idiom completion (both productive) discriminated among the three groups of L2 learners. Idiom correction discriminated among all four groups, including NSs outperforming Group 1 (AO 3–6). Most of the collocations examined in the Spadaro and Granena and Long studies, especially those in Spadaro’s tasks 4, 5, 6, and 7, were of a very limited kind found mostly in frozen or semi-frozen idiomatic usages. To the best of our knowledge, there has yet to be a CPH study focusing on a more representative set of collocations. However, the results are likely to confirm the claim made here that, together
Table 4.9  Lexis and collocations (Granena & Long, 2013).

<table>
<thead>
<tr>
<th>NSs (%)</th>
<th>AO 3–6 (%)</th>
<th>AO 7–15 (%)</th>
<th>AO 16+ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean %</td>
<td>Test</td>
<td>Mean %</td>
<td>Test</td>
</tr>
<tr>
<td>92.2 (5.4)</td>
<td>Idiom (unit) completion (productive)</td>
<td>85.7 (8.7)</td>
<td>Verbal collocations (receptive)</td>
</tr>
<tr>
<td>90.9 (6.7)</td>
<td>Idiom (unit) correction (productive)</td>
<td>84.9 (7.3)</td>
<td>Oral compound completion (productive)</td>
</tr>
<tr>
<td>90.3 (9.9)</td>
<td>Word/non-word (receptive)</td>
<td>83.5 (10.4)</td>
<td>Word/non-word (receptive)</td>
</tr>
<tr>
<td>90.0 (8.5)</td>
<td>Oral compound completion (productive)</td>
<td>80.8 (8.9)</td>
<td>Prepositional verbs (productive)</td>
</tr>
<tr>
<td>88.4 (4.5)</td>
<td>Prepositional verbs (productive)</td>
<td>78.4 (17.6)</td>
<td>Idiom (unit) completion (productive)</td>
</tr>
<tr>
<td>87.4 (7.5)</td>
<td>Verbal collocations (receptive)</td>
<td>74.5 (13.5)</td>
<td>Receptive core vocabulary</td>
</tr>
<tr>
<td>66.1 (19.1)</td>
<td>Receptive core vocabulary</td>
<td>71.1 (21.2)</td>
<td>Idiom (unit) correction (productive)</td>
</tr>
<tr>
<td>60.8 (9.7)</td>
<td>Productive core vocabulary</td>
<td>55.0 (13.1)</td>
<td>Productive core vocabulary</td>
</tr>
</tbody>
</table>

*Note: Standard deviations are provided in parentheses.*
with vocabulary, they, not grammar, constitute the biggest problem at advanced proficiency levels.

Reasons for the difficulty collocations pose are not hard to identify. As has been noted elsewhere (e.g. Boers & Lindstromberg, 2012; Long, 2015, pp. 307–316), at least nine sources of difficulty make collocations hard for adults: (i) L1/L2 differences; (ii) Collocations often consist of familiar words (turn up, ask around), so do not attract the learner’s attention; (iii) Insufficient frequency in the input; (iv) Semantic vagueness (have/make time, do the shopping, cut corners); (v) A time lapse between encounters can mean the initial memory trace has faded or been lost altogether before the next hearing or reading. (Encountering the sequence ‘PTR’ twice in a series of hundreds of letters makes it harder to recognize as a “collocation” than encountering ‘PTR’ twice within a few seconds or within, say, 20 letters.); (vi) Collocates are sometimes interrupted (War was finally declared on December 1st, 1986); (vii) Collocates of some words vary (do/complete/perform/accomplish a task); (viii) Cue competition exists for some deceptively similar collocations (pay the price/pay the cost, have/do/make/save/keep/mark time); (ix) The fact that many collocation errors do not cause communication breakdowns (arrive at/in London, make/take a photo, a little/small minority) means that deviant versions may not elicit negative feedback, leaving learners unaware of a problem and dependent on positive evidence or indirect negative feedback for its (usually much slower) resolution.

**L2 learning at advanced proficiency levels:**

**Summary of findings**

Research findings from CPH studies converge in showing that native-like L2 ultimate attainment among (late) L2 learners is not possible. These are studies with multi-task designs that have tested multiple linguistic domains and compared the extent to which the end-state of L2 knowledge resembles that of an NS. This does not mean that some advanced learners are not able to pass as NSs in everyday life or perform tasks like NSs. In fact, a number of them do. Neither does it mean that they are unable to learn certain linguistic items to native-like levels. What CPH research tells us is that native-like attainment across all the linguistic items within a subdomain (phonology, lexis, or morphosyntax) is not possible, and that the number of items that will be acquired to a native-like level will depend on AO.

In the studies reviewed above, the linguistic items that were not acquired even by very advanced early childhood L2 learners of Spanish were noun-adjective gender agreement, noun–adjective number agreement, subject–verb (person) agreement, and passives. This suggests that grammatical agreement and grammatical distinctions at the interface between two subdomains, such as syntax and semantics, pose learning problems for the advanced L2 learner. In the area of lexis and collocations, findings indicate that native-like production of multi-word units is impossible even among early childhood L2 learners. Phrase-level tasks that involved productive knowledge were consistently much harder for all L2 learners in the studies reviewed. The linguistic items that were acquired to native-like levels were
unmarked word order, a syntactic feature, marked word order (provided acquisition started before age 7), and some interpretable features, such as aspect and the subjunctive (provided again that acquisition started before age 7). In the area of lexis, receptive knowledge at the word level did not seem to pose a problem for L2 learners. Also, the patterns of association between lexical items in the mental lexicon were similar between NSs and very advanced L2 learners.

**Types of items resistant to learning even at advanced proficiency levels**

Identification of classes of late acquired, possibly never acquired, morphology, syntax, lexis, and collocations is of interest for their potential for predicting items and features likely to prove problematic for all adult learners and/or in many languages. Clearly, recalcitrant problems for adults in particular languages, such as Korean quantifiers or Russian verbs of motion, may be uniquely difficult (for adult English speakers, at least) and so of limited generalizability, with implications for other languages only at a very abstract level. To what extent is it possible to generalize from CPH studies or other research on advanced learners to new L1 and L2 pairings, thereby alerting learners and instructors alike to difficulties to which they will need to devote special attention?

A possible starting-point is to consider the L1-L2 pair in question and compare the two languages with respect to the linguistic feature of interest. The extent to which the two languages resemble one another is a factor that can help predict acquisition problems, but only in some cases. L2 learners can equally struggle with features that are absent or present in their L1. For example, Meisel (2009) mentions that gender assignment and agreement constitute a major acquisition problem for German learners of French, despite German being a gender-marking language. L2 learners, therefore, do not always rely on their L1 grammatical knowledge.

Generalizations are risky, but several proposals have been made about classes of problematic features in recent years. One of the best known is that of Sorace (2003, 2005), who has flagged phenomena at the “interface,” e.g. where morphology and syntax interact with semantics and pragmatics, as the source of lasting problems often found in end-state grammars. Examples in L2 Spanish include word order and morphology in certain discourse (pragmatic) contexts, e.g. topicalization and clitic-doubling. English-speaking learners of Spanish persevere with the canonical L1 and L2 SVO order and seem to avoid VOS, which can delay later developmental stages where word order and object-marking are inextricably entwined in certain discourse (pragmatic) contexts, as in *La ventana la rompió el niño* (*The window it[sing., fem.] broke the boy*). Note, however, that as is so often the case, sources of difficulty rarely occur in isolation. Spanish clitics may occur at the interface, for example, but they are also string-internal, communicatively redundant, and often so lightly articulated in fast spoken Spanish as to be virtually inaudible. In other words, perceptual salience is implicated once again.
In a detailed examination of the fossilization literature, Long (2003) pointed out that (among several other requirements) for a claim about permanent cessation of development of one or more linguistic constructions at the level of individuals to go through, supposedly fossilized structures/usages need to co-occur with other areas of an interlanguage that are demonstrably continuing to develop. Only then can putative fossilization (or stabilization) be differentiated from maturational constraints, which, if real, operate on all learners after certain ages of onset, and across whole linguistic subdomains (phonology, lexis and collocations, morphology and syntax) or all items meeting certain characteristics within a domain, not on particular structures. If that were not part of the definition, fossilization would be a redundant construct. Building on that idea and referring to it as “selective fossilization” (although “selective” is redundant if selectivity is part of the definition), Han (2009) proposed the Selective Fossilization Hypothesis. The idea is that the “acquisitional complexity” (Han & Lew, 2012) of particular structures/usages, but not others within the same individual, is predictable from a combination of L1 markedness and the robustness of L2 input. Marked forms are infrequent and variable; unmarked forms are frequent and consistent. L2 input is robust when features are of high frequency and low variability; low frequency and high variability, conversely, make input non-robust. Fossilization (or stabilization) is more likely to occur when comparable structures/usages are marked in the L1 and input on the L2 equivalent is non-robust. It is not clear how general qualities of the L1 and L2 like these could ever account for differential success with the same items at the level of individuals, but Han’s model offers a principled way of predicting at least some of the linguistic problems of advanced learners.

**Avenues for future research**

Given the relative neglect of advanced proficiency and the advanced learner until recent years, much work remains to be done. To begin with, further descriptive studies of the kind reported here are needed to identify and confirm classes of morphological and/or syntactic items problematic for advanced learners across a typologically wider range of languages (English, French, German, Mandarin, Russian, Korean, Arabic, Persian, Turkish, etc.). Where instructional solutions are concerned (as discussed in some detail in Long, 2017), the aim is to identify the least intrusive, but still efficient, forms of intervention for different classes of linguistic targets (salient/non-salient, marked/unmarked, etc.) for different kinds of learners, e.g. those with relatively greater strengths in explicit or implicit language aptitudes. Are traditional proactive explicit instruction, intentional learning, and noticing always needed, or even less intrusive reactive focus on form and intentional learning, or can the same results be achieved simply through enhanced incidental learning and detection? Aptitude–treatment interaction (ATI) studies and comparative evaluations are needed, in both experimental laboratory and classroom settings.
Even more work is needed on lexis and, especially, collocations. Both domains are increasingly important at advanced levels, yet even simple descriptive studies are few and far between. How to deal with vocabulary and collocations most efficiently is a major issue yet to be resolved by research on instructed SLA (ISLA) at any, not just advanced, proficiency levels. Prescriptions in the language teaching literature range from explicit treatment of individual items (Cobb, 2016; Laufer & Waldman, 2011) to a pure diet of incidental learning through extensive reading and/or listening (McQuillan, 2016; McQuillan & Krashen, 2008), as well as hybrid approaches (e.g. Nation, 2014). In addition to their lack of psycholinguistic credibility, the sheer size of the learning task and the limited time available in typical courses make all three prescriptions inadequate. A serious research program is needed to compare their effectiveness with that of various kinds of less intrusive, enhanced incidental learning (for discussion, see Long, 2017), a potentially more viable option, and with naturalistic exposure, e.g. through study-abroad programs. If salience, both objective and perceived, lies at the heart of most difficulties at advanced levels of L2 proficiency, albeit usually confounded with one or more additional factors, instruction at such levels will probably need to consist in large part of unobtrusive procedures capable of increasing the salience of the recalcitrant items for advanced learners.

REFERENCES


What Does Critical Period Research Reveal about Advanced L2 Proficiency?


Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? Reading in a Foreign Language, 26(2), 1–16.


5 Generative Approaches to Second Language (L2) Acquisition and Advanced L2 Proficiency

JASON ROTHMAN, 1,2 FATIH BAYRAM, 3 TANJA KUPISCH, 4,2 TERJE LOHNDAL, 5,2 AND MARIT WESTERGAARD 2,5

1 University of Reading  
2 UiT The Arctic University of Norway  
3 University of Portsmouth  
4 University of Konstanz  
5 NTNU Norwegian University of Science and Technology

Introduction

Relatively advanced proficiency in a second language (L2) is not uncommon; however, it is quite remarkable for an adult L2 speaker to become indistinguishable from native speakers. Undoubtedly, the apparent rarity reflects actual opportunity for success. The question thus becomes what underlies “opportunity for success” in this context. This question parallels the long-standing debate surrounding the application of the Critical Period Hypothesis (Lenneberg, 1967; Penfield & Roberts, 1959) to adult second language acquisition (SLA). Do ubiquitous differences that characterize child first language (L1) versus adult L2 development and ultimate attainment—which no one denies—follow from a biologically based critical period,1 after which adults are prevented from acquiring L2 competence in the same way L1 native speakers do? Is there a biologically induced age effect—in whole or in part—for accessing/using cognitive and/or domain-specific linguistic mechanisms that guide language acquisition in children? Conversely, is it possible
that both children and adults make use of the exact same mental (internal) mechanisms to acquire language, and that differences merely obtain as a result of extraneous factors that distinguish the learning tasks of the two sets of acquirers?

When considering the juxtaposed questions above, it is important to keep in mind what extraneous factors could be relevant and deterministic. Virtually all monolingual children (i) find themselves in an environment that inundates them with significant quantities of high-quality input from native speakers, (ii) will not have to deal with cross-linguistic influence/transfer from previously acquired languages affecting the path of development, and (iii) have a primordial need to acquire at least one language to help them encode the world around them and to use it to communicate their thoughts, needs, and desiderata. Alternatively, the typical adult acquiring an L2 (i) is not surrounded by quantities of native input anywhere near that of a child learning her L1, (ii) receives and must filter through significant amounts of non-native input (e.g. from classmates), (iii) has to deal with cross-linguistic influence/transfer from her L1 that can hinder as much as it facilitates, and (iv) does not have the same inherent need/intrinsic motivation for acquiring an L2 because her world is already completely encoded in linguistic terms and she can fully express herself in the L1.

It has been suggested that the above and/or other extraneous factors couple together in various constellations to explain specific subtypes of observable differences between children and adults. Such a view circumvents the need to claim that child versus adult acquisition processes are fundamentally different (e.g. Alemán-Bañón, Miller, & Rothman, 2017; Bialystok & Hakuta, 1994; Birdsong, 1992; Birdsong & Flege, 2000; Birdsong & Molis, 2001; Bruer, 1999; Epstein, Flynn, & Martohardjono, 1996; Flege, 1999; Hakuta, Bialystok, & Wiley, 2003; Rothman, 2008a; Schwartz & Sprouse, 1996, 2013; Slabakova, 2006; White, 1989, 2003, 2008). In Bialystok and Hakuta’s words, morphosyntax “remains accessible throughout life, even though the circumstances of our lives may muddy that access” (1994, p. 86). The claim is sustained by three distinct types of evidence, the first corresponding to the ageing brain in general and the second and third specifically to language acquisition/processing evidence from adults. To start, the whole notion of a critical period was originally predicated on the strong belief that the modular mind was deeply affected by loss of neurological plasticity after puberty, i.e. non-pathological, morphological changes to the structure of the adult brain affecting its ability to reorganize or redistribute function as needed. We know now, however, that the brain remains much more malleable throughout the lifespan than previously claimed, meaning there is no sharp decline in neurological plasticity culminating around puberty (see Fuchs & Flügge, 2014 for review). Evidence against (a strong version of) a critical period affecting sequential adult language acquisition comes from research showing that there are (i) very successful older learner exceptions to the general L2 acquisition outcome and (ii) significant behavioral and neurolinguistic/processing data that fail to show qualitative differences in the acquisition outcomes and/or real-time processing by adults for particular domains of L2 grammar for which measurable critical period effect are predicted.
It is possible, however, that adults simply no longer have the same abilities for language acquisition as children; that is, that the linguistic and/or cognitive learning mechanisms that guide child acquisition are no longer available—at all or to the same degree—after a specific time window in (early) adulthood (e.g. Abrahamsson & Hyltenstam, 2008, 2009; Bley-Vroman, 1989, 2009; Bylund, Abrahamsson & Hyltenstam, 2012, 2013; Clahsen & Muysken, 1989; Coppeters, 1987; DeKeyser, 2000; Granena & Long, 2013; Hawkins & Casillas, 2008; Hawkins & Chan, 1997; Johnson, 1992; Johnson & Newport, 1989; Long, 2005; Meisel, 2011; Schachter, 1988; Tsimpli & Dimitrakopoulou, 2007). Successes in adult L2 performance are acknowledged by researchers advocating a biologically induced age-related explanation for (some) L1 versus L2 differences. However, performance successes within L2 domains of grammar predicted to be unacquirable in adulthood—e.g. a property that has no equivalent counterpart in the L1—are understood as byproducts of adults’ generally high ability to learn patterns and/or their metalinguistic knowledge about the L2. In such a case, the underlying mental representations in child and adult learners would be different. Such a view is fueled by at least two factors: (i) there is, in general, an inverse correlation between increasing L2 age of onset and incidence of native-likeness overall (e.g. Abrahamsson & Hyltenstam, 2009) and (ii) the idea that any and all evidence of L1 versus L2 competence/performance outcome differences provides unassailable evidence for some type of critical/sensitive period. From such a perspective, the relatively few older L2 speakers who wind up indistinguishable from natives are considered anomalies or insignificant exceptions to an otherwise generalizable rule.

If evidence of any difference between children and adults is to be equated to a critical period effect, indeed there should be no debate on the matter, since no one denies the robustness of such evidence. However, we all of course care about the qualitative nature of why these differences obtain. In fact, unpacking the why itself will ultimately adjudicate between whether or not differences we observe reflect a true critical period for adult L2 acquisition. As pointed out by Rothman (2008a), claims about critical/sensitive periods map onto a “can versus cannot” as opposed to a “do versus do not” dichotomy. Furthermore, they make both global (general) and more fine-grained predictions. The fact that most adult L2 learners do not wind up indistinguishable from native monolinguals does not necessarily imply that they cannot acquire an L2 like native monolingual children. If the underlying cause of child L1 versus adult L2 differences truly is an outcome of biological (neurological) maturation, we could reasonably expect no exceptions at all since the claim is one of (im)possibility (neurological maturation makes it ‘cannot’) as opposed to predicting very low occurrence (as in typically ‘do not’). Global level predictions where any L1 versus L2 difference alone constitutes proper evidence in favor of critical/sensitive period(s) are well attested in the literature. At the more fine-grained level, however, things begin to break down. In particular, successes in the acquisition of discrete domains/properties of L2 grammar predicted to be impossible at any level of L2 proficiency pose significant challenges, e.g. complete target acquisition of domains of L2 morphosyntax, especially when accompanied by
native-like subtle semantic entailments that cannot be transferred from the L1 (Schwartz & Sprouse, 2013; Slabakova, 2006, 2016).

In the remainder of this chapter, we entertain the above possibilities with generative approaches to SLA (GenSLA) theories in mind, questioning how deep observations of L1 versus L2 differences go vis-à-vis data from high levels of adult L2 proficiency. To do so, we briefly introduce the main tenets of the generative paradigm as well as a cursory history of GenSLA and how they have, since the 1980s, addressed issues pertaining to advanced L2 proficiency and ultimate attainment. We argue and show that a unique type of data exclusive to GenSLA studies—specifically testing for Poverty of the Stimulus (PoS) effects in adult L2 acquisition—offers particularly clear evidence against the strong version of the Critical Period Hypothesis (e.g. Dekydtspotter, Sprouse, & Swanson, 2001; Dekydtspotter, Sprouse, & Thyre, 1999, 2000; Kanno, 1997; Montrul & Slabakova, 2003; Pérez-Leroux & Glass, 1999; Rothman & Iverson, 2008; Song & Schwartz, 2009). Aligned with the position that adult SLA is “Different? Yes. Fundamentally? No.” (White, 2008), we also introduce the reader to newer GenSLA models from the past decade or so that, in light of compelling evidence that Universal Grammar (UG) is still operative in adulthood, seek to explain why child L1 versus adult L2 as well as individual L2 speaker variation obtain even though there are no fundamental differences between child and adult language acquisition.

Main tenets of generative linguistic theory

Given the centrality of generative linguistic theorizing and its principal constructs to the GenSLA enterprise, e.g. Universal Grammar (UG) and PoS properties, it is important that we describe what they are and what motivates the proposals before continuing further. Of course, given the general readership and space limitations, we offer a somewhat simplified account. Akin to other brain-based modular subsystems that require external stimuli to unfold (e.g. vision), UG is argued to be a genetically endowed blueprint to the most generalizable facts about language; that is, it contains the linguistic information that is common to all human languages, labeled principles. As concerns linguistic learnability, the idea is that UG fills the gap left by what is learnable based on input and domain-general cognition alone. Equipped with UG, child learners are able to narrow down the search space for language learning by limiting their hypotheses about the target language from the superset of all logical possibilities to the subset UG allows; that is, only those that characterize potential human grammars. In listing a priori the limits on what is and what is not a possible grammar, UG has also been argued to identify and restrict the parameters of grammatical variation between languages (but see discussion below). Clearly, many domain-general cognitive, social, and computational principles shape linguistic development. According to the generative perspective, all these factors couple together to form and contribute to the task of organizing and making sense of input learners encounter (Rothman & Slabakova, 2017).
One well-known example of the interaction between principles and parameters is the Null Subject Parameter (NSP; Rizzi, 1982). The beauty of the original formulation of this parameter was that it assumed that grammatical properties swing in tandem (see D’Alessandro, 2015, for a recent review). Null subject languages correlated with the ability to extract the subject out of a finite embedded clause headed by an overt complementizer (1), they allow for subject inversion (2) (Kayne, 1980), and null subjects occur in languages with rich agreement (Taraldsen, 1980). Furthermore, both referential and non-referential (expletive) null subjects are licensed, which is not the case in non-null subject languages (3).

(1) ¿Quién1 dijiste que pro1/2 salió temprano? (Spanish)  
   who say-PRET.2.SG THAT leave-PRET.3.SG early  
   ‘Who did you say that left early?’ (Perlmutter, 1971, p. 103)

(2) a. È arrivato Gianni. (Italian)  
    b. *Est arrivé Jean. (French)  
    c. *Has arrived John. (English) (Roberts, 2007, p. 28)

(3) a. It rains frequently in April.  
    b. *pro rains frequently in April.  
    c. *Ello/Lo llueve a menudo en abril. (Spanish)  
    d. pro llueve a menudo en abril.  
   ‘It rains frequently in April.’ (Judy & Rothman, 2010, pp. 200–201)

A further property was added to this cluster by Montalbetti (1984). Consider the following data from Judy and Rothman (2010, p. 202).

(4) a. La niña1 cree que ella1/2 es la más inteligente. (Spanish)  
    b. La niña1 cree que pro1/2 es la más inteligente.  
       ‘The girl thinks that she/pro is the smartest.’  
    c. ¿Quién, cree que ella*1/2 es la más inteligente?  
    d. ¿Quién, cree que pro1/2 es la más inteligente?  
       ‘Who thinks that she/pro is the smartest?’

The data show that in Spanish, embedded null subjects have no absolute constraints for co-reference (see 4b and 4d). However, it is not possible for a variable overt embedded subject (quantified DPs or wh-words) to be co-referential with the matrix clause subject, although in general co-reference between a non-variable overt embedded subject (all other types of DPs/NPs) and the matrix subject is in principle possible (compare 4a–4c). That is, overt pronouns cannot be interpreted as bound variables, unlike pro. Montalbetti stated this as the following generalization and named it the Overt Pronoun Constraint (OPC).

(5) A bound variable interpretation of an overt pronoun is prohibited if pro is available in the same position.
Generative Approaches to Second Language (L2) Acquisition 77

English is then predicted to be different, since overt pronouns (the only ones applicable to English) can easily serve as bound variables.

(6) Who₁ thinks she₁/₂ is the brightest?

Knowledge represented by the OPC embodies an instance of PoS (more on this below). Evidence that (5) applies is not present in the input to the child, and it is not likely that it can be inferred from the available input either.

A particularly important aspect of the parametric line of thinking was to link easy-to-observe properties with hard-to-observe properties. For the pro-drop parameter, we see this in how extraction out of a finite embedded clause with an overt complementizer is linked to whether or not null subjects are allowed. It is easy for the child to observe whether or not null subjects are allowed, but it is not equally easy to observe the relevant restrictions on extraction. However, much work since the early 1980s has questioned the particular clustering that Rizzi (1982) argued for (see, among others, Haspelmath, 2008; Jaeggli & Hyams, 1988; Newmeyer, 2005; Rothman, 2009; Rothman & Iverson, 2007). Baker (2008, p. 352) states that “[h]istory has not been kind to the Pro-drop Parameter as originally stated.” Nevertheless, it serves as a good illustration of the original idea. Today, parameters are increasingly seen as emergent properties and not pre-defined by UG as such; only their format is (see e.g. Biberauer & Roberts, 2012; Westergaard, 2009, 2014).

Generative approaches to language acquisition have always been propelled by what is referred to as the logical problem of language acquisition. All typically developing children will be able to acquire language effortlessly despite input being impoverished or inconsistent at times. This suggests that language acquisition is guided by certain principles that are unique to language. Chomsky has always argued that there are innate principles specific to language that enable the child to acquire whatever language she is exposed to (see e.g. Chomsky, 1965, 1971, 1981, 1986, 1995; see also Fodor, 1983; Pinker, 1995). Support for domain-specific linguistic knowledge comes from the PoS argument (see Berwick, Pietroski, Yankama, & Chomsky, 2011; Schwartz & Sprouse, 2000, 2013, for updated reviews of PoS in general and its applicability to L2 acquisition). In essence, PoS concerns restrictions on structures in a given grammar that cannot possibly be derived from the input. PoS properties are argued to reflect “the innate schematism of mind that is applied to the data of experience” that, because there is no relevant information in the input to explain the resulting knowledge, “might reasonably be attributed to the organism itself as its contribution to the task of the acquisition of [linguistic] knowledge” (Chomsky, 1971, p. 26). Put differently, the input is insufficient in two senses: In scope since the input cannot provide evidence about all possible sentences that the child will encounter, and in quality because the input itself does not contain information about the kinds of representations that should be used in building a generative grammar of the language. The logic consists of the following steps, here based on Pullum and Scholz (2002, p. 18).
(7) a. Speakers acquire some aspect of grammatical representation.
b. The data the child is exposed to is consistent with multiple representations.
c. There is data that could be defined that would distinguish the true representation from the alternatives.
d. The data do not exist in the primary linguistic data.
e. Conclusion: The aspect of grammatical representation acquired in (a) is not determined by experience but by properties internal to the learner.

Space limitations prevent us from discussing the full range of evidence showing there are many such PoS instances with which children (and adults) are confronted. For a useful recent discussion, however, we direct the reader to Lasnik and Lidz (2017) for the case of L1 acquisition and Schwartz and Sprouse (2013) (as well as our discussion below) for how PoS also extends to the context of SLA.

**Generative approaches to L2 ultimate attainment**

The main focus of GenSLA has always been on describing and explaining the underlying second language grammar; that is, how the non-native grammar develops and ultimately manifests itself in mind/brain representations. A principled leitmotif of GenSLA studies, especially true in the first decades of its tradition (commencing in the early 1980s), sought to understand the interplay between universal knowledge (the role of/accessibility to UG in adulthood), knowledge that (potentially) comes from previous linguistic experience (L1 transfer), and knowledge that must have come from exposure to the target L2 (e.g. specific morphology and lexis). It is fair to say that the field of GenSLA started with a single main question: Do adult learners (post-puberty) continue to access UG as they acquire a second language? A second foregrounding question was soon added in parallel: What is the role of the L1 (transfer) in the process of second language acquisition? Both questions are of crucial importance and highly connected to one another; in fact, data needed to answer the first question are likely conditioned by the answer to the second. Given that the focus of this chapter, and indeed the entire handbook, is on advanced proficiency and ultimate attainment, it is prudent to point out that answers to each of the core queries also delimit the predictions for ultimate attainment. If it is the case that UG is inaccessible after the critical period, for example, then it would mean that particular L2 representations in adult acquired grammars are destined to be fundamentally different from L1 monolingual representations. Equally, if transfer from the L1 is prolific, if not complete, then this too would affect the entirety of the learning task for L2 acquisition and delimit the end result of the process. It would also partially predict what would be easier and more difficult to converge on—properties that are underlyingly the same between the L1 and L2 being easier, for example. Transfer would also account for some seeming successes without accessibility to UG in cases where knowledge could be transferred from the L1.
In this section, we will primarily refer to GenSLA research being conducted on advanced L2 learners and what, stemming from such research with this proficiency group, motivated the need for expansion beyond the two foregrounding questions. Before doing so, however, it is useful to remind the reader that the generative approach—any cognitive or other paradigmatic approach to SLA really—does not intend to address all questions worthy of serious consideration and investigation pertinent to SLA broadly speaking. For example, GenSLA, like connectionist/emergentist approaches, is not principally concerned with second language instruction, even though it might be able to offer insights to it and use data from instructed SLA to inform its theories (see Rothman & Slabakova, 2017). It is also useful to recognize that the main questions that pushed the GenSLA field forward in its beginning are partially theory internal, starting from the presupposition that there is an innate, domain-specific component to language, minimally operative in child language acquisition. As a result, much research has focused on providing answers to the question of UG accessibility in adulthood, which might at first glance appear irrelevant for researchers unconvinced by the necessity of a domain-specific linguistic module in the first place. Notwithstanding the potential answer to the theory-internal question of UG accessibility that studying discrete domains of L2 grammar can provide, the data themselves stand apart from the theory-specific question and are thus universally relevant. Collectively after 30 plus years, GenSLA studies provide a wealth of L2 data from corpora and very well designed experiments across an impressive array of grammatical properties and L1/L2 language combinations. The facts that such data represent must be explained parsimoniously by all theories claiming to cover the development and ultimate attainment of non-native grammars.

Access to Universal Grammar

Indeed, the most important research question of GenSLA during the 1980s through the late 1990s pondered whether or not UG was still accessible to adult learners. In the first iterations of this question early on, the answers were presented as dichotomous choices between “yes” and “no,” echoing most directly the Critical Period (CP) Hypothesis debates at the time. In other words, the generative linguistic equivalent to the critical period was essentially maturationally conditioned inaccessibility to UG. Degrees of success in L2 acquisition were not easily accommodated by the dichotomous choice. While it was taken as given that UG was operable in early child language acquisition (L1, L2, L3/Ln), determining whether adults indeed had continued access to UG or not was contentiously debated (e.g. Bley-Vroman, 1989; Clahsen & Muysken, 1986, 1989; Flynn, 1987; Schachter, 1988; Schwartz, 1987; White, 1989). UG inaccessibility claims from this era were critiqued for potentially jumping to premature conclusions based on limited types of data, primarily global differences between L1 and L2 developmental sequences and/or ultimate attainment. Early work taking a “no UG access” position had not yet fully considered the role that L1 transfer and other extraneous factors might play in obscuring any straightforward interpretation of L1/L2
differences. In the context of ultimate attainment and the very ability to achieve truly native-like knowledge of an L2, from a generative perspective at least, the answer to this debate is crucial in determining whether or not an L2 can in principle be fully acquired.

Some of the best, if not unassailable, evidence that accessibility to UG does not suffer from a critical period is L2 PoS competence: Some advanced L2 learners exhibit highly specific knowledge of linguistic restrictions that could not have been transferred from the L1, whose acquisition based on cues from the L2 input is seemingly impossible, and which is not taught in L2 classrooms (see Schwartz and Sprouse, 2013, for review). This type of evidence is left unexplained by domain-general cognitive considerations. Starting in the 1990s, a shift in GenSLA research occurred whereby probing directly for L2 learner PoS properties took center stage, mostly examining advanced stages of L2 proficiency. Kanno (1997), Pérez-Leroux and Glass (1999), Dekydtspotter and Sprouse (2001), Dekydtspotter and Sprouse, and Anderson (1997), and Rothman and Iverson (2008) among many others since, have shown that L2 grammars, despite significant differences from L1 grammars, are characterized by PoS effects, especially at advanced levels of proficiency.

Even in light of such evidence in significant quantities across many language pairings, there is no consensus in GenSLA that adults have direct access to all of UG; that is, the entirety of what UG provides to children. For example, Tsimpli and Dimitrakopoulou (2007) and Hawkins and Casillas (2008) take the position that while UG is indeed accessible in adulthood, all of the universal features it provides are not. According to both, only interpretable features—those with meaning (semantic) content—remain accessible, while uninterpretable ones—those relevant to (functional) syntactic operations, e.g. a feature that results in movement—are no longer available from UG and thus will not be instantiated in L2 grammars unless they can be transferred from the L1. Such proposals predict that even at the most advanced stages of L2 development, L1 and L2 grammars will reflect some fundamental differences in representation. However, what will be unacquirable is significantly reduced in comparison to what was claimed by theories of inevitable representational differences from the 1980s and 1990s. Under current accounts, at least some differences between adults and children must be explained on the basis of something other than UG (in)accessibility.

The role of L1 transfer

Despite UG accessibility (in whole or in part), there is no question that developmental path and ultimate attainment in adult L2 acquisition are observably different from child L1 acquisition. If it were the case that L1 transfer alone could explain all or most of the differences, then the second main line of query historically within GenSLA, that of determining the type, scope, and extent of L1 transfer effects, would provide sufficient explanation for the ubiquitous variation in SLA. It is clear that L1 transfer changes the overall learning task for adult L2ers, and it should therefore not be surprising that child L1 and adult L2 developmental
sequences differ. L1 transfer does explain a good amount of L2 variation, but certainly not all. Moreover, simultaneous bilinguals, i.e. child learners with two first languages, may also show acquisition orders that differ from those of monolinguals, which invalidates diverging acquisition orders as a diagnostic of critical period effects.

To highlight the explanatory value of L1 transfer as a main variable of difference, let us consider a classic example of this within an approach to L2 acquisition as parameter resetting. Previously, we discussed the Null Subject Parameter, which refers to the difference between languages like English, German, and Norwegian, which require the subject position be overtly filled, and languages like Spanish, Turkish, and Arabic, which allow for subjects to be either null or overtly expressed. At the initial state of L1 acquisition, parametric values are by definition underspecified, and any universally conditioned outcome (i.e. parameter settings or values) is equally available until exposure to a specific language (the L1) provides the child with sufficient evidence to choose the correct value pertaining to the target language. A child exposed to Spanish, for example, simply has to match one of the values of the NSP with the input of her exposure. Hearing a plurality of sentences in which there is no overt subject will lead her straightforwardly to the conclusion that Spanish is positively valued for the NSP. Alternatively, a child exposed to English does not hear sentences in which the subject position is ostensibly empty and therefore sets the parameter value to the minus setting.

To the extent that L1 transfer obtains, the learning task for an L1 Spanish learner of L2 English and for an L1 English learner of L2 Spanish for the NSP is not nearly as straightforward as that of a child learning his/her L1, nor are they equal to each other because the directionality of acquisition itself can also delimit the L2 learning task (White, 1985). A transferred L1 value means that parametric settings are already fully specified from the beginning of L2 acquisition, whereas in the case of child L1 they are by definition unspecified. Thus, the earliest stages of the L2 interlanguage are already specifically valued and this will not match the target input. The learning task is considerably more challenging than what the child has to do because the L2 learners need to undo and construct as opposed to simply construct. Moreover, as it turns out, not all “undoing” is equal. In the case of L1 English → L2 Spanish, the task is relatively uncomplicated and much closer to that of a child monolingual than in the L1 Spanish → L2 English direction. This is the case because resetting the parameter value in the former case requires expansion from a smaller (subset) value to the larger (superset) value, whereas the latter requires reconfiguration from a larger value to a more restricted (or subset) grammar. Spanish is considered to be the superset grammar because it allows the English requirement—overt expression of subjects—but, unlike English, also allows sentences with a null subject, provided that discourse requirements are met. Following from the Subset Principle (Manzini & Wexler, 1987), it is easier to expand a grammar than it is to reduce one from a setting that is too large. This is the case because there potentially would not be enough evidence to motivate grammatical restructuring in the superset → subset direction (i.e. cues that lead to choosing one setting over another might not be identifiable by the parser as a
result). An English speaker exposed to L2 Spanish will be confronted with a plurality of utterances that the transferred L1 value cannot handle (sentences with null subjects). In the case of Spanish natives exposed to L2 English, on the other hand, there will be fewer relevant parsing failures, as Spanish also allows sentences with overt subjects. The NSP is likely the most well documented parameter to have been studied in GenSLA over the past three decades, and L1 English → L2 Spanish is probably the most studied language pair. Work by Hilles (1986), Phinney (1987), Judy and Rothman (2010), and Judy (2011) examined the opposite direction (L1 Spanish → L2 English), and in some cases they compared and contrasted both directions under the same methodology. Whereas most work on L1 English → L2 Spanish has found evidence for early L2 acquisition of null subjects, the opposite direction does not fare as well. Earlier work by Hilles (1986) and Phinney (1987) already showed an advantage for the English → Spanish direction. Judy (2011) shows that although L1 Spanish speakers who are near-native L2 English speakers are quite good at rejecting referential null subjects in English overall, they are not nearly as good at rejecting ungrammatical null expletives. Judy’s explanation is that the transferred L1 value either cannot be reset due to its subset-superset relationship in this direction or it is much slower to be reset than the case of English → Spanish, as the evidence for resetting is much more subtle and thus takes longer to reach a threshold for restructuring.

The main point that this NSP example illustrates, beyond the obvious one pointing out the deterministic nature of L1 transfer, is that adult L2 learners do not all start the process of L2 acquisition at the same point. Moreover, it shows that the formal learning task for the same domain of grammar can vary considerably depending on the L1 starting point, even if all L2 learners have access to the same internal mechanisms that children do. Simply put, beginning the process with linguistic and other experiences alters the nature of the task. If L1 transfer can change the learning task, then there are likely other (experience-based) variables that contribute in similar ways to overall L1/L2 differences in ultimate attainment. Examining the extent to which L1 transfer and other experiences/factors differentially manifest themselves in L2 acquisition and thus, at least partially, explain L1/L2 differences, must surely be part of all cognitive-based SLA approaches. More widespread recognition of this within GenSLA since the late 1990s has spawned newer theories that seek to understand L2 variation, beyond UG accessibility and effects stemming from L1 influence. The next section is dedicated to these theories.

**Beyond UG access and L1 transfer**

By the mid-1990s it was exceedingly clear that the two main questions guiding GenSLA research would need to be significantly supplemented. Despite the observable fact that in many ways adult L2 acquisition differs from child L1 acquisition in quantitative and qualitative ways, the then 20 years of GenSLA research provided clear evidence that L2 interlanguage grammars/performances (i) instantiate abstract knowledge about the L2 that could not have been acquired on the
basis of the L2 input, (ii) reflect transfer from the L1 throughout the developmental process although L1 transfer alone cannot explain all L1/L2 disparities at any snapshot stage of L2 acquisition, and (iii) demonstrate that instruction and metalinguistic knowledge do affect performance, but seemingly have little effect on competence per se. However, although descriptive adequacy was high, explanatory adequacy—actually explaining how and why what was describable about L2 acquisition obtained—was far off from being achieved. The time was ripe to expand beyond the traditional foregrounding questions. That is, a critical mass of evidence from GenSLA studies (and indeed data from other paradigmatic approaches to SLA) made it clear that other variables/considerations needed to be incorporated better in GenSLA models to fully describe and explain the dynamic nature of adult L2 acquisition, especially fossilization at advanced stages of L2 proficiency. Although we would argue that GenSLA researchers have always been concerned with other variables than UG accessibility and L1 transfer effects (e.g. quantities and qualities of L2 input, differences in processing capacities between monolinguals and bilinguals, mapping difficulties between function and form and the like), it was not until the early 2000s that the call to understand the role of such variables became part of the core research program.

At the very turn of the millennium, several researchers had shifted their attention to the generally poor production of morphological paradigms in L2 acquisition, especially in otherwise very advanced and highly competent L2 learners (see Lardiere, 1998a, 1998b, 2007, for a very detailed case study example). This was especially true for researchers arguing that UG still directly guides adult L2 acquisition. After all, it is incumbent on such researchers to explain how the process of adult and child acquisition can be fundamentally equivalent with respect to access to UG while the developmental processes and outcomes of acquisition are so variable for adults, yet relatively invariable for children. While there are sure to be multifarious explanations for specific subtypes of L1 versus L2 differences—i.e. it is unlikely that there is a singular cause for all L1 versus L2 differences—it seemed prudent to focus on L1/L2 morphological disparities for several reasons, not the least of which was the robustness and ubiquity of the problem. The problematic case of L2 morphology is also particularly interesting because morphology is overtly and continuously taught to tutored L2 learners, it is reliably provided in available input and, depending on the specific morphemes, it may even be highly frequent. Thus, the fact that obligatory inflectional morphology—an overt phonological expression of functional syntactic categories—stands out as highly variable in adult SLA production, needed to be reconciled with (i) evidence of successes in L2 morphosyntactic PoS domains and (ii) high variability despite substantial frequency in the input and even metalinguistic knowledge.

On a theoretical level, understanding how and why L2 morphology presents such an obstacle for advanced proficiency L2 speakers was especially interesting given the assumed connection between morphology and syntactic development in the case of child L1 acquisition. Many L1 researchers at the time had been working under the assumption/proposal that overt morphology is directly linked to underlying syntax in acquisition, in fact, it was considered to be the driving
force for syntactic acquisition. Failure to show consistent and stable knowledge of morphological exponents was taken to indicate lack of acquisition of an associated functional category, whereas the point of eventual stability was understood as marking acquisition of the corresponding functional category (e.g. Clahsen, Penke, & Parodi, 1993/1994; Radford, 1990). Under an extension of this morphology-before-syntax account to L2 acquisition, omission and/or commission errors should be and were understood as a causal byproduct of a breakdown in the otherwise robust triggering relationship of morphology and syntax. Could it be, however, that the apparent alignment of morphology and syntax in children was essentially coincidental, by which morphology does not actually drive syntactic development? This was argued already by the Separation Hypothesis (e.g. Beard, 1987, 1995). The Separation Hypothesis maintained that morphosyntactic features can be present in a grammar, as seen by various reflexes in grammatical behavior, even when the corresponding morphology is absent or variable. Evidence included, for example, the syntactic features of empty (null) categories and the variable use of overt functional morphology in distinct registers and dialects of language heads. Potentially, adult L2 acquisition could be used to address this debate more generally. At the same time, a (newer) understanding that morphological production does not necessarily faithfully reflect the underlying syntactic representational system, if defensible on empirical grounds, might be able to parsimoniously reconcile (some) facts that otherwise sat in contradiction. And so, a syntax-before-morphology or a more general (partial) disassociation/separation between morphophonological output and morphosyntactic representation agenda was born in GenSLA (e.g. Lardiere, 2000; Prévost & White, 2000).

At the level of mere observation, L2 morphology presents rather differently than L1 morphological development. In typical child L1 acquisition, morphological issues tend to reflect developmental errors of omission in obligatory contexts that are eventually overcome. Throughout L2 development, on the other hand, errors of both omission and commission occur, the latter being the suppliance/production of wrong morphophonological agreement (e.g. a second-person verbal morpheme with a third-person subject). Although errors of both types lessen over time, they do not disappear completely even at advanced levels of proficiency. Despite the fact that the general trend applies to all L2 learners irrespective of the L1, the status of the L1 in terms of its morphological inventory (inclusive of the syntactic features selected and bundled on its morphophonology) matters as well. That is, in the domain of obligatory past inflection in English (−ed and its allomorphs), for example, Chinese natives produce even less overt past morphology (significantly less perhaps at the level of advanced L2 proficiency) than say German or Spanish speakers, because only the former lack such morphology in the L1 (e.g. Cabrelli Amaro, Campos-Dintrans, & Rothman, 2017; Hawkins & Liszka, 2003).

One of the first ideas put forward in the GenSLA syntax-before-morphology thread of research of the early 2000s is captured under the Missing Surface Inflection Hypothesis (Haznedar & Schwartz, 1997; Prévost & White, 2000), which basically states that morphological production even in highly advanced L2 speakers, given processing and other pressures on production, can under-represent underlying
syntactic representations. Evidence in support of this would be advanced L2 learners’ comprehension of functional morphology being demonstrably better than their variable production. White (2003) reviews a good deal of the relevant empirical research available at the time, showing that there is very good reason to believe that morphological productions can indeed underestimate syntactic competence. GenSLA studies continue to show this even today. Since the early 2000s, several other hypotheses have emerged, related to the same basic problem of explanation, including the Feature Reassembly Hypothesis (Lardiere, 2009), the Prosodic Transfer Hypothesis (PTH; Goad & White, 2004, 2006, 2009), and the Bottleneck Hypothesis (Slabakova, 2008). Paralleling changes in formal linguistic theory pertaining to the role and granularity of functional features as a basic unit of linguistic representation, the Feature Reassembly Hypothesis places the L2 problem at the level of complexities inherent to configuring the right mappings of feature configurations (or bundles) to specific lexical items in the L2. The main idea is that L2 learners have access to the entire inventory of UG features, not only the ones instantiated in the L1 functional lexicon. However, assembling the features in the correct configurations might prove extremely challenging for adults, not the least because the L2 will bundle features differently than in a speaker’s L1. The extent of difference in these mappings between the L1 and the target L2 represents one of the main learning tasks of L2 acquisition.

The PTH (Goad & White, 2004, 2006) follows from the same desire to try to explain why L2 morphology is such a vulnerable domain. It highlights, however, a distinct aspect to the feature mapping issue above, namely, the role that L1 phonological transfer can play in L2 morphological production. Essentially, the PTH claims that if L2 target morphology runs in disaccord to L1 constraints on phonological production (e.g. phonotactics) and/or the prosodification of the morphology to lexical words is different in the L1, then this can give rise to variation in L2 production. In other words, the underlying syntax that the morphology represents could be target-like (whether acquired during the course of L2 acquisition or simply transferred from the L1), but residual effects of L1 phonological constraints could negatively affect the production of L2 morphology. Let us consider again past tense morphological production in L2 English. We know from many studies that Chinese speakers of L2 English variably produce -ed in obligatory contexts (Hawkins & Liszka, 2003; Lardiere, 1998a, 1998b, 2007). According to the PTH, the reason for this, at the advanced level of proficiency, is not that Chinese learners cannot acquire the uninterpretable past tense feature mapped onto -ed (lacking in their L1), but rather because of L1 phonological transfer. Chinese does not permit complex codas and one of the most common allomorphs of -ed is [kt], as in walked, talked, and stalked. Moreover, Chinese does not have prosodic word adjunction, which is the process by which past tense morphology is adjoined to the root in English. If Goad and White are on the right track, not only would L2 speakers of English whose L1s lack past tense morphology have problems with past tense production, but indeed any learner whose L1 phonology could cause similar influence. In a recent paper, Cabrelli Amaro et al. (2017) tested this claim by comparing Chinese, Japanese, and Spanish highly advanced L2 speakers of English. Although Spanish and Japanese...
do have the past tense feature (unlike Chinese), both languages also have the same general prohibition on complex codas, and Spanish is also lacking prosodic word adjunction. The results show that all three groups variably produce English past tense, with no differences between the rates of suppliance in Spanish and Chinese speakers. The Japanese group, while still different from the native English controls, fared better overall, which the authors attribute to them only needing to overcome phonotactic constraints, whereas the other groups had more L1 effects to deal with. Another example showing that L1 prosodic structures may affect the production of L2 morphology provided by Goad and White (2009) relates to the use of English articles by L1 Turkish speakers. According to Goad and White, L1 Turkish speakers might have problems producing articles in L2 English simply because their L1 does not provide a prosodic template for that. The result is that in certain prosodic contexts, speakers either omit articles or stress them, as word stress would allow a prosodic representation that is available in Turkish (i.e. when they use the stressed numeral bir ‘one’). Snape and Kupisch (2010) measured the English articles produced by an advanced L2 English L1 Turkish speaker and found that indeed many of her articles were stressed.

Taken together, the above hypotheses related to morphological variability in L2 acquisition represent the growth of GenSLA studies—the pursuit of questions that complement the original ones that shaped the first decades of the paradigm—while still following the insights from and general trends in modern generative theory. There have been many other hypotheses put forward over the past two decades within GenSLA studies whose goals are similar, to describe and explain the reality of how and why L1 and L2 manifest so differently, without assuming that observable differences must mean that the processes are fundamentally different. Space does not permit us to go over other hypotheses in great detail; however, it is worth mentioning a few here so that the interested reader can look at them in greater detail. The Competing Systems Hypothesis (Rothman, 2008b), for example, highlights the possibility that in classroom SLA, the growth of a robust metalinguistic system (learning) in parallel with a system of linguistic competence (acquisition) can result in competition (in advanced learners specifically) with pedagogical rules taught that are less than accurate in linguistic terms. The Interface Hypothesis (Sorace, 2011), as another example, has been highly influential in recent years. This hypothesis draws our attention to the possibility that finite processing and attention resources that must be divided in bilinguals between activating and inhibiting the two languages can give rise to residual optionality in L2 performance, even when the underlying representations are otherwise demonstrably target-like.

**Conclusion**

This chapter started out with the observation that it is quite rare for adult L2 learners to achieve a level of proficiency that makes them indistinguishable from native speakers. Since all typically developing children do achieve native
proficiency in their L1, the difference between the two populations is obvious and undisputed. One important question in GenSLA is whether the two processes are fundamentally different, in that L1 children have access to an innate language learning mechanism (UG) that is no longer available for adults, or whether the attested differences are due to extraneous factors such as L1 transfer or lack of input and motivation. A fundamental difference would correspond to what is often formulated as a critical period for language acquisition early in life, after which there is a biologically determined cut-off point for access to the innate endowment for language. According to this view, it should be impossible for any adult learner to achieve complete native-like mastery of an L2 at a macro level as well as, at the individual level, for all adult L2 learners to acquire particular properties in the L2 (e.g. any property for which completely new L2 morphosyntactic features would need to be acquired). In this chapter, we have provided a discussion of these issues within GenSLA, taking into account certain historical changes in perspective that have gone hand in hand with developments within generative theory. Thus, while there was a focus on access to UG and the question of full, partial, or no transfer within the principles and parameters approach of the 1980s and 1990s, the field has more recently moved on to consider more diverse and fine-grained issues, e.g. the mapping of functional features between the L1 and the L2 and the consistent challenges of morphology in L2 acquisition. Furthermore, there is increased focus on the effect of variables such as qualitative/quantitative input factors and processing differences.

While there is healthy discussion and considerable disagreement with respect to these issues within the field of GenSLA, we have nevertheless been clear about our own perspective. Given current knowledge, we find convincing evidence that there is no biologically determined critical period for second language acquisition: Adult L2 learners have been found to acquire abstract knowledge that goes beyond the input (so-called PoS effects), indicating that there is continued access to the language learning mechanism in adulthood. Furthermore, to our knowledge, there is no single linguistic property that has been shown to be completely unlearnable in L2 acquisition. This means that the admittedly robust and commonly attested differences between children and adults must be due to L1 transfer and other extraneous factors. Thus, while the field continues to investigate these differences in order to increase our understanding of L1 and L2 acquisition, the two processes should nevertheless be considered to be fundamentally the same.

Acknowledgment

This research was partly supported by a grant from the Research Council of Norway for the project MiMS (Micro-variation in Multilingual Acquisition & Attrition Situations), PI Marit Westergaard, project number 250857.
NOTES

1. For ease of exposition and given limited space, we use the term critical period in a catchall way. That is, critical period is used herein as a proxy for either a singular absolute critical period for all domains of grammar, multiple critical periods applying at different times to distinct domains of grammar (e.g., an earlier one for phonology as compared to syntax), or even sensitive periods that apply gradually over time as opposed to a sharp critical period effect. This is because explaining the differences and the literature that debates them is peripheral to the points at hand, although we acknowledge that distinguishing between them for other important issues matters a great deal.

2. We wish to acknowledge that there is considerable debate as to what various generative theorists take to be the constitution of the innate language faculty. For some, the genetic linguistic endowment is truly minimal, perhaps only containing a single, overarching core operation that is truly domain-specific to language such as (recursive) Merge (e.g., Yang, Crain, Berwick, Chomsky, & Bolhuis, 2017) or a more generalized conception of syntactic recursion (e.g., Hauser, Chomsky, & Fitch, 2002). For others, the genetic endowment could not be so reductionist in light of non-recursive elements of language (phonology, morphology, and specific elements of syntax itself, for example) that seemingly cannot be explained as being specific to humans but not to language (they cannot follow from domain-general cognition) or not being specific to humans (they are attested in other animals, especially primate species) (e.g., Pinker & Jackendoff, 2005).

3. See Meisel (2011) for a more strict age-related divide between child and adult L2 acquisition with a cut-off age for fundamental differences for some domains of grammar (e.g., grammatical gender), argued to be as early as age 3–4.

REFERENCES


6 Interaction-Driven L2 Learning: Advanced Learners

NICOLE ZIEGLER¹ AND LARA BRYFONSKI²

¹ University of Hawai‘i at Mānoa
² Georgetown University

Introduction

Since Long’s first iteration (1981) and subsequent updates of the Interaction Hypothesis (1996), a growing body of second language acquisition (SLA) research has demonstrated a robust positive connection between conversational interaction and second language (L2) development (Mackey & Goo, 2007). Having evolved from a hypothesis to an extended framework (Gass & Mackey, 2006, 2007; Mackey, 2012), interaction research has expanded from initial examinations into the impact of conversational adjustment on learners’ L2 development (Gass & Varonis, 1994; Mackey, 1999) to investigations into a wide range of interactional factors and processes, including the role of a variety of individual differences such as working memory (e.g. Mackey, Philp, Fujii, Egi, & Tatsumi, 2002), aptitude (e.g. Trofimovich, Ammar, & Gatbonton, 2007), anxiety (e.g. Sheen, 2008), and motivation (e.g. You & Dornyei, 2016). Scholars have also examined the role of L2 proficiency (e.g. Watanabe & Swain, 2007), with results demonstrating a complex relationship amongst features of interaction, such as that between learners’ noticing of corrective feedback and advanced L2 proficiency (e.g. Lee, 2013; Li, 2014; Mackey & Philp, 1998).

Beginning with an overview of the interaction approach to SLA, including a brief history of the approach and its findings in relation to learner proficiency, this chapter examines interaction research that has made connections to advanced proficiency learners, including studies of corrective feedback, peer interaction,
The interaction approach

The interaction approach to second language acquisition (Gass, 1997; Gass & Mackey, 2006; Long, 1996) posits that the combination of exposure to modified input, output, and negotiation for meaning via the provision of positive and negative corrective feedback is essential to L2 development. This link between interaction and L2 development has been empirically tested in over a hundred studies since the 1980s (Abbuhl, Ziegler, Mackey, & Amoroso, 2015), providing robust evidence for positive benefits on a variety of grammatical and discourse features within a variety of contexts and proficiency levels (see Cobb, 2010; Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Mackey & Goo, 2007; Ziegler, 2016a for meta-analyses).

Key elements of the interaction approach

Input, which can be defined as the language that a learner is exposed to through listening, reading, writing, speaking, or other mediums, is a key component of the interaction approach. Operationalized as the positive evidence learners receive about the target language (Gass, Mackey, & Pica, 1998), input can be modified during interaction to be made more comprehensible, thereby adjusting to the needs of learners of varying proficiencies during the negotiation (Long, 1981; Mackey, 2012). Previous research has sought to examine the ways in which interlocutors modify the input to make it more comprehensible to language learners (Krashen, 1977, 1980), thereby providing learners the opportunity to confirm or reject their own hypotheses about what is possible in the target language. For example, studies have found that more advanced learners may be better able to leverage their attentional resources to notice features in the input than lower-proficiency learners (Gass, Svetics, & Lemelin, 2003), providing further opportunities for L2 development. However, as essential as input is in the process of language acquisition, Mackey (2012) points out it is the ways in which learners interact with input and their interlocutors through the interactional processes of negotiating for meaning, giving and receiving corrective feedback, and producing modified output, that are likely to lead to L2 development (Mackey, 2012).

Output, another key component of L2 development in the Interaction Hypothesis, is the language that learners produce themselves during interaction. Swain (1995, 2005) argues that the opportunity to produce language and correct non-target-like production after feedback allows learners to test out hypotheses regarding the target language, and directs them to attend to target language forms and to notice any gaps between their interlanguage and the target language,
thereby potentially promoting fluency and automatization (Swain, 1995). This modified output, or pushed output, which is defined as the reformulation of a learners’ utterance in response to feedback or self-monitoring (Mackey, 2012), is believed to facilitate L2 development by driving learners to modify their production in a more target-like manner (McDonough, 2005; McDonough & Mackey, 2006; Shehadeh, 2002; Swain, 1995, 2005). However, the process of modifying output can be as important to language development as the actual modification, with learners still benefiting whether they produce the correct target forms or not.

Negotiation for meaning, or adjustments made during conversation to achieve understanding between interlocutors, involves the elements of input, output, and corrective feedback, and is a key component of the interaction approach. Negotiation during interaction has been shown to facilitate language learning by allowing learners to receive comprehensible input and to modify their own output to be better understood (Mackey, 2012). Confirmation checks, clarification requests, comprehension checks, repetitions, prompts, and segmentations of words or phrases are all strategies interlocutors might employ while negotiating for meaning (Mackey, 2007). Negotiation might also take the form of corrective feedback, which provides negative evidence to the learner. Negative evidence, defined as input that supplies direct or indirect evidence of ungrammatical forms to the learner (Leeman, 2003, 2007), indicates to learners that there was an issue with their language production, potentially drawing their attention to gaps between their interlanguage (IL) and the target language (TL), thereby leading to L2 development. In addition, negotiation may prime learners to be more attentive to future input, raising their awareness of specific features of the target language and providing them with multiple opportunities to confirm or disconfirm hypotheses they have formed regarding the L2. Previous studies have also demonstrated a link between negotiation for meaning and proficiency level (Oliver & Mackey, 2003), with higher-proficiency learners tending to negotiate more with native speaker interlocutors than those with a lower proficiency.

Recent research has expanded the construct of negotiation to include interactional modifications occurring in response to other forms of implicit and explicit feedback, including recasts, which are a form of corrective feedback where the interlocutor rephrases all or part of a learner’s immediately preceding non-target-like utterance while the focus remains on meaning rather than form or object (Long, 2007; Mackey & Gass, 2015; Richards & Schmidt, 2002) and metalinguistic feedback (Mackey, 2012). Overall, research has shown that receiving feedback and participating in negotiation may support learners’ L2 development by providing both positive and negative evidence, with a growing body of research having demonstrated the positive effects of different types of corrective feedback on a wide range of features of L2 development in both classroom and laboratory settings (see Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007, for meta-analyses).

Although these studies have shown that both implicit and explicit forms of corrective feedback facilitate L2 development in a range of language learning contexts, the comparative efficacy of explicit feedback, such as metalinguistic feedback,
or more implicit forms of feedback, such as recasts, has been hotly debated in the current literature (Ellis, Loewen, & Erlam, 2006; Goo & Mackey, 2013; Lyster & Saito, 2010). For example, research suggests that recasts are facilitative of L2 development as they enhance the salience of the target feature and direct learners to contrast their erroneous utterance with their interlocutor’s reformulation (e.g. Goo & Mackey, 2013; Long, 1996, 2007), thereby focusing learners’ attentional resources on the target form. Importantly, regardless of efficacy compared to more explicit forms of corrective feedback, recasts are the form of feedback most often used by teachers in advanced proficiency classrooms (Zyzik & Polio, 2008). Some scholars suggest that this is due to the fact that implicit feedback does not interrupt the flow of meaning-making in the classroom and is more natural (Ellis & Sheen, 2006). However, studies have demonstrated that the influence of feedback, such as recasts, on L2 development is mitigated by a variety of factors, including target feature (e.g. phonology, morphosyntax, e.g. see Egi, 2007) and setting (Oliver, 2000), as well as learners’ individual differences (Mackey, 2012), such as the learner’s proficiency or developmental level (Mackey & Philp, 1998).

Second language proficiency, which can be defined as a learner’s linguistic ability in a given domain (e.g. speaking, listening, reading, writing) based on a given outcome measure, has been operationalized in a variety of ways in the interaction literature. Some studies identify learners as advanced by their current institutional status, such as by their enrollment year in a given language program (Li, 2009), by standardized tests, such as the TOEFL (Test of English as a Foreign Language) or TSE (Test of Spoken English; Lee, 2013) or the American Council of Teaching Foreign Languages (ACTFL) Oral Proficiency Interview (OPI), or by using independently designed objective and subjective measures (i.e. custom-made assessment targeting specific aspects of proficiency or self-reports or instructor ratings). Many studies have also investigated the relationship between interaction and proficiency, with the prototypical study dividing students into groups according to proficiency level, as measured by one of the possibilities above, and examining the impact of an interventional treatment by measuring an individual or multiple learning outcomes with a pre-, post-, or delayed post-test design. In order to triangulate findings, some studies additionally incorporate introspective measures of noticing or awareness, such as stimulated recall interviews (see Gass & Mackey, 2016), uptake sheets (Allwright, 1984), or other qualitative methods. The following section describes studies that have made connections to advanced proficiency learners in the domains of corrective feedback, peer interaction, TBLT, and computer-mediated communication, utilizing these methodologies.

Interaction and advanced learners

Previous studies in the field of interaction have sought to identify the types of interactions that most benefit learners at various levels and have suggested a need to alter the elements of interaction, such as corrective feedback, in relation to
students’ proficiency levels (Han & Jung, 2016; Panova & Lyster, 2002; Suzuki, 2005). For example, researchers have examined the relationship between the quality or quantity of interactional features and the explicitness of corrective feedback and uptake at various proficiency levels (e.g. Mackey & Philp, 1998). In addition, studies have investigated to what degree learners of different proficiency levels are more or less likely to notice feedback (Gass, 1997; Philp, 2003), as well as what kinds of feedback instructors of advanced learners tend to prefer and provide (Zyzik & Polio, 2008). Other studies have examined how proficiency plays a role when peers interact with each other (e.g. Kim & McDonough, 2008; Leeser, 2004). For example, research suggests that learners’ proficiency level may have an impact on the quantity and quality of interaction, with findings indicating that pairing high-proficiency learners with lower-proficiency learners may lead to more collaboration (Storch, 2001; Yule & Macdonald, 1990). However, Kowal and Swain (1997) found that lower-proficiency learners felt less comfortable interacting with advanced proficiency learners than with learners of similar proficiencies, suggesting that proficiency may impact both affective and cognitive factors. These studies have also examined the connections between proficiency and interaction in a variety of settings, such as classroom and computer-mediated settings (e.g. Sauro, 2009) and during task-based interactions (Ellis, 2009), as well as the relationship between individual factors, such as proficiency, and interactional features, like corrective feedback.

**Corrective feedback and advanced proficiency learners**

Previous work has uncovered a complex relationship between the effects of corrective feedback and L2 proficiency (see Nassaji, 2013). Some studies suggest that less proficient learners benefit more from explicit corrective feedback moves, such as metalinguistic explanations and prompts, than higher-proficiency learners (Ammar & Spada, 2006; Trofimovich et al., 2007). For example, Lin and Hedgcock (1996) demonstrated that learners receive sufficient amounts of feedback according to their own stage of development, with low-proficiency learners receiving more feedback than high-proficiency learners. However, high-proficiency learners were more likely to demonstrate uptake, or the integration of new features into their own interlanguage, of the feedback they received than their lower-proficiency peers. These findings and others (Williams, 2001) suggest that higher-proficiency learners might stand to benefit more from certain types of corrective feedback in interaction-driven learning than low-proficiency learners, due to the availability of more attentional resources (Gass et al., 2003), prior knowledge, and linguistic experiences (Philp, 1999).

For example, Li (2009, 2014) has examined the interplay between the effects of corrective feedback and L2 proficiency in a variety of studies with foreign language (FL) learners of Chinese, finding relationships between proficiency level and the explicitness and target of feedback. In Li (2009), 23 learners from two different proficiency levels, as determined by their enrollment year in the Chinese program, received either implicit feedback in the form of recasts or explicit
metalinguistic explanations. The students engaged in communicative tasks and received the feedback on their production of Chinese classifiers. Results demonstrated that proficiency influenced the efficacy of the different forms of feedback for the lower-proficiency learners, in that they benefited more from explicit feedback than implicit; however, the advanced learners benefited from both explicit and implicit feedback. These findings are consistent with the results of Mackey and Philp (1998), who found that for more advanced learners, negotiation for meaning with recasts was more effective than negotiation alone. However, the same feedback move, recasts, was not effective for less advanced learners. In 2014, Li updated his study to investigate the role of the linguistic target in mediating this relationship with 78 Chinese learners learning two different target structures. For one of the two target structures, Li again found that implicit recasts benefited the high-, but not the low-, proficiency learners. In addition, for more advanced learners, the effects of implicit feedback were more durable at the delayed post-test than the effects of metalinguistic feedback. However, for the other structure investigated, recasts were effective for both the high- and low-proficiency learners, suggesting that target feature may also play an important mediating role in the relationship between learners' proficiency levels and the efficacy of corrective feedback.

Overall, then, research seems to suggest that corrective feedback might have varying effects on learners of different proficiency levels. Findings from Gass et al. (2003) suggest that more advanced learners can leverage their attentional resources to better perceive corrective feedback than their lower-proficiency counterparts. It is also possible that more advanced learners are better able to store and retrieve feedback, due to their prior knowledge and linguistic experience (e.g. Atanassova, 2012; Philp, 1999, 2003). For example, Philp (1999) found that more advanced learners were more accurate in recalling implicit corrective feedback (recasts) after engaging in task-based interaction than were lower-proficiency learners. Additionally, Philp (2003) found that more advanced learners were more likely to accurately recall recasts. Similarly, the results of Atanassova (2012) demonstrated advanced learners were significantly more likely to report awareness of corrective feedback, as well as awareness of the target form. In addition, advanced learners' awareness of feedback was not negatively impacted by the type of feedback or target feature, a result that differs from previous research (e.g. Li, Ellis, & Shu, 2016). For example, studies have suggested that the recognition of corrective feedback may vary according to the target of the feedback. Carpenter, Jeon, MacGregor, and Mackey (2006), for instance, found that morphosyntactic recasts were less accurately recognized by advanced English as a second language (ESL) learners than phonological or lexical recasts. Further evidence comes from Li et al. (2016), who manipulated the time between error and feedback to see if immediate or delayed feedback affected uptake for learners at different proficiency levels. When 120 English as a foreign language (EFL) Chinese students, divided into high- and low-proficiency groups according to mid-term exam scores, performed two dictogloss tasks, the researchers found that immediate feedback was most effective for low-proficiency learners, while high-proficiency learners benefited
equally from both immediate and delayed feedback. However, the researchers attributed the difference to the cognitive demands of the task for learners of differing levels rather than to the timing of the corrective feedback. Evidence such as this suggests effects for proficiency in the uptake or noticing of corrective feedback moves during interaction.

Some studies have also uncovered different preferences for feedback that vary according to learner proficiency level. A qualitative study by Lee (2013) found that advanced ESL learners (as judged by their enrollment in a US doctoral program and TSE scores) preferred immediate explicit corrections during classroom interactions. According to questionnaire and interview data, the learners stated they did not feel frustrated by this form of corrective feedback and preferred having all their errors immediately corrected. This was in contrast to the beliefs of their instructors, who did not feel it was appropriate to correct all errors. Furthermore, an analysis of classroom recordings indicated that instructors primarily provided corrective feedback in the form of recasts, highlighting the mismatch between learners’ and instructors’ perceptions. These results are similar to those of Kaivanpanah, Alavi, and Sepehrinia (2015), who found that advanced learners preferred feedback that elicited self-correction, while teachers remained concerned about the negative impact of providing explicit and elicitative forms of feedback. Taken together, these studies point to a need for further research into the complex interplay between corrective feedback, uptake, and learners’ preferences for feedback.

**Peer interaction and advanced proficiency learners**

A large body of research has demonstrated the efficacy of interaction for facilitating L2 development (e.g. Keck et al., 2006; Mackey & Goo, 2007; Ziegler, 2016a), with findings showing that learners who work on a task together tend to perform better than they would working alone (Storch, 1999). As many communicative approaches to language teaching, like TBLT, involve collaborative group work, one of the challenges an instructor might face is how to group learners for tasks in a productive way, especially in classrooms with learners of mixed ability and proficiency levels. In order to understand the role of pair or group dynamics and how this might impact L2 performance and development, researchers have examined how peers can provide opportunities for input, output, negotiation, and corrective feedback, and how these interactional opportunities might vary according to proficiency or native speaker status (see Philp, Adams, & Iwashita, 2013; Pica, Lincoln-Porter, Paninos, & Linnell, 1996). García Mayo & Pica (2000), for example, found that advanced learners were able to provide as much input, feedback, and output to their peer interlocutors as native-speaker interlocutors. Moreover, the advanced learners provided more grammatically complex and accurate feedback to their partner than native speakers did, indicating that the quality of input varied across group and proficiency.

A number of studies have also examined the role of proficiency in learners’ production of Language Related Episodes (LREs), defined as the times during an
interaction in which learners discuss aspects of the language needed to fulfill some task or activity (Swain, 1998; Swain & Lapkin, 1995, 1998). For example, in one of the first studies in this line of research, Swain and Lapkin (1998) found that two advanced learners interacting during a task produced a wide range of LREs. These promising results inspired other researchers to investigate the effects of proficiency level on LRE rates (Leeser, 2004), with findings demonstrating that as proficiency level increased (according to course level), so did the number of LREs that occurred in learner-learner task-based interactions (Williams, 1999, 2001). In addition, as learners’ proficiency level increased, so did the frequency with which they were correctly able to resolve their LREs. Leeser (2004) also found that the number of grammatically focused LREs produced during learner-learner interaction increased with proficiency level. In this study, learners were grouped in the proficiency-based dyads high-high, high-low, and low-low. Groups with two high-proficiency learners produced more LREs focused on grammatical accuracy than on lexical items, whereas groups with two low-proficiency learners focused more on lexical items. The findings of this study suggest that proficiency groupings can impact the quantity of negotiation, as well as the opportunities for learners to focus on form, that occurs during a task. Furthermore, learners’ proficiency affected whether learners focused more on lexical or grammatical features during the interaction.

More recently, Kim and McDonough (2008) examined the impact of interlocutor proficiency on the production and resolution of LREs by pairing learners with intermediate and advanced interlocutors. In addition, this study also examined how pair dynamics might be influenced by interlocutors of different proficiency levels. Results indicated that learners produced and resolved significantly more lexical LREs when collaborating with advanced interlocutors, adding support to previous studies (e.g. Leeser, 2004; Williams, 2001). However, the occurrence of grammatical LREs did not exhibit the same pattern, with findings demonstrating no significant differences across interlocutor proficiency level. Furthermore, learners perceived interaction with advanced interlocutors to be especially beneficial, as they were able to receive answers to grammatical and lexical questions that might arise during the interaction, potentially supporting noticing and the development of metalinguistic knowledge. Similar results were obtained by Choi and Iwashita (2016), with learners producing and resolving more LREs overall when they interacted with an interlocutor of higher proficiency. Lexical LREs were produced more frequently than grammatical LREs, although in contrast to Kim and McDonough (2008), the occurrence of grammatical LREs also increased when learners interacted with a higher-proficiency interlocutor, providing support for Leeser’s (2004) earlier findings. Overall, the greater occurrence and resolution of LREs during interactions involving advanced interlocutors suggests that advanced interlocutors may be more likely to discuss language forms, as well as seek out solutions during interactive tasks, than lower-proficiency learners, thereby potentially increasing their opportunities for L2 development.

In addition to affecting the production and resolution of LREs, proficiency may also influence or interact with pair dynamics during learner interactions. For
instance, Yule and Macdonald (1990) found a difference in the amount of negotiation for meaning that occurred when high- versus low-proficiency learners were allowed to play the more dominant role in a group. During an interactive map task, the learner who was assigned the role of direction provider to another student who looked at a different map was provided with the more dominant role in the interaction. Results showed that when the high-proficiency learner was assigned the role of providing directions, little negotiation for meaning or LREs occurred, but when the lower-proficiency learner gave directions, much more negotiation was necessary to successfully complete the task, potentially leading to increased opportunities for noticing and subsequent L2 development.

However, other studies have found more complex effects from grouping students by proficiency level (Iwashita, 2001; Watanabe & Swain, 2007). In a study by Iwashita (2001), learners of Japanese were placed in the following dyads: low-low, high-high, and low-high, and performed communicative tasks. Results demonstrated no significant differences in the use of different types of corrective feedback, opportunities for modified output, or production of modified output in the learner pairs of various proficiencies. Using a repeated measures design to pair individual learners with interlocutors of higher and lower proficiencies, Watanabe and Swain (2007) examined the role of proficiency in different stages of interaction, including pair writing, noticing, reformulation, and text reconstruction. Findings indicated that when paired with lower-level interlocutors, learners produced more LREs during the noticing stage and had higher subsequent scores on their text reconstructions. Examinations of the patterns of interaction revealed that the pattern may play a greater role than proficiency, with learners engaged in collaborative interaction styles demonstrating higher scores and more LREs than learners with different, less collaborative interaction styles.

Similar results were obtained by Aldosari (2008), who found that pair dynamics played a more important role than proficiency or task type in terms of occurrence and resolution of LREs, with more collaborative styles occurring in matched, rather than mixed-proficiency, pairs. Similar to the findings of Kim and McDonough (2008), pair dynamics appear to be an important factor in the occurrence and resolution of LREs, with results suggesting that intermediate learners would benefit regardless of interlocutor proficiency if a collaborative dynamic occurs. In addition, in order to mitigate the potential discomfort that lower-proficiency learners may feel when working with more advanced learners (Kowal & Swain, 1997), advanced learners should be encouraged to maintain an expert rather than dominant role (Kim & McDonough, 2008). More recently, Dao and McDonough (2017) examined the effects of task role in mixed-proficiency interaction, finding that when lower-level learners held task information, they engaged in more LREs and had higher mutuality than when other task roles were employed, suggesting that task role may mediate the effects of proficiency on the occurrence and resolution of LREs in pair interaction. Overall, although these results indicate that proficiency may impact the quality and outcomes of interaction, they demonstrate that the presence of differing proficiency levels within groups may be mitigated or influenced by other factors, such as pair dynamics, painting a complex picture of how language
proficiency interacts with other individual factors to allow learners to benefit from corrective feedback and peer interaction.

**Technology, interaction, and advanced learners**

A growing body of research suggests that the positive benefits associated with interaction can and do occur in synchronous computer-mediated communication (SCMC), which includes real-time interaction, as in text-, video-, or multimodal chat, with results providing encouraging evidence for the use of technology to support and facilitate L2 development (see Ziegler, 2016a, 2016b, for a meta-analysis and review). Scholars have also suggested that the unique opportunities offered by technology-mediated environments might provide advantages over traditional language-learning settings, with studies demonstrating improved saliency (e.g. Smith, 2004; Ziegler, 2017), as well as increased opportunities for noticing (e.g. Kelm, 1992; Payne & Whitney, 2002; Pelletieri, 2000; Smith, 2003; Smith & Gorsuch, 2004; Toyoda & Harrison, 2002) and focus on form (e.g. Salaberry, 2000; Yilmaz & Yuksel, 2011). Although research has yielded promising results regarding the efficacy of interaction in SCMC, results may vary across modality, setting, and learner and interlocutor characteristics, including age, L1, and proficiency. For example, in one of the earlier studies examining SCMC, Toyoda and Harrison (2002) examined the text-chat discourse of five students enrolled in an advanced Japanese course. Their results indicated that negotiation patterns and interactional features found in face-to-face (FTF) interaction occur in SCMC, suggesting that text-chat provided developmental opportunities for advanced learners, as well as lower-proficiency learners (e.g. Pelletieri, 2000; Salaberry, 2000). Findings also demonstrated differences in interactional patterns between SCMC and FTF modes, such as variation in turn-taking strategies and the existence of different discourse features and the potential for added opportunities for learners to attend more closely to the form and content of the input, while still maintaining the real-time feel of conversation (Toyoda & Harrison, 2002).

More recently, van der Zwaard and Bannink (2014) examined the influence of type of communication mode (text-chat compared to video) on patterns of negotiation between advanced learners and native speakers. Results indicated that mode of communication played an important role, with text-chat negotiation being more to the point and resulting in the resolution of the communication breakdown. Advanced learners in this context also seemed more willing to ask more questions and persist in negotiation in order to resolve the trouble source in text-chat. These findings are similar to those of previous research suggesting benefits of written text-chat with intermediate and lower-level learners (e.g. Abrams, 2003; Beauvois, 1992; Chun, 1994; Kern, 1995). However, because these studies did not directly compare the interaction of learners of different proficiencies, it is not possible to draw firm conclusions regarding differences in the quantity or quality of interaction across beginning, intermediate, and advanced learners.

As with research examining interaction and proficiency in FTF contexts, studies in SCMC have also sought to deepen our understanding of the complex
relationship between interlocutor proficiency and pair dynamics. For instance, Kitade (2000) examined learners interacting with interlocutors of intermediate and advanced proficiency during text-chat, with learners stating they noticed more of what they had learned from advanced interlocutors during the collaborative activity. Sauro (2009), on the other hand, suggests that the high level of learners’ proficiency, determined to be advanced based on their successful completion of a Swedish state exam, may have resulted in learners’ noticing fewer instances of feedback provided by their interlocutors during text-chat. Sauro suggests that the length of some of the recasts may have been too long to be effective, leading to reduced saliency and noticing.

In addition, learner proficiency might influence not only learners’ frequency of noticing, but also learners’ provision of feedback. For instance, Sotillo (2005) found that there were more opportunities for error correction episodes during text-chat interactions where the interlocutor was an advanced learner compared to where the interlocutor was a native speaker of English. More error correction episodes also led to successful corrections during the interactions with advanced interlocutors, mirroring the results of research examining advanced proficiency learners and LREs in face-to-face environments. Advanced learner interlocutors also provided significantly more feedback than NS interlocutors, directing more attention to the production of form rather than focusing more closely on the message or information that the learner was attempting to convey.

Overall, the research examining advanced learners in SCMC environments is promising, with findings suggesting that advanced learners may continue to benefit from text-chat in particular. Considering that scholars have expressed concern that negotiation and interaction may not be as useful for advanced learners as for intermediate and lower-level learners (Long, 2016), the small body of research yielding positive results during text-chat is encouraging. However, as Ziegler (2013) points out, few studies have examined the effects of interaction in computer-mediated contexts on advanced learners’ development, highlighting the need for further research.

**Pedagogical implications**

As communication-focused language instruction such as TBLT continues to grow in popularity in language classrooms, language teachers and action researchers should continue to investigate the implications of these interactive pedagogies in authentic classroom settings. Most language classrooms today include learners at a variety of proficiency levels. Language instructors have the challenge of addressing the needs of each individual student while simultaneously meeting those of the class as a whole. For many teachers, this means addressing whether advanced students should be grouped together or if it might be more developmentally beneficial to mix students of different proficiency levels. Teachers must also consider how they might be best able to challenge advanced students while still meeting the needs of less-advanced students. The following section discusses pedagogical
implications of the interaction approach and offers recommendations in addressing these issues for language teachers or other stakeholders who work with advanced students.

The research reviewed above suggests that, in terms of feedback, explicit corrective feedback may work well for both low-proficiency and high-proficiency learners. Therefore, when classrooms consist of mixed-proficiency learners, explicit correction may stand to benefit the most learners at once. However, when learners are grouped homogeneously by proficiency, advanced learners might benefit more from the flow of meaning-making allowed by implicit correction. Language instructors might also consider varying their feedback style, so that learners of different levels can each benefit in turn from the feedback style for which they are developmentally ready. In addition, the literature on peer-to-peer interaction suggests that more advanced learners working together will engage in similar amounts of negotiation for meaning and LREs as they would if a native speaker was present. Therefore, when negotiation for meaning is the goal of a task, instructors might consider grouping students homogeneously by proficiency level and focusing more energy on the lower-proficiency groups. Alternatively, when groups are mixed, instructors might consider assigning the lower-proficiency learners the more “dominant” roles in a task, as in Yule and Macdonald (1990) and Dao and McDonough (2017), to ensure that the more advanced speakers do not dominate the activity. Additionally, instructors should encourage advanced learners to maintain an expert rather than dominant role (Kim & McDonough, 2008) to mitigate proficiency imbalances when working in heterogeneous groupings. However, as studies such as Iwashita (2001) point out, the proficiency levels of the members of a group may not matter for pushing learners to modify their output. Therefore, instructors might consider varying their grouping depending on the needs of the learners. Importantly, language instructors should be aware of the variety of factors, such as pair dynamics and other individual differences, that have been shown to influence the success of task-based and peer-to-peer interactions, and group students accordingly.

Research examining interaction in computer- or technology-mediated environments suggests that advanced learners may benefit from interacting via text-chat, especially when they interact with an advanced learner rather than a native speaker (Sotillo, 2005). In this context, advanced learners could work on resolving communication breakdowns and asking questions, areas that have been shown to be supported and enhanced by the text-chat environment. The judicious integration of text-chats and other technologies into the second language classroom is one exciting area of future investigation for researchers and instructors of advanced L2 learners.

Conclusions and future directions

The research described above has pointed toward a complex and developing understanding of the effects of interaction on learners at advanced levels of proficiency. In the domain of corrective feedback, research shows varying effects for learners at
different proficiency levels, with particular effects for proficiency in the uptake or noticing of corrective feedback moves during interaction. Based on these findings, language instructors may wish to vary their feedback style so that learners of different levels can benefit. In addition, findings from research on learner-learner interaction demonstrate that the proficiency of group members can impact the quantity and quality of negotiation and focus on form that occurs during a task, as well as what kinds of features learners focus on the most during task-based interactions. Considering these findings, instructors might aim to vary their grouping depending on the needs of the learners. Finally, research examining advanced learners in SCMC environments has suggested that advanced learners tend to benefit from text-chat by resolving more communication breakdowns and negotiating for meaning longer than learners of lower proficiency levels. Advanced learners also tended to provide more corrective feedback in SCMC when working in groups than when working with native speakers, suggesting that peer interaction may be equally or more beneficial than interaction with a native-speaker interlocutor.

Overall, the interaction approach to SLA has benefited from a robust body of research examining learners at a variety of proficiency levels. However, there are many promising directions for future investigations. For example, despite the large number of studies described in this chapter that have linked interaction with L2 development, not all researchers subscribe to the idea that interaction is the primary means by which language proficiency develops. The hypothesis was criticized at times for not addressing all aspects of the learning process, and some have taken issue with the validity of the hypothesis (for a review of these criticisms and responses see Ellis, 2003; Mackey & Gass, 2015). However, these criticisms were generally targeted at the earlier, stronger version of Long’s hypothesis rather than the updated version that focuses on particular aspects of the SLA process (Long, 1996). In addition, many studies in the interactionist tradition, and in the applied linguistics field in general, have focused on college-age students in university settings, with a heavy focus on English as the L1 or L2, highlighting the need for more research on younger or older learners as well as with target languages other than English. Furthermore, the majority of studies have examined learners of intermediate proficiency, with much less known about the effects of interaction on highly proficient learners. This has resulted in some researchers calling for methodological reforms to improve the generalizability of findings (Plonsky, 2013, 2014). More longitudinal and replication studies have also been called for to better understand the effects of interaction over time (Mackey, Abbuhl, & Gass, 2012). Additionally, researchers in the field have called for more integration of sociocognitive aspects of interaction, including the effects of social relationships and context on interaction (Lantolf, 2012), in order to deepen our understanding of the relationship between interaction and L2 performance and development.

One new and growing area of interaction research is studies that examine the interplay of individual differences, such as aptitude, with instructional
treatments. These “aptitude treatment interaction” (ATI) studies (see Li, 2015; Révész, 2011; Yilmaz, 2013, for examples) often combine quantitative and qualitative methods to match (or mis-match) learners to conditions thought to be most compatible with their aptitude (see Vatz, Tare, Jackson, & Doughty, 2013, for a thorough overview). While this is still an under-researched area in the field, future studies of interaction should consider drawing on the same rigorous methodologies employed in ATI studies in order to further investigate the role of proficiency in second language learning. By matching learners of various proficiency levels to conditions (for example, a certain type of corrective feedback, or type of task), researchers can enhance our understanding of complex relationships at work when varying teaching methods are employed with learners at advanced proficiency levels. Furthermore, we suggest that future research employ mixed methodology designs, as well as examine findings both quantitatively and qualitatively, thus providing a more holistic perspective on the process and product of L2 learning outcomes. For example, quantitative methods could employ empirical tests of learners’ aptitudes alongside proficiency, as well as utilize advanced technology, such as eye-tracking, to more precisely measure learners’ noticing of new forms at various proficiency levels. Qualitative data can be gathered via stimulated recall interviews (see Gass & Mackey, 2016) to further investigate the various processes learners engage in at advanced proficiency levels. Regardless of design, there is a clear need for more SLA researchers to directly investigate proficiency and to specifically design studies with advanced proficiency learners in mind. This line of inquiry has important implications for language teachers, as well as for students with higher language proficiency. Research that carefully compares and contrasts instructional treatments with learners at various proficiency levels may help educators make more informed decisions about the kinds of tasks and feedback they provide in their classrooms. This in turn will help learners discover how to best leverage their advanced proficiency to continue to improve their language skills.

This chapter has sought to synthesize the current research and understandings regarding language learners with advanced proficiency from an interactionist approach to second language acquisition. By examining studies of corrective feedback, peer interaction, task-based language teaching (TBLT), and interaction in computer-mediated contexts, we have aimed to provide a comprehensive overview of what is currently known about the relationship between advanced proficiency and interaction-driven language learning. While there is still much work to be done to understand this complex relationship, it is our hope that researchers and practitioners can utilize this chapter to explore new questions in this domain and apply them to authentic language-learning contexts, thus helping to drive the field forward in terms of both empirical and pedagogical progress.
REFERENCES


Interaction-Driven L2 Learning: Advanced Learners


Li, S. (2009). The differential effects of implicit and explicit feedback on second language (L2) learners at different proficiency levels, 19(1), 53–79.


7 Sociocultural Theory: Mediating Learners toward Advanced Proficiency

MATTHEW E. POEHNER
The Pennsylvania State University

Introduction

As formulated in the writings of Russian psychologist L. S. Vygotsky during the 1920s and 1930s, Sociocultural Theory (henceforth, SCT) offers a coherent perspective on human consciousness, including its fundamental roots in biology, its transformation through culture, and the role of education as a special, culturally organized form of activity that brings about new ways of understanding the world and acting to change it. Owing to the political context following the Russian Revolution, Vygotsky’s writings were banned for decades and only became known to Western researchers with the appearance of Mind in Society (1978), a heavily edited and translated synthesis of several of Vygotsky’s key ideas from lecture notes and other papers. Even in this work it is possible to grasp the scope of Vygotsky’s enterprise, which he explicated in greater detail elsewhere as outlining a scientific psychology rooted in the philosophy of dialectical materialism (Vygotsky, 1997). Given that Vygotsky operated at the level of philosopher of a new science as well as empirical researcher, that he engaged in teacher education, and that he participated in debates of psycho-educational measurement, it is perhaps not all that surprising that the discovery of his work has generated scholarship in so many fields. At the same time, it is precisely because Vygotsky was not focused on a very specific area within psychology that researchers wishing to extend his ideas to new domains have confronted both the opportunity and the challenge of determining the theory’s implications for problems and questions that Vygotsky himself did not fully explore.
In second language acquisition (SLA), the first published SCT paper is Frawley and Lantolf’s (1985) analysis of L2 discourse from a Vygotskian perspective, which was informed by the 1978 text as the major source of Vygotsky’s ideas available at that time. Since then, the field has seen countless papers, books, and conference presentations elaborating SCT and its relevance to examining fundamental questions in SLA. As more and more of Vygotsky’s writings have been made available in translation, L2 researchers gradually have come to a greater appreciation of his philosophical positions and their consequences for theory, research, and educational practice. Two distinct trends are discernible in the L2 SCT research literature: scholarship that employs concepts from the theory to explicate processes of L2 development and efforts to design educational environments that promote development. While both lines of research offer contributions to SLA scholars interested in advanced proficiency, the focus of this chapter will be the latter. This choice is motivated by the preponderance of innovative studies that have adopted this orientation since the work of Negueruela (2003) as well as by a conviction that this position resonates especially well with Vygotsky’s own views concerning the dialectical relation between theory and practice, in which the theory provides an appropriate orienting basis for practice but must be responsive to and indeed take account of the realities of practice. Lantolf and Poehner (2014) refer to this position as *praxis*.

In what follows, I present key principles in SCT that in themselves offer a prism through which one might productively study the development of advanced language proficiency. However, my primary purpose is to use this discussion as background for presenting pedagogical initiatives that are based in these principles and that seek not only to explain the development of advanced L2 proficiency but to actively promote it. These include *Systemic-Theoretical Instruction* (STI), rooted in Vygotsky’s (1987) proposal of abstract, scientific concepts as the proper content of schooling, and *Dynamic Assessment* (DA) and *Mediated Development* (MD) as interactional frameworks aligned with Vygotsky’s discovery of the *Zone of Proximal Development* (ZPD) and his analysis of its importance for education. Taken together, these initiatives function to move learners toward higher levels of proficiency, which as will be made clear, from an SCT perspective entails the appropriation of semiotic resources offered by the target language and intentional control over them during meaning-making activity.

**Principles of Sociocultural Theory**

**Internalization and mediation through signs**

It is important to recognize at the outset that unlike certain other theories employed by scholars in SLA, SCT is not a theory of language. Vygotsky himself was not a linguist interested in formulating a theory of language or communication but was devoted to understanding human psychology and its development. As Vygotsky’s enterprise was to articulate coherent principles for a scientific
psychology, he argued explicitly and passionately that all of the various subfields and specializations within psychology needed to be accounted for by a general psychology (Vygotsky, 1997). The general principles of psychology, he explained, needed to explicate phenomena as wide-ranging as basic general functions as well as abnormalities and difficulties and high-level reasoning and reflection. Thus, while Vygotsky’s available writings do not specifically address processes of L2 development, the fact that his theory is one of general psychology means that it must account for the development of all psychological abilities, including language. As will be explained, language enjoys a central role in Vygotsky’s analysis of human psychology.

SCT introduces a fundamental distinction between what Vygotsky (1987) termed higher and lower psychological functions. Higher psychological functions represent the transformation of lower, or natural, functions that humans share with other animals, such as memory, attention, and perception. These natural psychological functions are altered through the introduction of culturally available resources that enable individuals to control and direct their psychological activity. By analogy with the manner in which physical tools are employed to mediate activity in the material world, semiotic resources, notably language, serve as a symbolic tool to mediate psychological activity. As one concrete example of this process, Vygotsky (1987) describes the cultural behavior of tying string around one’s finger to remind oneself of something that needs to be done. According to Vygotsky, the bit of string has no inherent connection with what needs to be remembered; nonetheless, the string is imbued with meaning and comes to function as a sign. That is, the string/sign serves as a secondary form of environmental stimuli, one that is intentionally introduced, and that triggers and directs a natural process of remembering. Similarly, this process of mediating one’s psychological activity through a secondary stimulus can be carried out completely internally. Rather than relying on a bit of string to guide an act of remembering, people engage in verbal thinking. Verbal thinking, enabled through language, qualitatively changes the basic psychological function of remembering, allowing individuals to actively call certain phenomena to mind and to (re-)construct memories through narratives. Remembering becomes a culturally mediated psychological activity.

The process through which semiotic resources are appropriated was referred to by Vygotsky as internalization. Vygotsky and Luria (1994, p.155, italics in original) describe internalization as follows:

> what was an outward sign operation, i.e. a certain cultural method of self-control from without, is now transformed into a new intra-psychological layer and gives birth to a new psychological system, incomparably superior in content, and cultural-psychological in genesis.

Internalization is thus a “reconstruction of psychological activity” (1994, p.155) on the internal plane of what has previously been undertaken in cooperation with others. Toomela (2014) cautions, however, that internalization must not be taken to
mean that something literally enters an individual’s mind from the surrounding environment. He points to advances in neuropsychological research since Vygotsky’s time that allow for greater clarification of how the natural, lower forms of consciousness are transformed into cultural, higher consciousness and the role of meaning in this process. Toomela reports that it is only through sensory attributes that individuals are able to render psychological elements of their physical environment. As he puts it, “the environment comes into relationship with the individual mind through, and only through, sensory organs that transform the environmental events into neural signals” (Toomela, 2014, p. 114, emphasis in original). By way of example, he points to color, contour, and orientation as environmental elements that are perceived as sensory attributes through the visual system and that are later ‘synthesized’ into our experiences of objects and events. With regard to the social and cultural environment, Toomela notes that internalization is concerned not with sensory attributes of the environment but with “the emergence of semiotically mediated psychological structures” (p. 115). In a word, the development of higher forms of consciousness is driven by the internalization of signs. As Toomela puts it, “No external artifact can become an element of a psychological structure unless that artifact can be fully created both mentally and behaviorally. Only signs have such properties” (p. 116, emphasis in original).

Vygotsky (1987) explained the role of signs through analogy with physical artifacts that cultures create. The use of physical tools transforms our physical activity in the world, as can be seen in the difference between digging a hole with one’s hands, with a simple implement such as a spade or shovel, or by using an excavator. Signs function as symbolic tools that transform our psychological activity. Such symbolic tools may include language and counting systems as well as graphs, charts, models, and conceptual knowledge. Development from an SCT perspective is the process of internalizing symbolic resources. The process may begin with their use in a form that is external to us, as we come to understand their potential to mediate our actions, and it continues as the meanings that they carry are taken in by us, becoming part of our intra-mental activity. As will be explained, mediation through dialogic interaction with others and mediation involving specially prepared representations of concepts have both been identified by Vygotskian scholars as important means for development through schooling. This focus has also become increasingly prevalent among L2 SCT researchers. In the L2 field, where the aim of education is to promote learner language proficiency, learner internalization of high-quality conceptual knowledge is paramount. Relying on such understandings, learners are able to regulate their interpretation and use of language, making conscious decisions concerning which linguistic resources they draw upon to convey meanings. Specific examples of this work are discussed later in this chapter.

**Developing higher forms of consciousness through schooling**

Vygotsky regarded schooling as a specially created environment for the internalization of abstract conceptual knowledge. In entering debates over the relationship between education and psychological development, Vygotsky explicitly rejected
two views dominant in his day, one that held that the primary consideration for instruction is that it must await learner readiness (i.e. that the requisite development occurs prior to instruction), and the other that saw both instruction and development as resting upon the formation of associations and to therefore be one and the same (Vygotsky, 1987). Vygotsky advances his own position of obuchenie (teaching/learning) by first pointing to a precedent in a framework for school curricula during his time that he refers to as the theory of formal discipline. Vygotsky (1987, p. 198) characterizes this view as proposing that the study of academic disciplines “provides something more than the knowledge and skills that constitute the subject itself.” He offers as examples that this view held that studying classical languages promoted the development of capabilities needed for thinking in the human sciences more generally and that the study of mathematics offered similar advantages in natural sciences and technical fields. Vygotsky further notes that while he finds the idea progressive, the pedagogy through which the curriculum was realized tended to be traditional and therefore was not likely to yield the intended outcomes. He goes on to critique the work of Thorndike, whose research into the theory of formal disciplines not only failed to show any abilities that might be extended beyond a given discipline but also led to the extreme opposite view, namely that instruction of any kind cannot be expected to promote more abilities. In responding to Thorndike, Vygotsky makes two important distinctions. The first of these is to distinguish forms of development that do not in fact appear to impact cognitive functioning outside of a very specific domain. As examples, Vygotsky mentions learning to swim, ride a bicycle, or play golf. Curiously, each of these is a form of physical activity and so while one might speculate about potential general benefits to, say, gross motor coordination or balance, it does indeed seem unlikely that any would have an effect on cognition. Vygotsky’s second point is to disambiguate educational activity that is tightly focused upon developing a very narrow skill set, which he remarks pertains to some adult vocational education, and instruction that may occur during childhood and that “incorporates whole complexes of mental functions” (p. 200). It is this latter form of educational activity that Vygotsky regards as the proper focus of optimally guiding learner development of high forms of consciousness and that he explores in his discussions of obuchenie.

Given Vygotsky’s untimely death and the resultant incomplete state of many of his proposals, researchers have pursued a range of models for realizing obuchenie in practice (e.g. Bodrova & Leong, 2007; Davydov, 2004; Gal’perin, 1992; Karpov, 2014; Kozulin, 1998). Again, given that SCT needs to account for the development of all abilities, this body of research has pursued development in contexts that include early childhood education, basic literacy and numeracy during primary school years, the study of specific academic disciplines, and special education. While discussion of each of these models is well beyond the scope of the present chapter, it is important to note that they share a commitment to helping learners internalize conceptual knowledge that provides an orienting basis for action. That is, the goal of such work is not merely for learners to memorize definitions of abstract concepts but rather to acquire knowledge that they can draw upon (i.e. employ as mediation) for thinking and acting. Karpov (2003) expresses the significance of
Sociocultural Theory: Mediating Learners toward Advanced Proficiency

regarding knowledge and action as integrated in his comparison of both traditional and progressive pedagogies with one that is rooted in SCT. He explains that traditional approaches emphasize repetition of discrete skills and rote memorization of factual information while so-called progressive, discovery approaches to instruction require learners to function as scientists as they attempt to ascertain on their own the important principles, concepts, and relations within an area of study. As Karpov observes, considerable research has documented that under both systems many learners fail to reach the desired developmental goals (see also Egan, 2002). Karpov attributes this failure to the lack of conceptual understanding that enables learners to easily connect theoretical knowledge with procedural knowledge that together provide the basis for contemplating new and complex problems, formulating plans, and evaluating outcomes of their actions. Pedagogies that rely heavily on memorization and discrete skills often lead learners to success only when they encounter problems of a very specific type, while discovery approaches to learning leave many students with only partial understandings of important principles. He concludes by underscoring that while Vygotsky (1987) wrote of schooling as a unique environment for learners to come into contact with abstract conceptual knowledge in various disciplines, these forms of knowledge will yield psychological transformations and will “play such a mediational role only if they are supported by students’ mastery of relevant procedures” (Karpov, 2003, p. 68, emphasis added).

Karpov (2014) notes that in Russia during the decades following Vygotsky’s death, his colleagues and students carried out extensive work collaborating with teachers at the primary and secondary level to design and implement concept-based curricula across academic content areas (see also Gal’perin, 1992). Outside Russia, a leading proponent of such work is Kozulin, who has pursued concept-based research in the teaching of mathematics, literature, and science (e.g. Kinard & Kozulin, 2008; Kozulin, 1998). An important contribution of Kozulin’s research has been to bring Vygotsky’s writings on abstract conceptual knowledge into contact with the framework of the Mediated Learning Experience elaborated by Reuven Feuerstein and colleagues in Israel in their work to remediate basic cognitive functioning among learners with special needs (Feuerstein, Feuerstein, & Falik, 2010; Feuerstein, Rand, & Rynders, 1988). According to Miller (2011), Feuerstein’s theories developed independent of Vygotsky’s writings but exhibit remarkable parallels with regard to both how they conceive of human psychology and their approach to organizing education to promote the development of new ways of thinking. In Kozulin’s work, the commitment to specially designed materials that serve as mediating tools for learners and that carry meanings that learners eventually internalize is readily apparent. Moreover, Kozulin (1998) argues that the emphasis on learner dialogic interaction with an expert, or mediator, in Mediated Learning Experience fully aligns with Vygotsky’s understanding of the ZPD.

Briefly, Vygotsky (1978) argued that human psychology at any point comprises abilities that have already fully developed as well as those that have begun the developmental process but have not yet fully emerged. Independent performance of tasks, as occurs in most assessment situations, may reveal abilities that have
already developed, as it is in such contexts that individuals rely on what has already been internalized to regulate their functioning. Abilities that are still emerging, in contrast, may be understood through an individual’s participation in activities undertaken cooperatively with others. Specifically, an individual’s contributions to interpsychological activity and responsiveness to support that is offered reveal how the emerging abilities are developing. Attuning to these abilities-in-the-process-of-developing was crucial, in Vygotsky’s view, for obuchenie to succeed. As he explained, the ZPD provides essential information for an assessment of learner development by bringing to light the full range of their abilities while at the same time locating the proper focus of instruction aimed at promoting development (Vygotsky, 1998). According to Kozulin (1998), the Mediated Learning Experience realizes both of these goals. Certain sessions situate mediator and learner in interaction for the purpose of diagnosing what the learner is able to do independently and how responsive s/he is when tasks become too difficult and the mediator must intervene to provide prompts, feedback, models, and guidance. These sessions are described as DA because while the interaction is unmistakably instructional in character, the primary objective is to assess learner development. These results of the DA then inform mediator efforts to introduce new concepts and instructional materials, model how these can serve as tools, and guide learner internalization of their meanings. In this way, mediation through symbolic tools and through dialogic interaction come together to alternately bring into focus the diagnosis and promotion of learner abilities in an ongoing form of developmental education. As explained in the following section, these principles have been taken up and further elaborated by L2 researchers, particularly those concerned with the use of SCT as a basis for guiding L2 educational practice.

**L2 education to promote advanced proficiency**

Lantolf and Poehner (2014) argue that in his theoretical writings as well as through his work in special education and teacher education, Vygotsky himself was wholly committed to a vision of praxis in which theory guides practical activity but is also evaluated and refined according to the demands of practice. For these authors, SCT reveals its full relevance to L2 education and is further developed only by bringing its proposals into contact with the realities of assessing, teaching, and learning. Among the initiatives currently pursued by L2 researchers are STI, DA, and most recently, MD. Together, these frameworks offer a theoretically unified approach to supporting learner internalization of knowledge of the L2, in particular abstract conceptual knowledge, that they can draw upon to make reasoned decisions concerning how to express themselves. In short, these initiatives offer an approach to leveraging Vygotsky’s general vision of development through education in order to promote learner L2 proficiency.

The studies that are discussed are intended to illustrate how L2 researchers to date have employed SCT concepts and principles to support learner development in a variety of contexts. As Vygotsky’s theory is one that is concerned with all
forms of development and is not limited to any particular domain, these L2 studies should not be viewed as definitive of the populations, language features, or proficiency levels to which SCT is relevant. While details of the approaches need of course to be adapted to particular contexts, the commitment to conceptual knowledge and dialogic interaction as mediating learner development pertains to all learners of all languages and at any level of ability.

The content of L2 instruction as mediation toward proficiency: STI and MD

STI reorganizes L2 curricula around abstract linguistic concepts that may serve as symbolic tools with which learners can regulate their use of the target language (Negueruela, 2003). The central argument behind STI is that the central role assigned to morphosyntactic rules in many traditional approaches to L2 instruction does not provide learners an adequate basis for interpreting and conveying nuanced meanings through the L2 that are appropriate to communicative contexts. As Lantolf and Poehner (2014) explain, analyses of language that bring meaning to the fore may offer especially rich understandings of *znachenie* (conventionalized meaning) and for this reason STI researchers have frequently turned to work in cognitive linguistics for explanations of language concepts. Lantolf and Poehner continue that STI provides a pedagogical approach to help learners translate this conceptual knowledge into *smysl* (personal meaning) for the purpose of gaining greater control over the L2. Negueruela’s (2003) study initiated a line of research concerned with relating events and actions to one another in time and the resources available for doing so in various languages. For Negueruela, dealing specifically with L2 Spanish, this meant addressing the preterit–imperfect distinction that frequently poses difficulty to learners. According to Negueruela (2003), the preterit and imperfect are typically—and incorrectly—presented to learners by textbooks and instructors as two tenses that function following very specific sets of rules. In that model, learners proceed to memorize the rules and practice manipulating verb morphology in order to master the distinction. Pointing to research that documents misunderstandings of the preterit and imperfect that persist among learners at higher levels of proficiency, Negueruela argues that memorizing the relevant rules does not offer an adequate basis for understanding how together the preterit and imperfect offer possibilities for describing the past in relation to the present. Negueruela (2003) organized an STI program that introduced learners to verbal aspect as an abstract concept. A definition of aspect was provided and, more important, images were presented that depicted how relations among events change when one or the other (or both) is conveyed through the preterit or imperfect. Exemplar sentences in Spanish and English accompanied the images to help transition learners from a search for the ‘correct’ response to instead consider which forms most appropriately matched the meanings they wished to express. In particular, learner attention was directed to the ways in which the preterit and imperfect together serve to foreground or background events and the
consequences this has for meaning. Learners engaged in a series of tasks that involved selection of verbal aspect, and the visual materials and examples were available to them as tools to mediate their thinking. Following the work of Gal’perin (1992), Negueruela (2003) led the learners to also verbalize their understanding of verbal aspect and how it informed their linguistic choices as an additional step toward their internalization of the concept.

Lai (2012) similarly identified verbal aspect as a source of difficulty for learners of L2 Chinese, although given the differences with regard to how tense functions in that language versus in English, her STI project dealt extensively with offering L1 English learners of the language a conceptual understanding of Chinese tense. Following Lakoff and Johnson’s (1999) work in cognitive linguistics on spatial metaphors of time, Lai (2012) identified that a central challenge for L1 English learners of Chinese is that the two languages follow opposing spatial orientations. That is, in English the present time is conceptualized as the spatial location of the observer, with the future and past extending, respectively, to the observer’s right and left, such that a time line would place the future in front of the observer (to his/her right) and the past to the observer’s back (to his/her left). Lai notes that in Chinese, the past is actually communicated through the particle qián (front) and the future through the particle hou (back). Moreover, Chinese employs an additional spatial metaphor for time whereby shàng (up) is indicative of the past and xià (down) denotes the future. As Lai (2012) explains, knowledge of how these particles are rendered in English is insufficient for learners to understand their function in positioning events in time and in fact may appear counterintuitive. For the portion of her STI project that focused on tense in Chinese, Lai created a series of images portraying Chinese spatial metaphors for time. One contained a train with the relevant target language particles positioned at the front and back of the train. It was explained that according to this perspective, the train is moving past the observer, who is positioned roughly in the middle of the image. The front of the train contains qián, because it has already past (it is therefore in the past) while the back of the train displays hou and has not yet past the observer (thus it remains in the future). Another image represented Chinese spatial metaphors of time in a manner similar to the X-Y axes in mathematics, so that learners could see that the past is ‘right’ and ‘up’ while the future is ‘left’ and ‘down.’ For research purposes, Lai (2012) implemented her STI approach with a class of undergraduate university Chinese learners and compared their performance on tests of accuracy employing the past tense with those of learners in two other classes. In addition to quantitative analysis of the test scores, Lai also conducted in-depth qualitative analysis of learner verbalizations of their understanding of the Chinese temporal system and the reasons behind their choice to use particular forms during language tasks. Full discussion of the results of Lai’s research is beyond the scope of this chapter, but overall Lai (2012) concluded that learners who had participated in the STI program generally performed in a more systematic manner in their use of tense, prompting her to argue that a conceptual treatment of the topic is worth introducing at the earliest level of study.
Before turning to a third STI project, it is worth briefly considering Infante’s (2016) recent MD study as it paralleled Negueruela’s (2003) and Lai’s (2012) interest in the concepts of verbal tense and aspect. Infante’s work, however, extended this work in an important way by designing interactional sessions between an expert mediator and the learners in order to create opportunities to dialogically support learner internalization of conceptual knowledge of the L2. Specifically, Infante (2016) drew upon Feuerstein’s work to focus interactions with learners on helping them to recognize how STI materials may serve as psychological tools as they interpret instances of language as well as make reasoned decisions on how to express their own meanings. As in previous STI studies, Infante began by verbally explaining the linguistic concept, which in his case was the verbal tense-aspect system in English. His learners, which included a mix of secondary and university-level ESL students, then received a series of visual representations of tense-aspect and were asked to verbalize their interpretation of the images. Infante, functioning as a mediator, engaged dialogically with the students, prompting them to consider details of the images, drawing their attention to particular features, and posing leading questions. In subsequent sessions, the mediator demonstrated how exemplar sentences in English could be interpreted by referencing the materials as well as the images’ relevance to selecting forms appropriate to expressing particular ideas (e.g., *John was cuddling the baby when the doorbell rang*). After modeling how the materials can serve as tools, learners cooperated with the mediator to work through additional exercises before turning to co-revision of texts they had written in English. As the STI program continued, certain MD interactions also invited learners to perform particular psychological actions such as labeling internal elements of the conceptual images and verbalizing how each element contributed to the overall set of relations represented by the image. Another session asked learners to function in the absence of the materials but to attempt to visualize them and refer to them while revising their writing with the mediator. Some learners even devised their own model of the tense-aspect system. All of these interactions involved the presence of Infante as an interlocutor to prompt learner verbalization of their thinking and to guide and extend it. In this way, Infante’s (2016) research offers an important bridge between previous STI and DA studies, which have emphasized, respectively, symbolic mediation and dialogic mediation.

Finally, Kao (2014) turned to a much broader feature of communication in her STI study of L1 English students developing writing proficiency in L2 Chinese. Acknowledging the extensive literature on contrastive rhetoric, Kao argues against the traditional view attributed to researchers such as Kaplan (1972) that holds that each culture has its own specific style for organizing expository writing (e.g., while English is said to proceed in a linear path beginning with an explicit introduction of the thesis statement and continuing to present greater detail and supporting examples before offering a conclusion, Chinese is believed to follow an Oriental “spiral” pattern in which several ideas are introduced, repeated, and revisited). Building on more recent work in critical contrastive rhetoric, Kao (2014) devised
an STI program focused on the concept of a Chinese ‘theme’ as analogous to the English notion of ‘thesis statement.’ The Chinese theme, referred to as zhuti (‘main theme’) or zhuyi (‘main idea’), functions to delimit the topic of the text as well as announcing how the writer will approach it. However, in contrast with the English thesis statement, Kao (2014, p. 66) explains that

the method to develop a [Chinese] theme is to examine it comprehensively, or from multiple perspectives. It is unnecessary to narrow one’s focus to explore one point in depth. Instead, it is important to broadly discuss the theme or to divide the topic and discuss all of its aspects to ensure that every part has been exposed to the reader.

This conceptual frame was presented to the learners in Kao’s study, all of whom were recruited from a beginning level undergraduate university Chinese program. The concept of Chinese theme was further accompanied by visual representations of two distinct patterns in Chinese expository writing that differ with regard to their placement of the theme. The first of these, referred to as “Open the door and see the mountain” (kai men jian shan) is a more direct approach that, as the metaphor suggests, involves announcing the theme from the outset of the text, although as Kao (2014) notes, this is not always done in a manner as explicit as occurs in English presentations of a thesis statement. Instead, the door is opened to the reader and a path set out but the reader shares in the responsibility of actively forming connections across ideas and examples in order to arrive at a view of the mountain. The second approach, “Drawing the eyes of the dragon” (a form of Qu-cheng-zhuan-he), is much more indirect in that the theme is only reached at the end after the writer has elaborated the topic from several angles. Its name derives from the practice in Chinese art in which a painter would invest considerable time in drawing a dragon but would save until the very end the addition of the eyes that make it come to life.

A verbal description of the concept of Chinese theme was presented to learners in Kao’s (2014) study along with the two metaphorical organization patterns of “Open the door and see the mountain” and “Drawing the eyes of the dragon,” which were also both represented through images. These tools were available to learners as they read authentic Chinese texts and attempted to identify, in each, the theme and how it was developed by the writer. Learners generated their own ‘maps’ of the texts they read and discussed these with a mediator (the researcher), who provided feedback on their accuracy. Later, the learners worked to compose their own expository texts in the L2 and to explain, in English, which of the organizational patterns they had attempted to follow. According to Kao (2014), the overall results of the STI intervention were encouraging, although learners varied with regard to their success in moving from recognizing the patterns and how they differed from English to actually producing appropriate Chinese texts. Nonetheless, integrating such work at a relatively early stage in their study of the language indicates that a more sustained effort could impact the level of writing proficiency they ultimately attain.
Teacher-learner interaction as mediation for L2 proficiency: DA

While STI has emphasized the importance of mediation through language concepts and their representation in specially designed materials, DA has brought to the fore mediation as a dialogical process during interspsychological functioning between L2 teachers and learners. As explained, careful attention is given both to the extent of teacher, or mediator, involvement in prompting and guiding learners and to learner responsiveness during this process. The logic is that the greater the support learners require, the further they are from independent functioning. Determining the degree to which learner development of the abilities in question has already begun is relevant to many assessment decisions, perhaps most especially planning for future instruction. With regard to moving learners toward more advanced L2 proficiency, DA research to date has primarily targeted learner understanding of specific features of language and their capacity to deploy them in communicative contexts. However, as will be explained, studies have also focused on a more global conceptualization of language ability, such as oral proficiency and listening and reading comprehension.

The first in-depth application of DA principles with L2 learners was conducted by Poehner (2005) in a study aimed at control over tense and aspect during oral narration tasks among university-level learners of L2 French. The DA sessions occurred in a clinical setting in which a mediator met individually with students and played brief clips from popular films before asking them to reconstruct the events in the video in a narrative in the target language. This included sequencing the events and making choices concerning how to relate events to one another, which events might be foregrounded, and how dialogue and character reactions could be depicted. The mediator intervened to request that learners verbalize the reasons behind their tense and aspect choices and prompted them to reconsider their choices when the forms they selected did not match their intended meaning. As Poehner (2005) explains, weekly sessions with learners allowed the mediator to track changes in their thinking and their control over tense and aspect in French. Following this initial exploration of DA in the L2 domain, Poehner (2009) implemented a DA program in a primary school with learners of L2 Spanish. Not only did this project build on the earlier study by shifting from one-to-one interactions with learners to use of DA in a classroom setting, it also organized the quality of dialogic mediation in a more structured manner. That is, while the mediator in Poehner’s (2005) project followed an open-ended approach to supporting learners on the grounds that this optimized the possibility of attuning to learner emerging needs and difficulties, the classroom teacher who carried out DA in Poehner’s (2009) study scripted mediation in advance as a set of prompts and leading questions arranged from most implicit to most explicit. For example, when a learner produced an error during a classroom task the teacher’s initial response was to pause and allow the learner the opportunity to
identify and attempt to correct the mistake. If this was unsuccessful, the teacher repeated the learner’s utterance with a questioning intonation in order to prompt learner reflection on the statement. Subsequent moves drew learner attention to the specific parts of the response that were problematic and attempted to guide the learner toward an alternative. The process continued until, if necessary, the teacher revealed the correct forms and explained why they were correct. In this way, the teacher was able to ascertain the precise point at which learners were able to recognize the difficulty and begin to work to overcome it. In addition, she traced learners over time, noting changes to the number of prompts they required as they continued to work on a given language topic.

Levi (2012) extended L2 DA even further by bringing it into the context of large-scale language assessment. Moreover, rather than targeting learner control over specific features of language, her project was concerned with oral proficiency as a more general construct. Working within the context of English language oral proficiency interviews required for secondary school students in Israel, Levi carried out a three-stage form of DA in which a version of the interview was administered to students at two points in time, with an intervention occurring between them. Mediation during the four intervention sessions was twofold: students were provided with a detailed version of the rubric used to assess their oral proficiency and an expert mediator was present to provide additional support. Using the rubric, students were asked to watch recordings of either their own or another student’s initial oral proficiency interview and to evaluate the performance according to the rubric’s scales. In this way, learners began to gain greater awareness of specific features of their own language use and they worked to use the rubric dimensions to monitor their oral performances. Levi’s (2012) project stands out because it maintained a commitment to dialogic mediation, through the presence of the expert mediator during the intervention sessions, and demonstrated that in addition to its applicability to classroom settings, DA is also suitable in large-scale contexts.

A further innovation in the use of DA to monitor and promote learner language proficiency concerns its administration in a computerized format. Here the mediator’s responsibility to provide support and track learner responsiveness is performed by a computer program that embeds programmed feedback into the test. Leontjev (2016) illustrates the potential of computerized DA with L2 learners in his study aimed at diagnosing grammatical competence among Estonian learners of L2 English transitioning from primary to secondary school. Particularly important features of English that learners must demonstrate control over include wh-question formation and affix word derivation. Employing sentence-builder tasks, Leontjev’s computerized exam requires learners to select and modify forms in order to produce grammatical English sentences. When learners answer incorrectly, the computer provides feedback through an implicit prompt and then takes the learner to the next item on the test, which is a parallel item to the one in which the error was produced. If a learner again responds incorrectly, the process is repeated but a more explicit prompt is provided. In this way, the mediation is
similar to that described in Poehner’s (2009) primary school L2 Spanish study except that the prompts are generated automatically as part of the computerized test. An advantage of computerized DA is that it allows for the possibility of multiple test takers simultaneously engaging in a mediational process.

**Future directions**

To date, L2 STI programs as well as initial efforts to elaborate MD have focused on specific features of language, presenting them to learners as systematic concepts that can be referenced while determining meanings one wishes to convey. Much of the L2 DA research has similarly targeted learner control over particular linguistic forms, albeit with an emphasis on how those forms may be purposefully selected to convey meanings. As explained, some DA studies have begun to broaden their focus from component features of L2 proficiency toward more global constructs. For STI research to similarly take a broader view, it will be necessary to greatly expand upon the existing body of studies that each treat specified constructs in order to arrive at a Vygotskian concept-based curriculum that ranges across all levels of study, from novice to the most advanced. While such an enterprise is extremely ambitious and would likely require coordinated efforts among a team of researchers collaborating with teachers at every level of language instruction, it would not be without precedent. Indeed, the efforts by Gal’perin (1989) and colleagues over many years to implement Vygotskian pedagogical principles in the Moscow schools led to innovations in various academic domains, including language, and extended through primary and secondary levels. In the case of L2 education, an interesting question that would need to be considered if such a project were undertaken is whether it would be possible to develop an STI curriculum that would not be specific to any language but would operate at a sufficiently abstract level that it would pertain to all languages (e.g. all languages are able to convey a sense of time but how this is realized varies from language to language, involving adverbials, morphology, syntax, and so on). If a general concept-based curriculum were developed rather than one unique to every language, then the remaining work would involve determining how each of the concepts is operationalized in a given language.

Such extensive research programs notwithstanding, it is possible already to discern attempts to devise STI programs addressing more expansive aspects of communication that encompass a number of specific language features. For instance, framing his work within L2 sociopragmatics, van Compernolle (2012) implemented an STI intervention program organized around formality and address in interpersonal communication in L2 French. Through role-playing scenarios, during which conceptual materials including images and models were made available, learners were mediated toward an understanding of how alternating forms such as *tu* and *vous* could be employed for establishing and maintaining relations. The program then went on to consider the pragmatic implications for other structures, including negation, that can be deployed depending upon the relations that exist among speakers. For their part, Harsch and Poehner (2016) took up the concept of
intercultural competence as a target for mediation. In a project involving international students studying in the United Kingdom, individuals worked in pairs to interpret and comment upon ‘critical incidents,’ short vignettes of intercultural encounters and cultural misunderstandings that had been empirically derived from actual accounts given by other students. Students were directed to verbalize their understanding of the issue that was brought out by the critical incident and to discuss how the individuals in the incident either shared a mutual understanding of the issue or appeared to be at odds. They further worked to anticipate emotional responses of the participants and to propose how the situations might be negotiated. As this was organized as a DA, a mediator was present to offer prompts and feedback, to urge the students to extend their thinking, and to guide their thinking. As Harsch and Poehner (2016) explain, the procedure offers possibilities for determining individual students’ existing level of interculturality and their specific learning needs at the start of study-abroad experiences so that appropriate instructional support programs might be designed to help them move toward program objectives. Given the increasingly recognized importance of interculturality as a dimension of advanced language proficiency and of immersion experiences abroad as a privileged site for learner L2 development, it would seem that this is an area with considerable potential for future investigation and intervention.

Finally, Poehner and Swain’s (2016) analysis of Vygotsky’s proposal of human psychological functioning as comprising a cognitive-emotive unity has implications both for uses of SCT to understand processes of L2 proficiency development and for pedagogical efforts to mediate learners toward advanced proficiency. Vygotsky’s theoretical discussions of *perezhivanie* (lived experience), which he understood to be essential to understanding the integrity of individuals and their environment, have recently been attracting attention among SCT scholars as a corrective to the historical bias toward studying cognitive processes in isolation from emotions (see Mok, 2015). Poehner and Swain (2016) analyzed interactions from an L2 MD program and highlighted mediator efforts to guide learner decision-making and linguistic choices in a manner attuned to the student’s feelings of frustration as well as her sense of competence as a legitimate user of the L2. They also identified instances when the mediator’s attention shifted from cognitive aspects of the student’s performance and explicitly addressed the emotive, providing encouragement and recognition of her successes. Here again we see potential for further work to guide learners toward higher levels of proficiency.

REFERENCES


Part II  Advanced Proficiency and Performance: Multiple Dimensions and Contexts
8 Advanced-Level Grammatical Development in Instructed SLA

HEIDI BYRNES
Georgetown University

Introduction

By convention, a handbook chapter identifies the major themes of its chosen topic and then presents essentially settled findings. Such an approach presumes a stability of knowledge about grammar at advanced levels of L2 proficiency that does not presently exist, not least because remarkably little focused attention has thus far been devoted to the topic. But this lacuna opens up the unique opportunity for the language studies field to examine its fundamental assumptions about language learning in a way that reconceptualizes the entire extended sequence, from beginning to advanced levels, doing so in an integrated fashion. Three perspectives are crucial for accomplishing that task: (i) instructed language learning as a usefully delimited area of observation and action with regard to language learning, (ii) learners as key actors as they traverse the extended path of learning another language, and (iii) ‘grammar’ as a powerful observational lens. At the complex intersection of these three domains we should be able to understand more fully the role of grammar in all instructed language learning, but especially at the advanced level.
Problematicizing the interrelationships among advanced abilities, grammar, and instructed learning

The challenge of grammar

No matter what else we consider language to be, it is a complex and dynamically adaptive social-semiotic system of meaning-making, a quality that resides in the nature and interrelatedness of its lexicogrammatical resources – in other words, in its grammar (cf. Ellis & Larsen-Freeman, 2009). That fact justifies adopting a grammatical lens on substantive grounds. At the same time, an optic of grammar also reveals the existence of multiple descriptions and analyses of the grammar underlying real languages (Atkinson, 2011); each of these holds different epistemological and ontological positions, and each of them leads to different understandings regarding the nature of language learning—naturalistic as well as instructed learning.

Using Halliday’s (1996) distinction between grammatical descriptions, which he termed grammatics, and the phenomenon of grammar itself, I propose the following broad characterization of a grammatics suitable for advanced levels of performance: It should enable language professionals to imagine, research, and describe how language learners, over long sequences of learning, engage at multiple levels—cognitively, emotionally, socially, linguistically—with language learning as a meaning-making process. Because such an engagement is anything but additive and linear, such a grammatics must, in its very conceptualization, project a dynamic quality. In fact, dynamicity is likely to be particularly prominent at advanced performance levels, as competent L2 language users variably position themselves as languaging individuals who seek to use to greatest advantage their expanding linguistic, cognitive, emotional, and socially attuned resources in order to ‘mean more’ of what they, as individual actors, want to mean and can mean in a particular communicative situation. As they do so, we can imagine language learners engaging with others at three mutually interdependent levels: the micro level of social activity, the meso level of sociocultural institutions, and the macro level of ideological structures (The Douglas Fir Group, 2016).

More specifically, a grammatics for understanding and fostering advanced levels of proficiency should (i) describe how cognitively mature adults are, increasingly, able to express their meaning-making intentions and purposes because of and with their expanding L2 linguistic resources (Halliday & Matthiessen, 1999); (ii) consider how grammatical categories in the multiple languages to which the learners have access might influence their thinking for speaking (or writing) as they go about selecting and organizing information in light of the differential ways of encoding meaning that languages make available (Bylund & Athanasopoulos, 2015; Carroll & Lambert, 2006; Carroll, Murcia-Serra, Watorek, & Bendiscioli, 2000); (iii) incorporate options because languages are organized internally to allow for systemic choice, a quality that is the foundation of learners’ ability subsequently to make socially situated choices (Hasan, 2011); (iv) highlight the transparency of major grammatical categories, that is, the possibility of different movements
between and among them (e.g. grammaticalization, lexicalization, derivation, rank-shifting, embedding, grammatical metaphor); (v) characterize the complex interrelatedness of all grammatical categories in terms of complementarities, overlaps, and a certain openness of categories, rather than in terms of independent rules (Halliday, 2000); and (vi) present linguistic resources as patterns and bundles of probabilities and likelihoods that become ‘real’ meanings as these resources are enacted in particular communicative events (Matthiessen, 2009).

As is evident, for the proposed grammatics of advanced learning dynamic flexibility is fundamental, not an addendum. However, the picture is more expansive yet. To be sure, we rarely describe beginning-level learners as making deliberate and noteworthy semiotic choices; and yet, that is what they do as well, though with a different scope and depth and to a different degree. Such a perspective opens up the possibility of taking an integrated and expansive, long-term view of instructed language learning: No matter their performance level, all language learners continually engage in making and fine-tuning form–meaning and meaning–form associations in creative recalibration and approximation of fluid bidirectional linkages as they strive, over extended periods of learning, to expand their registerial and generic repertoire of language use (Matthiessen, 2006).

Instructed language learning tends not to be described in this way. This is so because our choice of grammatics, by and large, has not foregrounded meaning-making as the fundamental quality of language and language learning. In an educational setting the resulting constricted orientation has particularly serious consequences for the following reason: Instructed learning cannot draw on the myriads of mediating and channeling resources of the larger sociocultural context for consequential meaning-making in the same ways that naturalistic learning does (for discussion, see the Douglas Fir Group paper, 2016). But rather than concluding that instructed learning—and learners (!)—are therefore inherently deficient, we should seek to assure that our choice of grammatics captures as well as possible the inherently dynamic and socially embedded multifunctionality of language as a semiotic system. Indeed, for an educational program, such a dynamic, meaning-oriented grammatics is the foundation for a facilitative overarching educational philosophy and helps create a suitable frame of reference for its educative actions.

**The challenge and promise of advanced capacities**

As the language studies field now deliberately aspires to enabling instructed learners to reach advanced levels of ability, a meaning-oriented grammatics becomes indispensable. A crucial first step requires us to acknowledge that the discipline’s conceptual apparatus, vision of language learning, and scholarly habitus largely build on, validate, and naturalize a restricted optic on intermediate levels of performance that has sidelined considerations of the nature of meaning-making. However, current publication activity is beginning to query exactly that stance, even if indirectly, inasmuch as a rash of meta-analyses and comprehensive summaries in one major research area after another points to a saturatedness and
potentially unproductive self-absorption in the received categories of SLA. While that inquiry appears primarily to be a call for greater methodological rigor, at a deeper level it invites fundamental reflection on whether our research practices envision and portray language learning and the learner, especially the instructed learner, in a fashion that is able to describe—not to mention foster—the kind of competent language use that we associate with advancedness—surely both a desirable and a reasonable goal for instructed learning. Indeed, it would seem that the now increasingly accepted interpretation of the language learning process as dynamic, complex, and adaptive (Ellis & Larsen-Freeman, 2009) is poorly captured, perhaps deeply misrepresented, when its primary descriptors employ such metaphors as ‘skills’ by a ‘processor’ who faces multiple ‘processing constraints’ in dealing with ‘input’ and ‘output,’ and, fundamentally, is unable to handle meaning and form simultaneously. In the end, the person behind those processes is a pervasively deficient construct, at best suitable for clinical, decontextualized, often experimental investigation (cf. Keck & Ortega, cited in Keck & Kim, 2014, p. 142) but not for capturing the lively development of meaning-making repertoires that characterizes classrooms just like it characterizes naturalistic learning (Hall, 2017).

**The crucible of the context of instructed learning**

Finally, a dynamic view of language learning stipulates it as context-dependent (Larsen-Freeman, 2017). Far from serving merely as a non-distinct backdrop, educational contexts will involve numerous and diverse factors interacting in complex ways, leading to waxing and waning influences on learners, facilitative and obstructive encounters, emerging abilities that need time to develop fully, and considerable intra- and inter-individual variation in the effects of instruction. As already suggested, the first charge to language programs that want to enable their learners to link form, meaning, and use, then, is not to replicate the ‘real’ world of naturalistic learning; rather, it is to consider what kind of grammatics they need for creating an extended, articulated, conceptually integrated learning experience that supports learners in attaining advanced levels of L2 ability in an intellectually substantive and programmatic efficiency fashion. A dynamic, meaning-oriented description of grammar that takes use and function in social contexts as having shaped the very nature of the lexicogrammatical resources of a language offers a powerful, perhaps indispensable resource for working effectively toward those educational goals (cf. Byrnes, 2015, 2017; see also Chapter 2, Systemic Functional Linguistics and Advanced Second Language Proficiency).

As we foreground the openness and non-finality of language and language learning, three long-standing challenges for advanced learning can be reassessed: fossilization, the native speaker norm, and aptitude. First, fossilization is frequently invoked as a pervasive problem against which pedagogies must guard with particular vigilance and considerable energy in order to eradicate primarily morphological inaccuracies of surface features. That interpretation not only fails to recognize that the original theorization regarding fossilization addressed learners’
potentially limited ability to process discourse syntax and semantics (cf. Han, 2012); more damaging, it prematurely reiterates a preference for a deficient interpretation of the learner over considerations about the affordances that instructed settings could and should provide. Second, a dynamic non-finality and openness also questions holding up native speaker abilities as the gold standard for successful language learning; rather, it recognizes the inherently multicompetent nature of learning another language (Cook, 2013), rejects a teleological understanding of language development (Larsen-Freeman, 2014), and embraces the multilingual turn (Ortega, 2014). Aptitude, the third construct, all too frequently has been enlisted to “explain” (away) both learning success and failure, this although the typical classroom is not in the business of selecting its students by aptitude. More important, recent meta-analytical and summative research into aptitude points to at best “an overall moderate association with language learning” and that primarily in the initial stages of L2 development (Li, 2015, p. 385). Theoretically, that signals “that the concept of aptitude needs to be updated” (Skehan, 2015, p. 367); educationally, the construct’s ability to foster a deeper understanding of second language acquisition processes has yet to be ascertained, especially for more advanced levels of proficiency.

To summarize: I have laid out a framework that argues for a dynamic interrelationship among grammar, advancedness, and the context of instructed learning in order to understand more fully not only advanced-level grammatical development and advanced abilities but the longer trajectory of language learning. That orientation informs this review in its treatment and interpretation of existing research and in its inflection toward approaches that explicitly address central aspects of advanced-level proficiency in a theoretically principled manner.

Advancedness, where art thou?

The question of what constitutes advanced levels of L2 ability receives a wide range of responses. It may refer to foreign language (FL) learners who have just completed their language requirements; L2 users whose high score on formal testing regimes, such as the TOEFL (Test of English as a Foreign Language) or CEFR (Common European Framework of Reference for Languages) scale, qualifies them for inclusion; undergraduate and graduate students completing their degree requirements in a non-native language; language learners involved in high-level academic and professional language use; and, of course, heritage learners (see Chapter 10, The Prior Language Experience of Heritage Bilinguals; e.g. Achugar & Colombi, 2008; Brown & Bown, 2015; Byrnes, 2002, 2012; Byrnes & Maxim, 2004; Byrnes, Weger-Guntharp, & Sprang, 2006; Colombi & Schleppegrell, 2002; Leaver & Shekhtman, 2002; Ortega & Byrnes, 2008a). An overall descriptor of their abilities frequently states that advanced learners have acquired the major morphological and syntactic patterns of a language, a facility that results in a certain independence of language use. However, how to translate this considerable range of abilities and the notion of ‘independence’ or ‘proficiency’ into theoretically substantive and
educationally usable grammatical categories is far from clear. This chapter presents diverse foci that have been deployed while recognizing that, to date, advanced performance and the role of grammar in describing and facilitating it has received at best peripheral and incidental attention.

**Understanding and addressing local-level phenomena**

*Positioning accuracy, responding from rule-based notions of grammar*

An obvious focus regarding grammar at the advanced level addresses grammatical forms that have already been ‘learned’ or ‘should’ have been learned, with a focus on accuracy, complexity, sophistication, mastery/control of patterns, and expansion of repertoire. Among early studies looking into the syntactic and morphological accuracy of advanced L2 learners is Bardovi-Harlig and Bofman’s (1989) investigation of L2 English learners from five different language backgrounds. Its finding of a similar pattern of errors across languages seemed to point to a “universal design feature of language” and, more broadly, to a relative strength in syntax and relative weakness in morphology, inherently a language-specific phenomenon. In the end, the researchers reiterated Slobin’s stance that “neither the communicative position nor the formal position—both learner-centered positions—is necessary to interpret the strong syntax–weak morphology stage exhibited by the advanced ESL learners of the study” (p. 17)—perhaps a theoretically satisfying but an educationally inconclusive and ambivalent interpretation of the findings.

Educators, for their turn, have long interpreted this well-known state of affairs as requiring active remediation of a range of deficiencies (e.g. errors, overgeneralization, inappropriate use/misuse, defective rule application, unawareness of the scope/limitations of previously learned ‘rules,’ even fossilization; Richards, 2008), making situated pedagogical decisions in light of their professional experience. While researchers endeavored to offer pedagogical guidance, not surprisingly, advanced ability levels, in their very complexity and variability, rarely informed such pronouncements (Hulstijn, 1995). All too often, they merely offered commonsense recommendations or restated dominant psycholinguistically informed positions with little concern for their suitability across the long trajectory of language learning toward advancedness (Ellis, 2006). As a result, educators at the advanced level were simultaneously aware that the old ‘rules of grammar’ were unhelpful even as they saw that ‘grammar’ most definitely mattered at the advanced level.

With little guidance from theory and the research community, educators arrived at their own workable general pedagogical principles, among them: (i) expanding known forms with their presumably common meanings in common contexts to new meanings, that is, working with the rampant fact of polysemy that is particularly high with the use of some of the most common words in different contexts;
after all, it was increasingly clear that the meaning of grammatical phenomena was really located in larger syntactic units (run up a hill, run up a bill); (ii) taking ‘known meanings,’ such as modal meanings expressed through modal verbs, and introducing new forms, such as modal adverbs (possibly, perhaps) (cf. McCarthy, 2016); and (iii) working with late-acquired structures, among them relativization, subjunctives, and passive constructions.

Repeatedly, classroom realities and theoretical constructs met up in uncomfortable ways. For example, what constitutes a late-acquired phenomenon is not at all straightforward. If even beginning learners of German are quite capable of acquiring relativization (Byrnes & Sinicrope, 2008), a designation of ‘late-acquired’ is at least misleading. It becomes troublesome when it validates, through the power of theories, learner limitations. Among them are theories about presumed difficulty, as they are expressed, for instance, in the Noun Phrase Accessibility Hierarchy, or in the Perceptual Difficulty Hypothesis, or Pienemann’s process-ability hierarchy. And it becomes counterproductive when evidence from learner data (see Byrnes & Sinicrope, 2008) uncovers a complex set of relationships that include as particularly prominent features dynamic and shifting relationships between syntactic and lexicogrammatical features, genre-related probabilities of language use at the simplex and complex clausal level, and, within that set of options, learner choices that show significant variation for different learners, even though certain “syndromes” … register- and genre-related bundles of features, are discernible. (p. 134)

Turning to the construct of ‘difficulty’ seemed to offer a somewhat more sympathetic reading of the challenges learners (and teachers) faced, particularly when it was given a cross-linguistic, comparative inflection. Such an approach to the teaching of grammar—the what, the when, and the how—of course has a long history in the language studies field, in fact might be said to inform all approaches, no matter their theoretical persuasion or pedagogical realization. By taking a long view on learning, surprising opportunities for more sophisticated ways of understanding the teaching of ‘difficult’ grammatical phenomena seemed to open up, which pointed to multiple factors as influencing those difficulties, including the learning experience itself (for the difficulties in defining ‘difficulty,’ see Collins, Trofimovich, White, Cardoso, & Horst, 2009). As DeKeyser (2016) concludes: “What we need, then, is research that does not ask about the beneficial effect of aptitudes or instructional contexts or the difficulty of particular structures, but that looks at the interaction of two or more of these variables. Many paradoxical findings in the SLA literature stem from the lack of interaction research” (p. 358).

Taking an explicitly cognitive approach that highlights the importance of salience (psychophysical salience, salient associations, and context and surprisal) provided yet another way to look at core issues in long-term instructed learning. Once again, it foregrounds the multiple interactions among linguistic complexity, cognitive difficulty, individual differences, and instruction and development, all phenomena that teaching at the advanced level puts in plain view: “Salience is adaptively complex, involving multiple agents at multiple levels in interaction.
We need to acknowledge this complexity and to adopt theoretical perspectives concerning language emergence, dynamic systems theory, and language as a complex adaptive system” (Ellis, 2016, p. 349).

One can safely assume that Master’s (1997) expansive pedagogical framework for teaching the English articles *a*, *the*, and the zero article, from beginning to very advanced levels, did not get its inspiration from considerations such as these. Nonetheless, his conclusion that “the complex, multi-componential nature of the English articles requires that they be introduced gradually and over a long period of time and that they cannot possibly be taught in a single lesson” (p. 228) is a down-to-earth translation of such thinking into an extended program context.

**Exploring the complex embeddedness of grammatical features**

As already stated, the opportunity to address complex symbiotic interrelationships in the area of grammar is dramatically increased under a longitudinal trajectory. For example, Collentine, Collentine, Clark, and Friginal (2002) focused on how awareness of the larger syntactic environment enhanced second semester Spanish students’ ability to benefit from subjunctive instruction. While learners’ enhanced ability “to identify the syntactic conditions where the subjunctive was not a candidate” did not also translate into better mood selection overall, the researchers rightly speculated that the instructional effects might have been magnified with advanced-level students, thereby encouraging a complex relational understanding of grammatical phenomena from the earliest stages of language learning.

What insights might be gained when morphosyntactic issues are contextualized at higher performance levels is well illustrated by Geeslin’s (2003) extensive treatment of copula choice between *ser* and *estar* with data from advanced and native users of Spanish. The following issues came to the fore: (i) the variability regarding language choices not only as the ‘negative/flawed’ variability of instructed learners, but as inherent to contextualized language use; among the native users, “more than half of the test items produced variation” (p. 739); (ii) the centrality of choice, a choice that is affected by a host of features, among them a range of semantic and pragmatic features of the discourse context, in addition to variation based on the lexical and semantic features of the adjectives themselves (cf. p. 726); (iii) the possibility that remaining inaccuracies by advanced learners are unlikely to be addressed with pedagogies that disregard the multiple linguistic influences on form selection. Taken together, these insights thoroughly reposition earlier notions of accuracy, learner deficits, learner development, learners’ engagement as meaning-makers and, ultimately, appropriate learning goals—and, simultaneously, question traditional notions of grammar. At the same time, they leave little room for the kind of non-engagement with regard to complex grammatical phenomena on the part of faculty in many an FL literary-cultural studies program that Polio and Zyzik (2009) documented.
There is ample evidence that other constructs show similar system-wide entanglements. For example, Schumacher (2011) shows that German, generally not considered an aspect language, nevertheless expresses aspectual meanings morphologically, not only lexically. Specifically, the meaning differences between the simple past (Imperfekt) and the present perfect (Perfekt) are generally relegated to usage in written narratives versus oral telling or whether a narrative does or does not express speaker/writer distance. However, in certain contexts, the present perfect clearly carries the aspeactual meaning of ongoing conditions, which can be further reinforced through co-occurrence of a less frequent version of the passive that conveys the sense of condition (Zustandspassiv).

Given this complex interrelatedness of morphological features across larger expanses of language, across several linguistic strata, and into the pragmatics of language use, Geeslin’s (2003) conclusion is noteworthy:

For second language acquisition researchers, this means that learner language is not acquired by gradually applying a simple rule correctly. Instead, stages of acquisition are marked by the incorporation of new linguistic details at the levels of syntax, semantics, or pragmatics and a continuous reranking of the importance of these features. Thus, noticing a contextual feature and its relationship to a particular form (e.g., knowing that individual frames of reference go with estar) does not mean that this information will lead to native-like use, because frame of reference is never the only contextual feature present, and it can be “overruled” by other discourse features. Thus, learners must gradually adjust each constraint in light of how it interacts with several others. (pp. 750–751)

**Observing advanced-level grammatical abilities through the CAF lens**

One of the most prolific research agendas in SLA has focused on issues of complexity, accuracy, and fluency, commonly abbreviated as CAF, in relationship to the performance of both oral and written tasks. While advanced ability levels were neither its impetus nor its focus, over time issues of advancedness could not but come to the fore. Specifically, addressing processing at the beginning/intermediate levels, the two dominant, psycholinguistically oriented theoretical edifices—Robinson’s Cognition Hypothesis (Robinson, 2011) and Skehan’s Limited Attentional Capacity Model (Skehan & Foster, 2001)—set forth very different hypotheses regarding the effects of task complexity on accuracy. At the same time, both subscribe to a zero sum interpretation of the availability of attentional resources to either form or meaning, one or the other. In light of the previous discussion regarding the critical importance of highly nuanced choices on the part of advanced learners that inherently involve the simultaneous consideration of form and meaning, not to mention socially situated use, the findings obtained in CAF research require nuanced interpretation with regard to our understanding of advanced-level phenomena.
One such note of caution is voiced by Serafini and Sanz (2016) when they uncover evidence for a decreasing impact of cognitive ability (e.g. working memory capacity) as proficiency increases. Three vignettes further illustrate that point: First, in a comprehensive overview of complexity research, Bulté and Housen (2012) conclude that complexity as handled in CAF research “lacks adequate definitions supported by theories of linguistics, cognition or language learning,” a situation that is further exacerbated by shortcomings in its operationalization which makes problematic whether it “has been, and can be, validly, reliably and efficiently measured in empirical research” (p. 41). The conclusion of the volume’s editors is then particularly sobering: As research constructs, complexity, accuracy, and fluency would seem to be useful and valid but that “this is where the consensus ends and the controversy begins” (Housen, Kuiken, & Vedder, 2012, p. 300).

By contrast, a focus on advancedness might assess matters quite differently. To begin with, the subtly assumed linear relationship between syntactic complexity and proficiency must be rejected inasmuch as ‘complexification’ may occur in different subsystems of the lexicogrammatical system (phrase, clause, sentence) at different levels of proficiency, and therefore is less or more predictive of development at different locations of the overall learning trajectory (Byrnes, 2009, 2014). Furthermore it is becoming increasingly apparent that whatever we mean by ‘complexity’ will require a functional, that is, a meaning orientation that foregrounds the consequences of communication demands, has both linguistic and cognitive-semantic underpinnings, and must consider instructed development, L1, proficiency level, modality, and genre/task/content in mutually interrelated ways (Byrnes, 2014; Ortega, 2015).

Second, studies of diverse aspects of CAF more often than not address research questions that are primarily theory- and methodology-driven even as they invoke the educationally oriented construct of task. But, as Van den Branden, Bygate, and Norris (2009) note, the main focus has been “the use of tasks to reveal iterative answers to highly theoretical questions, generally focused on immediate processes considered to contribute to language acquisition” (p. 8). Largely missing are studies that would show how the instructional use of tasks “enables the student to acquire language ability and apply it in performing a variety of authentic tasks to criterion” (p. 8). While linking a discussion of grammar with task-based teaching might appear to be jarringly incongruous, that is true only with a narrow interpretation of grammar. In fact, a task-based approach in particular, with its keen interest in real-world performance abilities, should benefit from meaning- and use-oriented approaches to grammar that would uncover the situated potency of available grammatical resources (e.g. Byrnes, 2015; Byrnes, Maxim, & Norris, 2010; Colombi, 2006; de Oliveira & Schleppegrell, 2015; Ryshina-Pankova, 2015; Schleppegrell, 2004, 2006).

Third, the presence and effects of intricately interrelated phenomena is foregrounded particularly insistently by Norris and Ortega (2009) in their call for an organic approach to the investigation of CAF. A much-cited article, it is rightly interpreted as a way to strengthen approaches to measurement in CAF research
toward multidimensional and multifaceted as contrasted with unidimensional measures; it also foregrounds a need to attend to dynamicity and variability in CAF research as contrasted with linearity or even collinearity. If one also heeds the authors’ concerns about an impoverished operationalization of CAF that fails to recognize the reality of “a dynamic and interrelated set of constantly changing subsystems” (p. 555, emphasis added), the real question is what we take to be the nature of those subsystems and how they might be dynamically interrelated in the ecology of long-term instructed learning.

Understanding advancedness as a discourse phenomenon

Asserting and positioning a shift to discourse

Understanding advancedness as a discourse phenomenon has a long history in the language studies field, especially in the environment of pedagogical recommendations (for an early treatment, see McCarthy & Carter, 1994). There are interesting parallels to the current situation when, in the earlier shift from form-focused to more communicative language teaching, Celce-Murcia (1991) observed that the real issues were not so much methodological in nature in terms of when and how to teach grammar. Rather, they were ultimately conceptual inasmuch as “grammar should never be taught as an end in itself but always with reference to meaning, social factors, or discourse—or a combination of these factors” (pp. 466–467). While advancedness is now generally taken to be a discourse phenomenon (Hughes & McCarthy, 1998; Paltridge, 2012), the deeper theoretical and educational implications of such a shift remain to be addressed.

From ‘discourse’ to ‘genre’

To begin with, because ‘discourse’ is difficult to translate into educational practice, the real impact on professional thinking has been through the construct ‘genre’: Since the appearance of Swales’s (1990) Genre analysis, there has been a veritable avalanche of genre-related publications. Most frequently focused on writing, they link advancedness with diverse genres in different disciplinary contexts, often bundled under English for Specific purposes (ESP; Belcher, Johns, & Paltridge, 2011; Parodi, 2010), or English for academic purposes (EAP; Paltridge, 2004); to various professional contexts (e.g. business, medicine, law; Swales, 2000), genre use in the classroom (Johns, 2002); and to specific components of genres, such as introductions, summaries, and conclusions. More recently, investigations have incorporated such foci as the role of metadiscourse in various contexts of writing (Hyland, 2005) and identity construction in academic discourse (Flowerdew & Wang, 2015). Both descriptive analyses and more writing instructional publications (Candlin, Crompton, & Hatim, 2016) can readily be found.
Implications of a genre orientation for advanced instructed language learning

Given the enormous range of contexts where the construct of genre is applied, it comes as no surprise that it has received multiple definitions (for an excellent and accessible summative treatment, see Tardy’s introduction, 2011, to a guest edited issue on genre use in the North American context). For the present discussion two perspectives that genre foregrounds are particularly important: its tendency to focus on more advanced performance levels and its admittedly often indirect probing into what role grammar should be assigned throughout all of language learning. I explore them in two steps, addressing more programmatically and pedagogically oriented issues in this and theoretical concerns in the subsequent section.

First, moving into discourse and genre questions the orthodoxy of communicative language teaching and its focus on a transactional orality. Among other shortcomings it had presented the system of language in severely truncated form by ignoring the expansion in meaning-making resources that is at the heart of literate languages and the diverse culturally embedded genres that we take to be part of an advanced learner’s repertoire. Importantly, these genres are formalized and normalized through educational systems and are transmitted through educative activities. Communicative language teaching had done this by privileging a concern with sentence-level phenomena, anchoring them in an overtly dialogical environment through interaction-based approaches to pedagogy and the claimed benefits for language learning and development of ‘the negotiation of meaning,’ of ‘noticing the gap’ or moving from ‘input’ and meaning-oriented processing to ‘output’ and syntactically oriented processing.

Along with implicit and explicit learning these became central constructs that, at least discursively if not in reality, established the meaning-oriented bona fides of such an approach. That left little room and little need to imagine and appreciate the fundamental dialogicality of all communication, including, quite emphatically, written texts with their distinctively different set of resources for realizing dialogicality. As Bakhtin (1986), using the distinction between primary and secondary (ideological) genres states: “a one-sided orientation toward primary genres inevitably leads to a vulgarization of the entire problem (behaviorist linguistics is an extreme example)” (p. 62). In the end, that shortened optic ignored the movement from orality to literacy and the pivotal role that education plays in enabling L1 and L2 language users to develop the linguistic resources necessary to expand their meaning-making capacities from the commonsense knowledge of everyday life to general educational knowledge, to disciplinary knowledge, each with very specific linguistic manifestations (Halliday, 1993, 2002).

Such sidelining has serious consequences for teaching at the advanced level. As Halliday (1999) notes, the specialized knowledge structures of the disciplines, represented in their preferred genres, are highly technical, with multiple layers of abstraction, with a range of linguistic resources for value judgments (particularly in the humanities), and the possibility of expressing technicality, not just with
technical terms but with a reconstrual of reality through a synoptic form of semi-
osification, as contrasted with the flow of everyday life expressed verbally (see also below, and Chapter 2, Systemic Functional Linguistics and Advanced Second Language Proficiency). What stands out is that "educational knowledge is almost entirely construed in language; even when it is presented in other forms" (p. 85) (e.g. scientific diagrams, mathematical formulas), because the learning process itself verbalizes them and draws on specific lexicogrammatical resources to do so. Unless language is understood and analyzed that way and also taught that way, not least by pointing out how grammatical phenomena are bundled in particular genres, the claim that learning another language is ultimately about learning new ways of thinking and knowing is quite literally 'unthinkable' (but see Coffin & Donohue, 2014; Halliday, 2004; Martin, 1991, 2007; Ryshina-Pankova, 2015; Ryshina-Pankova & Byrnes, 2013; Schleppegrell, Achugar, & Oteiza, 2004).

Though largely atheoretical in outlook, contemporary corpus linguistics, particularly the insights gained through multidimensional register analysis, followed by the statistical procedure of factor analysis, provides important data on just such matters. It has readily confirmed that written discourse has a much greater range of linguistic variation, most likely pointing as well to the availability of a much greater range of resources than is available in spoken language, including prefabricated patterns (e.g. among others, Biber, 1988; Biber & Conrad, 2009; Biber & Gray, 2016).

Second, the extensive engagement with diverse genres has led to an increased understanding of the genre- and discipline-specific nature of the lexicogrammatical resources that advanced language users have (and need) for competent language use in various cultural contexts. Shaw's (2012) survey of the formal features that characterize academic writing provides a useful summary of diverse features in different genres. Descriptions typically begin with Halliday and Hasan’s (1976) treatment of coherence and cohesion in terms of the major cohesive devices of reference, substitution, ellipsis, and conjunction. Thereafter, they highlight such features as a restricted set of tenses and somewhat specialized functions for them; predominance of the third person of verbs and specialized uses of first-person I/we; frequency of the passive voice; deverbalization (with a lower rate of finite verb expressions and a rise in participle phrases, nouns, and adjectives that encapsulate similar areas of meaning); specific functions of modality to express hedging of claims; complex nominals with prenominal and postnominal modifications; diverse aspects of cohesion; engaging with the reader and engaging the reader; handling intertextuality, which involves variously incorporating knowledge presented in other texts, for example, by accepting, extending, or refuting it in order to bolster one’s own claims; and, ultimately, taking a stance on a particular topic.

More expansive, because it goes beyond academic writing, is Crossley, Roscoe, & McNamara’s (2014) cluster analysis. Focusing on successful L1 writing, the researchers arrived at four different successful writing styles: action and depiction style, academic style, accessible style, and lexical style. These styles show overlapping but nevertheless distinct linguistic indices, among them: affective and
perceptual processes; causal features; cognitive processes and mechanisms; cohesion; content features; descriptors; and diverse lexical indices. They also show that what constitutes successful writing is multidimensional and cannot be described through the presence or absence of only certain predefined features. With regard to grammar and its teaching, the analysis points to the need for an expansive and differentiated understanding of the interrelationship among all lexicogrammatical resources that a language makes available, followed by pedagogies that are carefully contextualized with regard to major text types and, ultimately, specific genres (see also Jarvis, Grant, Bikowski, & Ferris, 2003, for a similar investigation of L2 writing).

Further substantiation for an understanding of language as a semiotic resource rather than a fixed inventory and for the increasing importance of language users’ situated choices in competent language use, is available in Berman and Verhoeven’s (2002) study across seven languages that investigated both speech and writing with children at different ages in a semi-longitudinal design. The authors highlight this overarching fact: “the ability to recruit linguistic forms in the context of extended discourse … has a long developmental history” (p. 14). They emphasize as well that an understanding of text structuring and content (e.g. in narrative as compared with expository texts), propositional attitudes, and discourse stance requires a focus on a range of complex interactions that show shared developmental patterns but also increasing within-genre variation with more mature speaker-writers who “deploy a wider variety of lexical, syntactic, and rhetorical devices for differentiating stance; and they also make more metatextual commentary” (p. 37).

**Toward a functional understanding of grammar**

A discourse turn in the language studies field, further circumscribed through the construct of genre, requires yet another step, this time into theoretical territory. As already suggested, quite different assumptions undergird the construct of ‘genre’ in various traditions. Thus, where North American rhetorical approaches direct their attention to the social and institutional contexts within which genres occur, the ESP tradition and systemic functional linguistics (SFL) are more concerned with the linguistic features that realize particular genres (Hyon, 1996). But what is meant by ‘linguistic’ approaches itself needs scrutiny. Thus, it is not sufficient to advocate a social theory of genre with its concern for social context of situation, meaning, and form—applied linguistics has long had those interests. Yet, this is where the bulk of contemporary discourse and genre-oriented thinking is located, a position that makes it possible to uncover a seemingly unlimited number of interesting facts about what goes on in diverse textual environments but that, ultimately, leaves them as disjointed phenomena that are not conceptually organized into a larger framework.

The real issue, instead, is whether one assumes that language is extrinsically or intrinsically functional. In the case of extrinsic functionality, the lexicogrammatical
resources of a language ‘reflect’ but do not actually create meaning through system-embedded wording choices. If, however, one assumes languages to possess intrinsic functionality, one must uncover the ways in which a language’s grammatical system necessarily and inseparably describes contexts of language use, such as through three pervasive metafunctions: the social relations between participants in discourse, the social action that is achieved, and the management of the interaction as a social event. Such a metafunctional organization in terms of interpersonal, ideational, and textual metafunctions must, of course, be traceable in the very design of language as a social semiotic system, and therefore must pervasively inform the description and analysis of language, that is, the grammatics that the theory develops.

**Understanding genre within a functional grammar—implications for grammar teaching**

As detailed in Chapter 2 of this handbook, SFL, as a natural functional theory of language (Halliday, 2014), takes this latter position when it posits three strata internal to the system of language: Semantics as the system of meaning and lexicogrammar as the system of wording make up the content plane and phonology/graphemics make up the expression plane. There is a crucially important additional assumption of a solidary rather than an arbitrary relationship between semantics and lexicogrammar, which stipulates that meaning and lexicogrammatical form work together to articulate the constitutive features of social contexts (i.e. tenor, field, and mode). The internal strata of a language are organized in terms of systemic choices within system networks, rather than in terms of rules that are most conspicuously manifested at the textual level through the particular way in which genre is substantively theorized in SFL as contrasted with being variously described in other genre approaches (for details, see Chapter 2).

That theoretical anchor at the same time renders SFL distinctively suitable for teaching grammar at the advanced level. When genre is seen as a “recurrent configuration of meanings and a culture as a system of genres” (Martin, 2009, p. 13) and, furthermore, is interpreted in terms of the staged, goal-oriented, and social qualities of how it deploys its linguistic resources, it offers two critical components for successful learning: the necessary limiting context and a genuine invitation for dynamic meaning-making that encourages situated choices, thereby facilitating acts of meaning-making that are sufficiently bounded and sufficiently open—in other words, facilitating learning and development.

By now there is a wealth of research in the second language environment, both ESL and FL, which supports the feasibility, usability, and efficacy of language teaching with such an approach to genre and grammar. Importantly, the phenomena that have been addressed range from macro text-organizational to quite local-level issues of modality and mood or case use (Liamkina, 2008). I list the following as representative of that grand sweep: (i) the all-important movement from oral to written style, with the central characteristic of reconstrual of actions as
conceptual entities, especially through nominal structures and much higher information density (Achugar & Colombi, 2008; Halliday, 2002); (ii) diverse aspects of textuality and information structure, for example linguistic resources used for theme and rheme development, as defined in the theory (Halliday, 2000; Ryshina-Pankova, 2006); (iii) the importance of hyperthemes in order to uncover the argumentative structures of texts as a kind of chain of reasoning that is fostered, among other options, through expanding the clause complex by means of elaboration, extension, and enhancement (Ravelli, 2004); (iv) the function of periodicity as enabling texts to “create expectations by flagging forward and consolidate them by summarizing back” (Martin & Rose, 2007, p. 184), which leads to waves of information occurring in crests of textual prominence and troughs of lesser importance (akin to signaling nouns in ESP, e.g. Flowerdew & Forest, 2015); (v) the creation of an authorial voice by anticipating, in an imagined dialogue with likely readers, their possible responses, something that requires subtle management of appraisal and diverse forms of knowledge claims, particularly through the resources of epistemic modality (Coffin & Donohue, 2014); (vi) the importance of projection, a way of bringing in other voices from other texts (Woodward-Kron, 2009); (vii) the functional role of clause complexing as creating diverse logicosemantic relationships between clauses; (viii) extensive treatment of one of the key features of advanced abilities, namely the expansion of the noun clause through pre- and postnominal modifications (Fang, Schleppegrell, & Cox, 2006; Ryshina-Pankova & Byrnes, 2013); and (ix) a genre-oriented understanding of variable tense use, for instance in narratives (Byrnes & Sprang, 2004; Liamkina & Ryshina-Pankova, 2012).

A natural functional theory of language and advanced-level grammar teaching

By way of summary, I offer these characteristics of a natural functional theory, and of SFL in particular, as a way of understanding and fostering grammatical development at the advanced level in instructed settings.

The copiousness of the theory responds to a key concern, namely the enormous variation regarding needs and interests that characterizes teaching at the advanced level. It also recognizes the agency of learners as social personas who are meaning-makers. Through its two major forms of semiosis—congruent and synoptic—which are particularly well exemplified in grammatical metaphor, the theory inherently has the long-term developmental orientation that is required for understanding and enhancing grammatical performance over long instructional stretches. By drawing on a theoretical framework that explicitly and in a principled way links lexicogrammatical patterns to discourse semiotics and social use, teachers can enhance their abilities to teach grammar and learners can expand their meaning-capacities through grammatical categories that have been made transparent for their meaning-making qualities and have become available for shared metacognitive reflection. Such a differentiated metacognitive awareness of semiotic resources, always an unbreakable link of meaning and form and always
an invitation to make choices in use, is fundamentally different from the extensive (and ultimately unresolved and unproductive) deliberations about implicit and explicit teaching and learning and the implied threat to a highly valued automaticity of processing attributed to an awareness that facilitates meaningful choice. For advanced learners and advanced learning, these categories are unlikely to be insightful, not to mention productive for teaching and learning. Finally, because of the possibility of taking different pedagogical stances, for instance by incorporating cognitive-semantic approaches to teaching grammar at this level (e.g. Liamkina, 2008; Liamkina & Ryshina-Pankova, 2012), considering the contributions that usage-based approaches might make (Tyler, 2010), and drawing on corpus-based data (Flowerdew, 2009), a variety of pedagogical approaches becomes available.

Directions for future research

In considering useful directions for future research, I return to the overarching argument of this chapter, namely that a focus on advancedness and the role of grammar in instructed learning can facilitate a desirable expansion of horizons for the language studies field. Given the enormity of the project, one should expect and welcome contributions from different sectors whose unifying characteristics might be described like this:

• A functional meaning-orientation for the analysis of language, where a natural functional approach that explicitly theorizes and empirically researches the relationship between contexts of use and meaning as instantiated, in a principled way, in the lexicogrammatical resources of a language, is likely to be most productive;
• An emergentist ontology and epistemology that is inherently dynamic and privileges long-term development over short-term learning, bounded variability over rule-application, and situated choice over correctness;
• Serious engagement with (and, therefore, taking seriously) key dynamics of language programs, where the existence and substance of an educational philosophy regarding long-term learning, as manifested in differentiated curricular thinking and pedagogies, are critical if educationally useable findings are to result;
• Rethinking research foci and research methodologies, including what constitutes rigorous inquiry, once a longitudinal trajectory and the fundamental presence of situated choices by learners is no longer bartered away (Byrne & Callaghan, 2014; Hiver & Al-Hoorie, 2016; Larsen-Freeman, 2017; Norris & Manchón, 2012; Ortega & Byrnes, 2008b);
• Use of well-theorized and significant developmental turning-points in language development in order to set up research agendas that both justify and facilitate the enormous research effort involved in substantive longitudinal studies;
• Links with other areas of applied linguistics that take a functional approach, even if they do not necessarily build on an expansive theory of lexicogrammar, with foci ranging from vocabulary acquisition to pragmatics to corpus-linguistics;
• Incorporating new impulses regarding what constitutes valid assessment of language abilities, including grammatical abilities (Purpura, 2014; Timpe Laughlin, Wain, & Schmidgall, 2015);
• And, in any of these efforts, substantive not peripheral collaborative engagement by language educators.

REFERENCES


from http://jan.ucc.nau.edu/~jgc/research/indiana/.


Halliday, M. A. K. (2000). Grammar and daily life: Concurrence and


Larsen-Freeman, D. (2014). Another step to be taken—Rethinking the end point of the interlanguage continuum. In Z. Han & E. Tarone (Eds.), *Interlanguage: Forty years later* (pp. 203–220). Amsterdam: John Benjamins.


Advanced-Level Grammatical Development in Instructed SLA

language assessment (pp. 100–124). Oxford: John Wiley & Sons, Ltd.


9 Individual Differences in Advanced Proficiency

PAULA WINKE AND SUSAN M. GASS
Michigan State University

Introduction and literature review

Advanced language learners are needed globally to meet the needs of a multilingual and increasingly interconnected world (see Thorne & Reinhardt, 2008), but they are also needed locally to perform professional jobs that require advanced language competence (Aoki, 2013; Kennedy & Hansen, 2015; Moeller, 2013). Research has shown that students themselves desire advanced competencies to increase their professional employment prospects (Loomis, 2015). Language programs, recognizing the need for highly proficient speakers globally, locally, and individually, have (as one of their goals) the objective to produce language learners who obtain the status of advanced.

There are a number of ways of conceptualizing advancedness. One possibility is to think of a language learner who can use the language contextually in sophisticated ways (Ortega & Byrnes, 2008, p. 8). According to Ortega and Byrnes, advancedness is linked to “aspects of literacy, to diverse manifestations of cultural competence, choice among registers and multiple speech community repertoires, voice, and identity in cross-cultural communicative settings” (p. 8).

A second way to define the term (advanced) is to do so loosely. For example, advanced learners can be defined relative to their lower-level peers within the institution in which they are learning (Ortega & Byrnes, 2008, p. 7); that is, they can be seen, at the college-level, as students who have continued studying past the minimum requirement. Thus, third- and fourth-year college students and majors (who are past a two-year requirement) in the language may be seen as advanced language learners. Ortega and Byrnes noted that a more concrete way to define a learner as advanced is through well-designed, standardized testing, that is, “tests that come with conventionalized and publicly available regulations for
administration, scoring, and reporting” (p. 7). This is important because researchers and language teachers know that not all students in, for example, third- and fourth-year college courses are clinically advanced in their language skills. Students can reach third- and fourth-year classes by passing through the lower levels of coursework without obtaining high grades and without reaching linguistically and pragmatically advanced levels of communicative ability. This is because many universities, even when there is a specific language requirement, allow students to fulfill that requirement by seat time and by minimally receiving passing grades rather than through a proficiency outcome.

By concretely identifying, through standardized testing, which language learners are advanced, researchers can better investigate how to educate learners to the advanced level and beyond and can better understand the components of advancedness. Norris (2006) noted that after identifying advanced language learners through standardized tests, applied linguists can focus on “developing a curriculum that captures the complexity of the construct, and delivering instruction that helps learners to get there” (p. 167). Norris further suggested that applied linguists need to participate in “broad-based discussions about what educating for advanced FL [foreign language] capacities requires” (p. 168), and testing for advancedness is part of that participation and educational evaluation. Norris wrote that within language learning programs, applied linguists need to assess the learners with standardized measures to ensure that advanced learning is occurring and consequently help ensure that language programs are meeting their goals. Brown (2009) concurred, and wrote that “instructors must learn what constitutes advanced-level proficiency and must design their courses toward reaching those defined outcomes” (p. 535).

In the current study, we administered standardized tests of speaking, reading, and listening to uncover the proficiency levels of undergraduate students studying Chinese, French, Russian, and Spanish. The tests were independent from the curricula, and were parallel across the four languages. In all, we tested a pseudo-random sample of 1,311 students learning Chinese (N = 177), French (N = 377), Russian (N = 82), and Spanish (N = 675). In this chapter, we describe the few (147 out of 1,311) that tested at the advanced level or higher, and what characteristics those learners had. We did this to understand more fully what factors contribute to advanced learning proficiency.

A significant portion of our investigation involved an analysis of a detailed survey given to all students with the particular goal of understanding the individual differences that may contribute to their language learning trajectories. Many language programs rely (intentionally or not) on factors that may be external to the program to bring students up to an advanced level of language ability. For example, experience abroad in a country where the language is spoken natively is a factor that helps language learners achieve advanced levels of language performance, especially in speaking. Other factors, such as one’s heritage status, use of authentic L2 resources outside of class, and motivation also play a role in advanced-level attainment. Below, we briefly review these factors, as we will investigate these in combination with advanced-level status to best understand how advanced students come to be advanced in the language programs at hand.
Abroad experience

Without a doubt, studying abroad contributes greatly to the development of language learners’ linguistic and cultural competence. This is especially true when (as explained by Davidson, 2015, p. 117) “appropriate program designs and student support conditions [while abroad] are met” (Davidson, 2010; Kinginger, 2008). In studies by Davidson (2010, 2015), studying abroad increased language learners’ proficiency in a variety of target languages with very few individual exceptions (that most certainly could be accounted for through measurement error), and the impact of studying abroad was particularly great for learners of less commonly taught languages (native English speakers learning languages such as Arabic, Chinese, Russian, and Persian). As Davidson (2015) wrote, “it is difficult to construct a comprehensive curricular model within our existing educational system without recourse to one or more immersion models” (p. 145). Linguistic success can additionally be bolstered through study abroad by the development of lasting and satisfying relationships with host-country nationals while abroad (Spenader, 2008). Thus, even if specific or large language learning gains are not directly observable post study abroad, the idea is that motivation may be increased and learner autonomy may be fostered through the study-abroad experience, which in turn may help maximize one’s future language learning potential and sustain one’s motivation to continue learning to the advanced level.

Heritage status

There is no simple way to define heritage language learners. First, it is crucial to recognize the variance that often falls under the umbrella of heritage speaker. Montrul and Polinsky (2011, p. 59), citing Au, Knightly, Jun, and Oh (2002), noted, “[h]eritage speakers are notorious for the tremendous variance within their populations—from very high proficiency cases where some registers may be affected, to so-called overhearers.” And, at times, language proficiency is not even a central part of the definition. For example, according to Valdés (2001, p. 38), “it is the historical and personal connection to the language that is salient and not the actual proficiency of individual speakers. Armenian, for example, would be considered a heritage language for American students of Armenian ancestry even if the students were English-speaking monolinguals.”

For those with some level of language knowledge of their heritage language, it is clear that they have language learning behaviors and needs that are different from those of non-heritage foreign language students (Kondo-Brown, 2005). A standard definition is one by Valdés: “The term ‘heritage’ speaker is used to refer to a student who is raised in a home where a non-English language is spoken, who speaks or merely understands the heritage language, and who is to some degree bilingual in English and the heritage language.” Polinsky (2008) brought in language acquisition order as part of her definition of heritage language. In her definition the heritage language is the one “[w]hich was first for an individual
with respect to the order of acquisition but has not been completely acquired because of the switch to another dominant language. An individual may use the heritage language under certain conditions and understand it, but his/her primary language is a different one” (p. 1). Finally, Carreira (2013) considered heritage language learners as “speakers who acquire their first language in a home where a language other than that of the country of residence is spoken (p. 1). As described by Carreira, for heritage language learners, their first language (L1) acquisition is “interrupted as the language of the country is encountered and then becomes dominant, resulting in an incompletely acquired first language” (p. 1). However, when the heritage language learner learns the L1 in a foreign language learning context, his or her social conditions may favor the acquisition of the language. The heritage student may have a supportive home environment for learning the language, and the heritage learner may find a large and networked community of speakers that support his or her learning efforts. They may be advanced in their speaking already or in certain contexts, but their speaking may also be different from their non-heritage classroom peers, or even from native speakers of the language (see Kagan, 2012). As Kagan reviewed, their motivation for learning may differ considerably from their non-heritage peers. Thus, learning to the advanced level may be assisted by one’s heritage status, although, as reviewed by Kondo-Brown, not all studies have supported this view.

**Use of authentic L2 resources outside of class**

Researchers have long shown that access to and use of authentic L2 resources facilitate language learning. This is because, as Blake (2013) described, “L2 learning requires a lot of time on task. Crucial to this L2 processing is the extent and nature of the input received” (p. 2). Blake noted that increasing one’s access to the target language is the best way to increase input, and the most popular and familiar way to do this is through study abroad. But Blake noted that within the computer-assisted language learning (CALL) field, authentic materials through online resources can be used to foster autonomy in language learning without going abroad. Bardovi-Harlig and Bastos (2011) found in their study that one important aspect leading to successful conversational competence (i.e., the correct use of conventional expressions) was frequent interaction with speakers of the language outside of the formal learning environment. They noted that “social media and computer-mediated communication would be relevant additional sources of input in conversational registers” (p. 376). They speculated that things such as email, internet use, TV watching, and movies may factor into language learners’ development of advanced discourse competencies.

**Motivation**

As described by Dörnyei (2005, 2009), motivation is accepted by both teachers and researchers as a key factor that influences the rate and success of second and foreign language learning. It is what starts the language learning process and what
sustains it. It must be fostered, and it must change over the course of one’s language learning pathway. It can be mediated by the language learning environment, one’s peers, cultural and historical events, and by smaller units such as individual tasks in the language classroom. Motivation can be measured in multiple ways, but without a doubt, it contributes to the attainment of advanced-level language learning success. It may be a necessary precursor to such attainment, because studies have shown that without motivation, language learning tends not to occur (Dörnyei, 2009).

Taking into consideration the above-mentioned factors of (i) abroad experience, (ii) heritage status, (iii) use of authentic L2 resources outside of class time and (iv), motivation, in this study we investigate the profiles of advanced language learners in the Chinese, French, Russian, and Spanish language programs at our university. Through this study, we investigate the profiles of advanced language learners to see if any combinations of these factors map together to form predictable groups of advanced language learners. We follow advice from Ortega and Byrnes (2008) and Norris (2006) to use standardized testing to identify advanced learners so that we can then participate in the discussion about what educating for advanced capacities requires, especially within the context of the programs at our university. We also run analyses to determine which of the factors contribute most to predicting advanced learner profiles. The questions that guide this study are:

1. What are the profiles of advanced-level foreign language learners across the Chinese, French, Russian, and Spanish language programs?
2. How do the individual difference factors of (i) abroad experience, (ii) heritage status, (iii) use of authentic L2 resources outside of class time, and (iv) motivation contribute to the various profiles of advanced foreign language learners?

Method

Participants and materials  The data for this chapter come from a larger dataset that includes language testing of students studying French, Chinese, Russian, Spanish. We were the Principle Investigators on a federal grant from the National Security Education Program (NSEP) called the Language Proficiency Flagship (see https://celta.msu.edu/projects/proficiency-flagship/). For that project, we assessed the proficiency of more than 1,000 Chinese, French, Russian, and Spanish language learners in the first through fourth years of the undergraduate language programs. Testing occurred five times over a period of three years (fall 2014, spring 2015, fall 2015, spring 2016, and once more in spring 2017). The data for this chapter come from testing from spring 2016. The tests used were official American Council on the Teaching of Foreign Language (ACTFL) tests from Language Testing International (LTI, https://www.languagetesting.com/), the company that packages and sells official ACTFL proficiency tests. Chinese, French, Russian, and Spanish instructors took their intact classes into the computer labs during class times to have them first fill out a background questionnaire, and then to take the ACTFL Oral Proficiency Interview Computerized (OPIc). The instructors did this
during a six-week testing window during March and April of 2016. The students were then asked to come into the language lab outside of their class times but within the same window to take the ACTFL Reading Proficiency Test (RPT) and Listening Proficiency Test (LPT). In total, in spring 2016, 1,311 students across the four programs participated.

LTI reported all ACTFL test scores to us via the proficiency scale as defined by the ACTFL Proficiency Guidelines (Swender, Conrad, & Vicars, 2012). That means that for each individual and for each test, we received a score-band assignment such as Novice low, Novice mid, Novice high, and so on (up to Superior) as described within the Proficiency Guidelines 2012 document (ACTFL, 2012). We first selected from the LTI test score data file those who achieved an Advanced rating (Advanced low, Advanced mid, or Advanced high) or higher (Superior) on any individual test (or multiple tests). In total, 147 students obtained Advanced (from low to high) or Superior test scores in at least one of the three skills being measured: speaking, reading, and listening. The majority of these 147 had Advanced (and we mean Advanced low or higher) in reading \((N = 125)\); a smaller number also or separately achieved Advanced in listening \((N = 44)\) or speaking \((N = 43)\). We would like to note that ACTFL refers to the score band of Advanced with a capital A, but henceforth in this paper we will refer to Advanced learners as (plainly) advanced learners, and by that we mean any student who scored Advanced low or higher on an official ACTFL test or tests.

As noted earlier, all students were asked to fill out a background survey online as part of the test administration. The Michigan State University Flagship team designed the survey to obtain longitudinal information on a large variety of language-learner variables. For this study, we picked 15 survey questions to represent the four key variables we are interested in: (1) experience abroad in a country where the L2 is spoken natively, (2) identity as heritage/non-heritage speaker, (3) use of authentic L2 resources outside of class time and apart from language-class homework, and (4) motivation for learning the L2. Based on prior research (reviewed above), these four learner variables play an important role in the process of second language acquisition, and may help with the learners’ ultimate attainment of L2 proficiency. Table 9.1 contains the four variables and the 15 survey questions that tapped into these variables.

Using the “Merge Files” method provided in SPSS 22.0, we added the scores or responses from these 15 survey questions to the advanced students’ test score profiles (in the test score database) with all cases matched via the students’ university identification numbers.

On the question of abroad experience, each person was scored categorically (no/yes or 0/1) on three levels: abroad experience (in a country where the language is spoken natively); experience studying abroad; and experience living with a host family while abroad. Any individual who had studied abroad or lived with a host family while abroad necessarily received a 1 for having abroad experience. But for those who had been abroad, it was possible that they were not on a study-abroad program, nor did they stay with a host family.

Although there are three survey questions under the category of heritage speaker, we did not use them as individual indicators of heritage status in our
analysis. Instead, we took a conservative approach and combined them into a single variable (0/1) to indicate the status of being a non-heritage or heritage speaker. We did so because no survey question directly asked the students whether they were heritage speakers of the L2. Given the lack of a direct question, we defined the construct of heritage speaker in a broad sense; specifically, those who answered yes to any of the three questions. Non-heritage-status students were those who answered no to all three questions. Thus, we are fairly certain that our non-heritage group (group 0) contains only non-heritage learners. The heritage learners (group 1), however, represent a wide range of individuals on that construct.

The third category (i.e. use of authentic L2 resources) consisted of eight continuous questions on the survey, each indexing the frequency with which a particular type of L2 resource was used. We used a six-point Likert scale, and we asked students
to choose the number that best represented how often they used authentic L2 resources to assist their L2 learning outside of class (and not as homework): “1” to “6” respectively stood for “Never,” “Once a month or less,” “A few times a month,” “Weekly,” “A few times a week,” and “Daily.” We correlated the score from the eight questions. Given the strong correlation between online-video watching and TV and movie watching ($r = .76$), we collapsed those scores into a new video watching index that was, for each learner, the average of these two questions. As a result, we included in total seven types of authentic L2 resources in our analysis of this variable.

The last variable is motivation, for which there was only one survey question. That question asked about the students’ motivation in learning the target language. Again, students responded using a 6-point Likert scale, where “1” represented “not interested in learning the language,” and “6” represented “strongly interested in learning the language.” Table 9.2 contains the final set of variables that we included in our analysis: four categorical and eight continuous variables.

After matching the score profiles of the 147 advanced students with their survey data, we removed the cases that did not have any survey data available and thereby obtained the final dataset for this present project. In total, we had 142 advanced language learners with survey question data, 41 (29%) were advanced in speaking, 41 (29%) in listening, and 121 (86%) in reading. We further coded the 142 students into the following seven groups according to their advanced skill profile: those who were advanced in (1) speaking only ($N = 18$), (2) reading only ($N = 75$), (3) listening only ($N = 1$), (4) both speaking and reading ($N = 8$), (5) both speaking and listening ($N = 2$), (6) both reading and listening ($N = 24$), and (7) all three skills ($N = 14$).

**Data analysis** The major analysis that we carried out for this project was a cluster analysis. We used the TwoStep clustering algorithm provided in SPSS 22.0 to perform the cluster analysis because TwoStep is a better tool to deal with a mixture of categorical and continuous data, as compared with K-Means and hierarchical clustering (IBM SPSS, n.d.). Due to the exploratory nature of this analysis, we did not specify the number of clusters to be formed. Instead, we relied on the software program to test different solutions and choose the best one. We selected Schwarz’s
Bayesian Information Criterion (BIC) as the clustering criterion, but the results stayed the same when we switched to Akaike’s Information Criterion (AIC). Regarding the maximum number of clusters, we accepted the SPSS default value of 15, which means that the program may only pick the best-fitting solution among the solutions that had 15 or fewer clusters. Because of the co-existence of categorical and continuous variables, we applied Log-likelihood as the distance measure, rather than the more often-used Euclidean measure.

To minimize information loss, we added all of the learner variables from Table 9.2 in the analysis. But before doing that, we standardized the continuous variables. We did not opt for noise handling because we did not observe any outliers in the data. Because we expected each cluster to be mostly unique in terms of the learners’ characteristics, we assumed the distribution of skill profiles would vary by cluster. To test our hypothesis, we called for a skill-profile description for each cluster in the SPSS output based on the seven types of skill profile that we coded earlier (i.e., reading only, listening only, speaking and reading, speaking and listening, reading and listening, and all three skills). We were then able to run a comparison across clusters to see whether each cluster was featured by a somewhat distinct skill-profile distribution.

Three assumptions apply to TwoStep clustering when the distance measure of Log-likelihood is selected. First, continuous variables should follow a normal distribution; second, categorical variables should follow a multinomial distribution; and third, variables should be independent from each other (IBM SPSS, n.d.). Our data did not meet the first two assumptions, but we continued our analysis because the algorithm is relatively robust against violations of these assumptions (Norušis, 2011). For testing the third assumption, Mooi and Sarstedt (2011) suggested consulting the correlation matrix and eliminating or replacing highly correlated variables (i.e., \( r \geq .9 \)). The strongest correlation that we found in the correlation matrix was the one between online-video watching and TV and movie watching (\( r = .76 \)). Although it was below the threshold set by Mooi and Sarstedt (2011), as noted earlier, we decided to replace these two variables with a combined variable of video watching because, conceptually, we believe the two variables overlap to a great extent.

**Results and discussion**

*Frequency distributions* The frequency distribution of the seven skill profiles across the learners is shown in Table 9.3. Over half of the advanced learners were advanced in *reading only*. The second largest group was the *reading and listening* group, followed by the *speaking only* group, and the *all three skills* group. Those who were advanced in speaking and a receptive skill (either reading or listening) made up only 7% of the entire sample. Based on the summary statistics, we infer from our data at least two patterns in L2 skill development: (i) It was easier to achieve a high level of proficiency in reading than in the other skills; (ii) Advancedness in receptive skills (i.e., reading and listening) tended to go hand in hand, whereas advancedness in speaking, which is a productive skill, tended to be
independent of receptive skills unless a learner had a balanced development in all three skills.

Table 9.4 displays the frequency distribution of the four categorical survey variables. The majority of learners were non-heritage speakers ($N = 111$); heritage speakers only accounted for 22% of the entire sample ($N = 31$). Regarding the students’ abroad experience, more than two thirds had been abroad ($N = 98$), 52 had been on a study-abroad program, and 55 stayed with a host family during their time abroad. There was a large overlap between those who had studied abroad and those who had homestay experiences ($N = 40$). There was also a nearly equally large number of students who had been abroad for non-study purposes ($N = 46$).

As can be seen in Table 9.4, the relationship between heritage-speaker status and abroad experience is as one would expect. Among the non-heritage speakers, 63%
had been to a place where the L2 is spoken (N = 70). The number of those who had been abroad was lower in the heritage-speaker group (N = 28), but the percentage was much higher (i.e., 90%). Most non-heritage speakers went abroad for a study-abroad program, whereas heritage speakers went abroad for non-study reasons. Less than 50% of the heritage speakers indicated that they stayed with a host family during their time abroad; by contrast, homestays were higher among non-heritage speakers. In the non-heritage-speaker group, the homestays overlapped with study abroad.

Based on our examination of Table 9.4, four groups of learners emerged from the data: (i) heritage speakers who had been abroad, (ii) non-heritage speakers who stayed with a host family during their time studying abroad, (iii) non-heritage speakers who had been abroad for non-study purposes, and (iv) non-heritage speakers who had not been abroad. However, we need a formal statistical analysis to help us determine the robustness of this pattern, and we explain that analysis next.

Descriptive statistics: Continuous survey variables and score profiles Table 9.5 shows the descriptive statistics for the eight continuous variables. Overall, the advanced students were highly motivated to learn the L2, as we suspected would be the case based on the literature (Dörnyei, 2005, 2009). In fact, it appears there is a ceiling effect for motivation: Out of 6 points, the average motivation score of all advanced learners was 5.64, with a small standard deviation of 0.56.

By looking at the data in Table 9.5 across skill-profile subgroups, one can see that students who were advanced in speaking made more frequent use of authentic L2 resources outside of class than those who were advanced in the receptive skills only, especially those who were advanced in reading only. In other words, it may be that to be advanced in something other than reading in these language programs, extra L2-contact outside of class may be necessary, as foreshadowed by Blake (2013) and Bardovi-Harlig and Bastos (2011). The fact that reading is the skill that most learners have advanced proficiency in is likely a reflection of university-based curricula where literature is the focus. The paucity of speaking opportunities in these classes and the paucity of feedback (other than recasts) that one finds in advanced-level classes (Zyzik & Polio, 2008) in most universities may be the source of the low numbers of advanced speaking scores.

We think these data suggest that the language programs’ formal L2 reading instruction is adequate for students to obtain advanced reading without much help from other resources, perhaps because motivation appears to be high for all students. But advanced in reading plus another skill (speaking or listening) seems to require more engagement with the language outside of class, or may rely on study abroad, as suggested by Davidson (2015). One can see that high, authentic engagement with the language outside of class seems to be related to advanced speaking and reading and advancedness on all three skills. Advancedness in listening, however, seems elusive. The data underscore the rarity with which students in our database obtained advanced in listening, and how advanced skill tends to co-occur with reading or speaking.
### Table 9.5 Descriptive statistics for the continuous variables by skill profile.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Oral interaction</th>
<th>Podcasts &amp; news</th>
<th>Books &amp; novels</th>
<th>Email writing</th>
<th>Social media</th>
<th>Text chatting</th>
<th>Video watching</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Speaking (N=18)</td>
<td>4.88 (0.86)</td>
<td>2.71 (1.69)</td>
<td>3.06 (1.61)</td>
<td>2.47 (1.42)</td>
<td>3.59 (1.91)</td>
<td>3.59 (1.66)</td>
<td>3.44 (1.69)</td>
<td>5.53 (0.72)</td>
</tr>
<tr>
<td>2. Reading (N=75)</td>
<td>3.67 (1.59)</td>
<td>1.83 (1.06)</td>
<td>2.47 (1.26)</td>
<td>2.04 (1.05)</td>
<td>2.32 (1.67)</td>
<td>2.61 (1.52)</td>
<td>2.17 (0.96)</td>
<td>5.64 (0.56)</td>
</tr>
<tr>
<td>3. Listening (N=1)</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00</td>
<td>3.00</td>
<td>1.50</td>
<td>5.00</td>
</tr>
<tr>
<td>4. Speaking &amp; reading (N=8)</td>
<td>4.75 (1.67)</td>
<td>2.50 (1.60)</td>
<td>4.13 (1.36)</td>
<td>3.25 (1.17)</td>
<td>3.75 (1.58)</td>
<td>4.25 (1.58)</td>
<td>3.44 (1.88)</td>
<td>5.50 (0.76)</td>
</tr>
<tr>
<td>5. Speaking &amp; listening (N=2)</td>
<td>5.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>5.00</td>
<td>5.33</td>
<td>4.50</td>
<td>6.00</td>
</tr>
<tr>
<td>6. Reading &amp; listening (N=24)</td>
<td>3.42 (1.56)</td>
<td>2.75 (1.29)</td>
<td>2.83 (1.17)</td>
<td>2.46 (1.14)</td>
<td>2.67 (1.49)</td>
<td>3.25 (1.68)</td>
<td>2.90 (1.39)</td>
<td>5.71 (0.46)</td>
</tr>
<tr>
<td>7. Three skills (N=14)</td>
<td>4.64 (1.60)</td>
<td>2.86 (1.79)</td>
<td>3.71 (1.59)</td>
<td>3.79 (1.53)</td>
<td>3.29 (2.09)</td>
<td>3.38 (1.94)</td>
<td>3.29 (1.44)</td>
<td>5.71 (0.47)</td>
</tr>
<tr>
<td>Total (N=142)</td>
<td>3.94 (1.59)</td>
<td>2.26 (1.39)</td>
<td>2.84 (1.42)</td>
<td>2.45 (1.31)</td>
<td>2.76 (1.79)</td>
<td>3.06 (1.68)</td>
<td>2.67 (1.37)</td>
<td>5.64 (0.56)</td>
</tr>
</tbody>
</table>

Notes: Scores are means; in parentheses (when N>2) are standard deviations.
Cluster analysis When we ran the cluster analysis, the best solution in SPSS 22.0 was four clusters. The silhouette coefficient, an index of cluster quality, was .2, which means that the average within-cluster distance was slightly smaller than the smallest average between-cluster distance (see Norušis, 2011). The silhouette coefficient of .2 is not large enough to be considered a good solution, but it is within the acceptable range, thus we proceeded in using the cluster analysis to investigate individual difference profiles in advanced learners.

There were four clusters that appeared within the 132 cases in the final cluster solution (there were 10 learners that the program did not fit in any cluster; they can be considered cluster outliers). The four clusters are ordered by the computer program in terms of fit: The statistically tightest cluster is the first cluster, and the statistically loosest cluster is the last one. The first cluster was the smallest ($N=22$) and had 16.42% of the cases. The second cluster was the largest ($N=41$), with 31.60% of the 132 cases. The third ($N=32$) comprised 23.88% of the cases. And the final cluster ($N=38$) was 29.10% of the cases. The ratio of the size of the largest cluster to that of the smallest cluster was 1.86. Although we did not get four clusters of equal size, the ratio of sizes was tolerable in that the clustered cases were not unproportionally assigned to an extremely large cluster, and that the size of the smallest cluster was decent enough for us to draw meaningful inferences. The number of outliers was small and thus not concerning for the analysis, in that the number was less than half of the smallest cluster.

The importance of different predictors can be found in Figure 9.1. Video watching proved to be the most important predictor, which means that it made the greatest

![Figure 9.1 Bar graph for cluster predictor importance.](image-url)
contribution in differentiating among the different clusters. Second most important was the triadic variable of abroad/study-abroad/homestay experience. By contrast, interest in learning the L2 (motivation) had the lowest importance. This is most likely because this variable had little variation; all learners were mostly at ceiling (all advanced learners had high motivation for learning the L2).

Table 9.6 lists the learner characteristics of each cluster. Here we list the clusters by N size. The largest cluster was of non-heritage speakers who had been abroad for non-study purposes (Cluster 2). The next cluster was of non-heritage speakers who had no experience abroad (Cluster 4). The next cluster was of non-heritage speakers who had both study-abroad and homestay experience (Cluster 3). And finally, one cluster was of heritage speakers who had been abroad for non-study purposes (Cluster 1).

For Cluster 1, the heritage-speaker cluster, the average frequency of their use of the authentic L2 resources was exceptionally high. In the domains of video watching

### Table 9.6 Summary statistics for cluster analysis.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Categorical variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abroad experience</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Study-abroad experience</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Homestay experience</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Heritage speaker</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Continuous variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of L2 resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Video watching</strong></td>
<td>2.01 (0.68)</td>
<td>1.88 (0.64)</td>
<td>2.94 (1.08)</td>
<td>4.93 (0.94)</td>
</tr>
<tr>
<td>Social media</td>
<td>2.10 (1.26)</td>
<td>1.92 (1.46)</td>
<td>2.94 (1.48)</td>
<td>5.23 (1.07)</td>
</tr>
<tr>
<td>News &amp; podcasts</td>
<td>1.68 (0.85)</td>
<td>1.62 (0.67)</td>
<td>2.69 (1.28)</td>
<td>4.05 (1.65)</td>
</tr>
<tr>
<td>Email writing</td>
<td>2.00 (0.87)</td>
<td>1.67 (0.93)</td>
<td>3.12 (1.21)</td>
<td>3.68 (1.17)</td>
</tr>
<tr>
<td>Books &amp; novels</td>
<td>2.37 (1.22)</td>
<td>2.26 (1.14)</td>
<td>3.22 (1.36)</td>
<td>4.09 (1.27)</td>
</tr>
<tr>
<td>Text chatting</td>
<td>2.61 (1.48)</td>
<td>2.56 (1.54)</td>
<td>3.19 (1.51)</td>
<td>4.50 (1.68)</td>
</tr>
<tr>
<td>Oral interaction</td>
<td>3.51 (1.55)</td>
<td>3.72 (1.64)</td>
<td>3.94 (1.50)</td>
<td>5.32 (0.84)</td>
</tr>
<tr>
<td>Learning interest</td>
<td>5.63 (0.58)</td>
<td>5.49 (0.60)</td>
<td>5.75 (0.44)</td>
<td>5.64 (0.66)</td>
</tr>
</tbody>
</table>
and the use of social media, the differences in mean frequency scores between Cluster 1 and other clusters were over 2 points. To translate such differences into real life, the heritage-speaker students noted that they watched L2 videos and used L2 social media in the target language on a weekly basis, but their non-heritage-speaker peers indicated they used the two types of L2 resources only monthly or less. This could be evidence that heritage language learners have learning behaviors that are different from non-heritage students (as described by Kondo-Brown, 2005), and that they have a large networked community of speakers that support their learning efforts (Carreira, 2013; Kagan, 2012); our data may suggest that this network is beneficially digital. The data may suggest that the heritage learner’s digital social network and knowledge and use of authentic L2 resources helps him or her acquire the language to the advanced level.

Among the three clusters of non-heritage speakers (Clusters 3, 2, 1), the mean frequency scores on authentic L2 use from the study-abroad cluster (Cluster 3) was on average 1 point higher than from the other two clusters. For most types of authentic L2 resource use, students in Cluster 3 used them a few times a month, whereas Clusters 4 and 2 used them once a month or less. One may see this as a sustained benefit from studying abroad, especially if students, while studying abroad, are learning how to use authentic L2 resources, and they continue to use them post study abroad, as these data seem to suggest.

It should be noted that although Clusters 2 and 4 did not seem to access extra (outside of class) L2 input from TV, social media, news, or books, they had frequent face-to-face conversations with speakers of the L2, and also had high motivation for learning the L2. These positive features probably partially explain why students in Clusters 2 and 4 were able to achieve an advanced level of proficiency in the L2. Similar features were also found in Clusters 1 and 3. But, as we will show next, Clusters 2 and 4 students mostly achieved advanced in reading only.

Finally, we examined the skill profiles of the students in each cluster. Pie graphs in Figure 9.2 show how the seven different skill profiles are represented within each cluster. Within Cluster 1 (the heritage-speaker cluster, N = 22), we can see that most students were advanced in speaking only. Although those who were advanced in reading only constituted the second largest group in Cluster 1, we also see a decent proportion of students who were advanced in speaking and reading and all three skills. Clusters 2, 3, and 4 (the non-heritage clusters) were similar to some extent. They were all dominated by students who were advanced in reading only, with the second largest group being the learners who were advanced in reading and listening. Cluster 4 (non-heritage, no abroad experience, low use of authentic L2 resources) was most dominantly advanced in reading only.

Figure 9.3 illustrates the same clustered data from a different perspective. This time, pies represent skill profiles and slices represent clusters. We did not include the pie graphs for the listening only and the listening and speaking groups due to the small number of observations in those groups. Again, Cluster 1 (the heritage group) seems to be weighted across the profiles rather well, with the most advanced profiles on speaking (14 out of 21 within this cluster were advanced in speaking). Cluster 4 (non-heritage, non-abroad), by contrast, demonstrated advancedness
largely in the receptive skills (N = 35 out of 38), as indicated by its large share of reading only (28), listening only (1), and reading and listening (6) advanced-skill profiles. Across the three non-heritage clusters (Clusters 4, 3, 2), as experience abroad becomes a more important factor (from Cluster 4 with no abroad experience, to Cluster 2 with abroad experience but no study-abroad experience, and Cluster 3 with study-abroad experience and homestays), the use of authentic L2 resources concomitantly increases, and, as one can see, the mixture of advanced profiles (away from a predominately reading only profile) increases, until with Cluster 3, the advanced profile distribution is similar to that represented by the heritage learners in Cluster 1.
In summary, the TwoStep clustering algorithm demonstrated four clusters, each with high motivation for learning the language: (1) heritage-speakers who had been abroad and used authentic L2 resources at a high level of frequency (Cluster 1, N = 22); (2) non-heritage speakers who had been abroad but used the L2 resources at a mid-level frequency (Cluster 3, N = 32); (3) non-heritage speakers who had studied abroad but seldom used authentic L2 resources (Cluster 2, N = 41); and (4) non-heritage speakers who had never been abroad and seldom used authentic L2 resources (Cluster 4, N = 39). Regarding the skill profiles of the four clusters, students in the heritage-speaker cluster were more likely to be advanced speakers, whereas students in the other three clusters were more likely to be advanced readers. Students who had studied abroad and indicated that they used authentic L2 resources more frequently had a better chance of making balanced progress in all three skills. Students who had never been abroad and indicated that they infrequently used authentic L2 resources tended to achieve an advanced proficiency level in reading but not in the other skills. Thus we agree with Davidson (2015) that there are difficulties in constructing a comprehensive foreign language curricular model, and in particular, a model that promotes complete advanced-skill development. As we reviewed above, Davidson mentioned that the difficulties are more apparent in connection with less-commonly taught language programs, but we believe these difficulties arise even in programs such as Spanish when the upper-level classes have a strong focus on reading and less on the others skills (as found by Zyzik & Polio, 2008).

Conclusion

In our conclusion, we wanted to first go back to our definition of what it is to be an advanced user of a foreign or second language. As we summarized at the beginning of this chapter, an advanced language learner can use the language contextually in sophisticated ways (Ortega & Byrnes, 2008, p. 8), and his or her advanced status is related to “aspects of literacy, to diverse manifestations of cultural competence, choice among registers and multiple speech community repertoires, voice, and identity in cross-cultural communicative settings” (p. 8). Advanced language learners are multilingual and needed to take professional jobs that require advanced language competence (Aoki, 2013; Kennedy & Hansen, 2015; Moeller, 2013). But in our study, we found very few truly (robustly) advanced language learners. Only 14 out of 142 advanced learners in this study (that is, 10%) were actually advanced in all three skills (listening, speaking, and reading). Most advanced learners in our study were advanced in reading alone (N = 75, 53%). The attainment of an advanced learner status is even more elusive if we look at the population that took the tests in spring 2016. Out of 1,311 students who tested, only 14 (1%) obtained a rating of advanced across three skills. However, some students opted not to take all three tests, and another number of students (not counted within the 1,311) did not take any of the tests. Thus, there could be more within the population at large that are advanced across all
tests. But to us, it is troubling that so few are demonstrating, across-the-board advancedness.

Revisiting our definition of the advanced learner, we do not think that one who is advanced in reading only is necessarily one able to fulfill the full repertoire of what it means to be advanced. Can those who are advanced in reading only actually use the language contextually in sophisticated ways? Likewise, are the 21 (14%) advanced learners in our study who did not receive an advanced score in reading advanced in relation to “aspects of literacy,” and “to diverse manifestations of culture competence” (Ortega & Byrnes, 2008, p. 8). Probably not. They may also be lacking in “voice and identity in cross-cultural communicative settings,” as also defined by Ortega and Byrnes (p. 8). Thus, by assessing our language learners to measure advancedness, as recommended by Ortega and Byrnes and by Norris (2006), we have become aware that advancedness, in our students, is not an absolute skill as mostly defined in the literature. Our question is no longer how students progress from the plateau of intermediate-level language learners to advanced ones. Rather, the question for us is, how do language learners move from somewhat advanced (advanced in one or more skills but not all) to fully advanced in all skills?

At our university, with our data, we see proficiency at the advanced level is often fractionally obtained and mostly incomplete. In our study, only 42 (30%) of those with advanced skills would actually be allowed to work professionally (as a language teacher, for example) in the target language, because advanced skill for professional licensing is assessed and measured mostly through the skill of speaking alone. Advancedness in skills other than speaking is insufficient for professional work in most situations (see Aoki, 2013; Kennedy & Hansen, 2015; Moeller, 2013). On the other hand, with advanced in speaking but not in reading, an advanced learner would theoretically not meet most definitions of advancedness, as literacy appears to be key to theoretical advancedness (Ortega & Byrnes, 2008). Thus, in using standardized assessment to take a snapshot of advanced foreign language learners (to ensure that advanced programs are doing their job, as suggested by Norris, 2006), we find we are not satisfied. The programs appear to be able to create advanced readers, regardless of abroad experience, heritage status, and use of authentic L2 resources outside of class (as evidenced by Cluster 4), but other advanced skill sets appear elusive, especially listening and a combination of all three (listening, speaking, and reading). It seems that language programs need more engagement with the language outside of class (abroad experience, use of authentic L2 resources outside of class) to achieve higher or more robust advanced learners, ones who are able to use the language in sophisticated ways.

Two of our results intrigued us in particular. First, we found that out-of-class video usage was an important predictor of advanced clustering. And second, we found that heritage language learners tended to be the most frequent users of social media, videos, email, and text chatting in the target language outside of class time (and apart from homework), and that those who returned from studying abroad were the second most frequent users of such online tools. We also noted that these two clusters of students (Cluster 1, heritage learners with high social
Individual Differences in Advanced Proficiency

media use; and Cluster 3, non-heritage learners who studied abroad with home-stays and who use social media in the target language to a medium degree, and more so than Clusters 2 and 4 did) had the most complex advanced profile mixes (see Figure 9.3), with more individuals obtaining advanced status in skills other than just reading. We believe that these results point to the benefits of language engagement outside of class, as Blake (2013) stressed, and this important engagement includes digital media use. This type of language engagement is often not fostered from within the classroom as much as it could or should be. Rather, it may be fostered now through heritage connections or study-abroad experiences. It could be that during study abroad, language learners are making, as Spenader (2008) noted, important and lasting connections to the host-country nationals while abroad; we speculate that these connections may be digitally lasting in nature. We thus believe that we should encourage the language programs to teach students how to connect to speakers of the language through social media (and how to connect with speakers in person when possible) and how to find authentic (and routinely watch) video sources so that the language learners will have better chances of using and engaging with the language outside of class on a regular basis. Perhaps second to motivation (which was high for all), video use and social media use in the target language outside of class may indicate high engagement with the language and language community (it may be related to a particular kind of motivation for learning), and such engagement may be a necessary precursor to advanced communicative skills—in particular, speaking.

We believe the most important take-away from this study is that advanced language learning can happen in regular university language programs, and it does. We agree with Moeller (2013) that advanced is an achievable goal. The task for language educators is to make sure it happens more often and more frequently. A large task is to ensure that advancedness is across all skills, and that it is obtained by all learners regardless of their heritage status or abroad experience. Prior researchers have shown that advanced-level skill attainment is possible without study abroad. For example, Brown (2009) showed that with robust in-class programming of argumentation and debate, students on the cusp of being advanced in speaking can pass that threshold. We believe that another push toward advancedness without study abroad may be in the use of social media and the regular consumption of video and online text. We agree with Thorne and Reinhardt (2008), who noted that foreign language teachers need to “combine the best of the analytic traditions of schooling with the life experiences and future needs of today’s foreign language students” (p. 562). In particular, they advocated for the inclusion of digital literacies within the advanced curriculum to “provide a realia counterweight to the prescriptivist versions of grammar, style, and vocabulary in foreign language texts that typically are not based upon actual language use” (p. 562). We agree that advanced-level instructors may need to pick up on the instruction of social media and other online and video-based realia to ensure more robust advancedness in their students.

We would be remiss if we did not point out that our conclusions are based on an analysis of those with advanced proficiency in at least one skill. Without a
comparable analysis of those with intermediate-level proficiency, we cannot be sure that these findings are unique to our population. For example, it might be that some variables (e.g. video watching) are common to groups or individuals who do not reach advanced proficiency. But nevertheless, our analysis shows the benefits of testing for advanced proficiency (as suggested by Ortega & Byrnes, 2008, and by Norris, 2006). Programs need to do so to identify the advanced learners so that they can also learn from them, as we have. There are many areas that future research might be able to investigate, building on the current work. First, our definition of a “heritage” speaker was not as precise as one would like. Heritage speakers have numerous profiles, and a more nuanced approach might reveal other characteristics that lead to advanced proficiency. An interesting finding was the prediction of out-of-class video watching to an advanced profile. Because our data come from a large questionnaire, we were not able to delve deeper into this somewhat surprising finding. For example, was extensive video watching stemming from class assignments or was this completely voluntary and independent from coursework? And, of course, video watching itself covers a wide range of activities and timeframes. This, too deserves further investigation.

Finally, we need to mention the narrow scope of advanced proficiency taken in this chapter. We want to point this out because this narrow scope also appears in a large amount of the literature dealing with proficiency testing. The assessments that we conducted were focused on linguistic skills (reading, listening, speaking) and did not take into account other aspects of proficiency, such as cultural knowledge. Cultural knowledge, of course, is an important part of being a proficient speaker, but it is a component that is often ignored in standardized testing. We hope that the current research will lead to further studies into what it takes to be (and what it takes to become) an advanced speaker of a second or foreign language.

Acknowledgment

This research was funded by the National Security Education Program’s Language Proficiency Flagship Initiative (grant # 2340-MSU-7-PI-093-PO1) awarded to principal investigators Winke and Gass.

REFERENCES


The Prior Language Experience of Heritage Bilinguals

CRISTINA SANZ¹ AND JULIO TORRES²

¹ Georgetown University
² University of California, Irvine

Introduction

This chapter is mainly concerned with heritage language learners (HLLs); that is, classroom learners of a heritage language (HL), usually a minority language, that is one of the learners’ languages along with the societal majority language. As such, HLLs distinguish themselves from second language learners (L2Ls) in that they arrive at an instructed setting with prior experience of the language being taught. This distinction is important, as scholars have highlighted the need for different approaches to language teaching for HLLs (e.g. Beaudrie, Ducar, & Potowski, 2014; Kagan & Dillon, 2012; Torres, Pascual y Cabo, & Beusterien, 2017). The present chapter will provide the reader with important information on differences between HLLs and L2Ls. It will also identify the characteristics of advanced HL speakers’ linguistic knowledge, as well as pedagogical interventions and their differential effects on each group of learners. Finally, given HLLs’ bilingual experience, we will consider whether this experience enables them to benefit from an advantage in learning further languages (Lado, Bowden, Stafford, & Sanz, 2018; Sanz, 2000).

An important part of any introduction is a definition of key terms. While this is true for any field, it is especially important in bilingualism research, where even the term bilingualism is used to label so many different ways of being bilingual. So we are going to devote the following paragraphs to distinguish among L2Ls, HLLs, and balanced bilinguals. We use here the term balanced bilinguals to describe...
highly proficient and biliterate bilinguals, but we agree with Grosjean (2008) that true balanced bilingualism does not exist given the number of contexts in which bilinguals make use of one or two of their languages.

Everybody agrees bilinguals are individuals who can use two languages. Advanced L2 learners and HLLs are bilinguals. The disagreement comes with any adjective that we may attach to narrow down such broad definition. L2 learners, including advanced L2 learners, are referred to as emergent bilinguals, while HLLs have in the past been referred to as semilingual, and more recently, incomplete bilinguals. Both groups are usually compared to monolingual native speakers and tend to be characterized by what they cannot do; however, while emergent suggests growth in the right direction, semilingual and incomplete suggest lack of development, stagnation. Neither group is labeled balanced bilingual for several reasons. In the most strict sense, a balanced bilingual is always a simultaneous bilingual; that is, a person who learned two languages simultaneously from birth and maintained comparable proficiency and levels of literacy and use in both. In a less strict sense, a balanced bilingual is an individual with full and equal control over two languages; the definition implies equal use of both languages in terms of frequency and context (formal and informal), it also suggests comparable capacity and use of reading and writing skills. A balanced bilingual is not an oral bilingual or a receptive bilingual. By this definition, emergent bilinguals and HL speakers are not balanced bilinguals, as their languages differ notably in proficiency, frequency and context of use, and levels of literacy.

Furthermore, and unlike ‘HL’ or ‘L2,’ ‘balanced’ does not refer as much to the bilingual’s past or future language proficiency as to the current comparable status of two languages in the individual’s linguistic repertoire. Even though L2Ls and HLLs are bilinguals, they differ from each other in substantial aspects having to do with the past, the present, and the future proficiency along with the use of their two languages. The concept that distinguishes the two is that of native language. By definition, HLLs are in the classroom to (re)learn a first language (L1) while L2 learners, as their label suggests, start learning a foreign or second language once their first language is well established and are in the classroom to add a new language to their repertoire. L2 learners are often literate in their L1 and take full advantage of it when learning the L2; in fact, it is not uncommon in classroom contexts to find that exposure to written language is far superior to exposure to aural input. In contrast, HLLs are literate in the majority language, but are often illiterate or have limited literacy skills in their HL. Therefore, advanced L2Ls and HLLs differ in terms of their past and their present, but will both reach balanced bilingualism? Given appropriate social and educational circumstances, L2Ls will keep developing their L2 while maintaining their L1, and HLLs will maintain the majority language while further developing their HL. Following these trajectories, only possible in the appropriate social and educational circumstances, both groups have the potential to become balanced bilinguals. In fact, if those circumstances existed, HL learners’ past and present would be one characterized by full, balanced bilingualism.
Individual differences in prior language experience

In the field of second language acquisition (SLA), it is accepted that learners’ individual differences play a major role in language performance and outcomes (Dörnyei & Skehan, 2003; Gass & Mackey, 2012; Sanz, 2005). Prior language experience as an individual difference has been examined mainly in the acquisition of a third language (L3) (Bowden, Sanz, & Stafford, 2005). Theoretically, researchers working within this strand have been preoccupied with describing how bilinguals as experienced language learners approach the task of learning an additional language in comparison to their monolingual peers. For example, Sanz (2000) showed that biliterate Catalan-Spanish speakers showed greater gains in learning L3 English morphosyntax. Stafford, Bowden, and Sanz (2011) compared early and late bilinguals and found that late bilinguals retained more information of L3 Latin.

However, while HLLs certainly learn additional languages, recent research efforts have been devoted to examining how they (re)learn a childhood or HL that does not enjoy a majority status in their country of residence. As such, their use of the HL is typically restricted to their home and community in informal contexts while they use the majority language more extensively including in academic settings. Due to this prior language experience, the HL oftentimes becomes their weaker and less dominant language in adulthood, especially in cases where HL children have no access to early bilingual education. Overall, HL speakers’ grammars exhibit patterns of simplification and restructuring when compared to their monolingual native counterparts (Benmamoun, Montrul, & Polinsky, 2013; Montrul, 2016a; Silva-Corvalán, 1994). Researchers have argued that adult HL grammar outcomes can be due to underlying linguistic phenomena including: incomplete acquisition, attrition, transfer from the dominant language, the quantity and quality of the input, and/or that HL speakers’ path to acquisition just differs from that of monolingual speakers (see Rothman, Tsimpli, & Pascual y Cabo, 2016, for a recent discussion on these issues).

Therefore, an important question to consider is how HLLs’ prior language experience, as described above, interacts with pedagogical interventions to advance their language performance and development in comparison to L2Ls. A key difference between HLLs and L2Ls is that HL speakers oftentimes bring functional knowledge of the HL into the learning environment (Valdés, 2001). It is generally agreed that HL speakers acquired their L1 through a different internal process than their non-native counterparts because, as native speakers, heritage speakers are first exposed to the HL at home in the course of growing up, through natural interaction and without exposure to explicit discussions of how the language works. It is therefore an immersive, implicit context, very different from the formal classroom where L2 learners usually learn their L2. Even though it would be wrong to claim that learning in a naturalistic context implies learning in the absence of attention to form, and that a formal context precludes the possibility of learning implicitly, differences in context of acquisition may be responsible for differences in performance in these two groups. In fact, evidence exists that HL speakers perform
better in less explicit and aural experimental tasks in comparison to their L2 speaker peers (Bowles, 2011a; Montrul, Davidson, de la Fuente, & Foote, 2014). This prior language experience most likely bears consequences for how both groups of learners attend to, process, and store target language input in the classroom.

HL and L2 speakers also demonstrate parallel gaps and variability in their grammar systems, especially as highlighted by evidence in the area of morphosyntax. For example, Montrul et al. (2014) stated that “many of the non-native patterns displayed by heritage speakers resemble the grammatical patterns typical of adult L2 learners who are either in the process of learning the L2 or have ceased development (and fossilized)” (p. 119). A recent study by Swender, Martin, Rivera-Martínez, and Kagan (2014) looking at oral proficiency exam data from intermediate L2Ls and HLLs of Spanish and Russian, found a lack of structural control, limited vocabulary, and linguistic breakdown when producing higher-level grammar for both L2 and heritage speaker participant groups. Again, despite differences in learning contexts and age of acquisition, factors leading to non-targetlike production, such as transfer, can be similar for both L2Ls and HLLs. A final point in common between L2Ls and HLLs that differentiates them from monolingual native speakers of majority languages is that for both HL and L2 speakers, full development of linguistic abilities may not always be guaranteed (Montrul, 2008).

Brain research suggests that L2 language training under implicit conditions can lead to native-like neural activity at high proficiency levels (Morgan-Short, Steinhauer, Sanz, & Ullman, 2012), so the possibility is there. However, issues of identity and motivation may be stronger factors for HL speakers than for L2 learners, for whom the L2 is not so personal and conflicted.

Obviously, an important variable in L2 and HL development is proficiency. Thomas (1994) defined language proficiency as “a person’s overall competence and ability to perform in L2.” Hulstijn (2011) highlighted the importance of defining language proficiency, as it is a key component in addressing individual differences in language development for adult native and non-native speakers. In regards to HL bilinguals, scholars have argued for distinctions between language proficiency and language dominance while also recognizing their interrelatedness. Language dominance can refer to the speaker’s frequent use of a particular language in a number of settings and the easiness with which they access one language over the other, whereas proficiency is specifically characterized as the speaker’s linguistic knowledge (e.g. grammar, pronunciation, fluency) at a given point in time (Birdsong, 2014; Montrul, 2016a, 2016b; Silva-Corvalán, 2014). As a group, HL speakers are characterized by the heterogeneity of their proficiency: At one extreme one finds some HL speakers who are biliterate, balanced bilinguals, at the other HL speakers who can only understand but not speak the HL, with the entire gamut in between. Polinsky and Kagan (2007) discussed that the range of HL speakers’ proficiency levels should be viewed according to how they resemble or deviate from a monolingual baseline.

A number of factors can contribute to this group’s linguistic heterogeneity, including age of arrival and, consequently, length of residence in their new country if they are first generation; attitudes to the HL and motivation to maintain it or
relearn it; and access to input and opportunities for interaction, especially among second- and third-generation HL speakers. A key factor is the role of formal education, as some HL speakers may have been schooled in their L1 prior to arrival in the new country, or may have attended evening or weekend classes in their new neighborhoods, or even, if they are very lucky, may have attended a dual-language immersion program that supports the maintenance and development of the HL. Additionally, those who only speak the language at home and are not formally educated in it tend to use a variety that deviates from monolingual varieties, as the majority language takes over. This is due to how the majority language takes over, restricting access to input and interaction in the HL.

In line with the theme of this handbook on the advanced language learner, we will now focus on reviewing 14 studies, mostly for Spanish, that measured proficiency and reported findings for advanced proficient HL speakers. We searched for articles in the Linguistic and Language Behavior Abstracts database, in the Heritage Language Journal, and in Google Scholar. Our selection of these 14 studies was based on the following criteria: (i) the studies were published in peer-reviewed journals, (ii) the studies used some criterion or instrument to measure proficiency in the HL, and (iii) the studies reported specific findings for the advanced proficient group. The studies are summarized in Table 10.1.

Proficiency measurements

As seen in Table 10.1, the measurements to gauge participants’ proficiency are many, and include the American Council on the Teaching of Foreign Languages (ACTFL) Oral Proficiency Interview (OPI) or its computer simulation, online diagnostic tests, years of formal classroom experience, cloze passages, grammar proficiency tests, and a Spanish proficiency test (in Montrul, 2004), which is a short version of the DELE, Diploma de español lengua extranjera (Diploma of Spanish as a Foreign Language). Notably, and in line with Thomas’s (1994) comprehensive definition of proficiency, these measurements are all global, rather than focused on discrete grammatical functions. In addition, a few of the studies (Alarcón, 2011; Montrul & Bowles, 2009; Pascual y Cabo, 2016) relied on questionnaires to elicit participants’ linguistic background information (e.g. frequency of language use, education, age of onset of the L2) and self-rated proficiency in listening, reading, speaking, and writing. Montrul and Bowles (2009) and Pascual y Cabo (2016) provided both self-rating and scores from the proficiency test.

In his review of Spanish HL writing research, Torres (2016) recommended researchers to rely on more than one proficiency measurement in an attempt to triangulate proficiency measures and better understand the interplay between proficiency levels and heritage participants’ language performance and outcomes. For example, in Estremera and Torres (2014), HL participants’ self-ratings appeared to strongly correlate only with their performance on more explicit and written experimental tasks, suggesting that the utility of self-ratings may depend on the experimental task. Future studies ought to also incorporate some measurement of
Table 10.1 Previous research examining linguistic structure among advanced-level heritage learners.

<table>
<thead>
<tr>
<th>Study</th>
<th>Language</th>
<th>Proficiency measure</th>
<th>Summary of findings</th>
</tr>
</thead>
</table>
• handled formal situations  
• produced abstract vocabulary  
• produced more null than overt subjects  
• omitted the preposition *a* with animate direct objects |
| 2. Montrul (2004) | Spanish | DELE (Montrul’s adapted version) | • demonstrated robust knowledge of unaccusativity in Spanish |
| 3. Montrul (2005) | Spanish | DELE (Montrul’s adapted version) | • omitted dative object marker *a* with *gustar* (like) verbs |
| 4. Montrul & Bowles (2009) | Spanish | DELE (Montrul’s adapted version) | • accepted ungrammatical double object constructions |
| 5. Alarcón (2010) | Spanish | Online diagnostic test | • exhibited positive attitudes and high motivation toward Spanish  
• reported significant interactions in Spanish  
• produced a high percentage of clitics  
• discriminated between grammatical and ungrammatical items on a written acceptability judgment test |
| 6. Montrul (2010) | Spanish | DELE (Montrul’s adapted version) | • demonstrated and produced knowledge of Spanish gender agreement |
| 7. Alarcón (2011) | Spanish | Grammar proficiency test Participants’ self-ratings | • demonstrated and produced knowledge of Spanish gender agreement |
| 8. Montrul & Perpiñán (2011) | Spanish | DELE (Montrul’s adapted version) | • lacked superior recognition of tense-aspect and mood morphology  
• made semantic distinctions between preterit and imperfect  
• lagged behind advanced L2 learners in distinguishing between indicative and subjunctive mood |
| 9. Davidson & Lekic (2012) | Russian | ACTFL OPI Test of Russian as a Foreign Language (TORFL) | • increased their overall proficiency level as a result of Russian Overseas Flagship program |
| 10. Mikhaylova (2012) | Russian | Cloze passage | • fell along a continuum in interpreting Russian aspect |
| 11. Albirini (2014) | Arabic | Years of classroom experience | • produced more accurate sentential negation in standard Arabic |
| 12. Leal Méndez, Rothman, & Slabakova (2014) | Spanish | DELE (Montrul’s adapted version) | • demonstrated knowledge of clitic right dislocation in discourse appropriate situations |
| 13. Swender, Martin, Rivera-Martinez, & Kagan (2014) | Russian and Spanish | ACTFL OPIc | • narrated with description in the HL  
• handled formal situations  
• produced more abstract vocabulary  
• demonstrated an agentive interpretation of Spanish *gustar* (like) verbs |
| 14. Pascual y Cabo (2016) | Spanish | DELE (Montrul’s adapted version) |
language dominance, like the promising Bilingual Language Profile (BLP) instrument (Birdsong, Gertken, & Amengual, 2012). This measurement proved to be productive in a recent study by Amengual (2016), who found that Spanish-dominant bilinguals exhibited more Spanish tap-trill features like those found in most monolingual Spanish speakers.

Given that the most common measurements are the ACTFL OPI and Montrul’s (2004) adaptation of the DELE proficiency test, we briefly describe them here. The ACTFL OPI consists of an interactive interview; the individual’s oral performance is then compared against a set of benchmarks (ACTFL, 2012). Variations of the OPI have also been developed (SOPI, a recorded simulated oral proficiency instrument, and OPIc or Oral Proficiency Interview – Computerized) for which the live interview session has been eliminated for practical reasons. According to ACTFL (2012), speakers at the advanced level can communicate effectively in informal and formal scenarios and on concrete topics related to events of interest to the individual; they can narrate in different time frames and provide organized and connected paragraph-length responses. In contrast, intermediate-level speakers are limited to daily life topics and sentence-level speech. A number of L2 studies and foreign language programs have utilized the OPI to measure participants’ proficiency levels (e.g. Bowden, 2016; Chambless, 2012; Hernández, 2010; Iwashita, 2006; Malone & Montee, 2010). Additionally, as Table 10.1 demonstrates, three HL studies (Davidson & Lekic, 2012; Kagan & Friedman, 2003; Swender et al., 2014) adapted the OPI, SOPI, and OPIc as proficiency measures driven by an interest to address educational issues pertaining to course placement issues and program outcomes (e.g. Russian Overseas Flagship program).

Interestingly though, despite its popularity, none of the studies that included advanced proficient HL bilinguals relied on the OPI. In the case of Spanish HL studies, researchers preferred instead to use Montrul’s adapted version of the DELE, possibly for convenience and for replicability purposes. DELE is the Spanish official language accreditation diploma. Montrul’s short version consists of a fill-in-the blank vocabulary section and a cloze passage section that add up to a total of 50 points. The scores assign participants to three possible proficiency categories: low proficiency (scores 0–29), intermediate proficiency (30–39), and advanced proficiency (40–50). An online copy of this DELE is available through a research digital repository hosted in the National Heritage Language Resource Center’s website (http://nhlrc.ucla.edu/nhlrc/category/data). Many Spanish HL formal and applied studies have implemented this proficiency measure (e.g. Bowles, 2011b; Cuza & Frank, 2011; Henshaw, 2015; Leal Méndez, Rothman, & Slabakova, 2014; Montrul, 2004, 2005, 2010; Montrul & Bowles, 2008; Montrul & Perpiñán, 2011; Pascual y Cabo, 2016; Pascual y Cabo & Gómez Soler, 2015; Torres, 2013; Torres & Sanz, 2015).

However, researchers have also raised concerns regarding this assessment tool as it taps into metalinguistic knowledge and so puts HL speakers at a disadvantage (Carreira & Potowski, 2011). Montrul (2016b) defended it, arguing that studies have yielded correlations with participants’ performance on linguistic tasks, and cites two studies (see Montrul, Foote, & Perpiñán, 2008; Montrul & Ionin, 2012).
that yielded a Cronbach’s alpha value of .80. More recently, Chomón Zamora’s unpublished paper (2015) found a significant correlation between Montrul’s assessment test and Ortega’s Elicited Imitation Task (EIT; Ortega, 2000); developed for L2 Spanish learners, the EIT has received empirical support as a valid internal and external instrument for measuring L2 proficiency (see Bowden, 2016). Unlike the adaptation of the DELE, the EIT measures both comprehension and production, as it requires participants to listen to a number of sentences increasing in syntactic complexity, and to repeat the sentences after a short delay; for this reason, the EIT is a more appropriate test to measure HLLs’ proficiency, especially for individuals with low literacy levels. An additional advantage of the EIT is its availability in different languages (e.g. Mandarin, Korean), thus facilitating cross-linguistic comparisons within and across studies.

### Advanced proficient heritage speakers

The studies summarized in Table 10.1 either addressed the advanced HL speakers’ educational needs or consisted of formal accounts of HL speakers’ grammars. Regarding educational needs, Alarcón (2010) reported on how advanced HL students of Spanish exhibited a high degree of motivation and positive attitudes toward studying the HL while mostly expressing a need to develop academic writing skills in the HL. Educational experiences were beneficial in pushing the linguistic development of advanced Russian and Spanish HLLs, as evidenced in Davidson and Lekic (2012) as well as Swender and colleagues (2014).

Scholars in formal approaches to HL’s morphosyntax agree that certain linguistic structures remain unchanged across proficiency levels. For example, Montrul (2010) and Leal Méndez et al. (2014) found that HL speakers at all proficiency levels demonstrated robust knowledge of Spanish clitics and of unaccusativity (Montrul, 2005). Conversely, other linguistic structures are vulnerable for HL speakers across the proficiency spectrum, as is the case with dative object marking (DOM) in Spanish as demonstrated in Montrul (2004) and Montrul and Bowles (2009). In Pascual y Cabo (2016), advanced and intermediate proficient Spanish HL participants assigned agentive interpretation to gustar (like) verbs, unlike monolingual native controls. Comparing advanced proficient HL speakers and L2Ls, studies of Russian and Spanish verbal aspect identified comparable linguistic behavior in both groups of bilinguals (Mikhaylova, 2012; Montrul & Perpiñán, 2011). Also, while advanced Arabic HL speakers demonstrated superior knowledge of sentential negation, their performance paralleled that of L2 participants (Albirini, 2014). However, advanced HL speakers exhibited suboptimal performance on the Spanish indicative and subjunctive mood distinction in comparison to their advanced L2 peers (Montrul & Perpiñán, 2011).

In sum, the overall findings from studies on the advanced HL speaker suggest that examining distinctly advanced proficiency is necessary to further understand its complexity. Yet, based on the available evidence, a clear picture does not emerge as advanced HL speakers behave along a continuum: While they can exhibit
linguistic patterns that align with those found for low/intermediate HL speakers, L2Ls, and monolingual native speakers, they also show patterns that depart from those of the other groups. Often, the specific patterns differ not so much in their nature per se as in the degree or the frequency to which they perform certain linguistic functions that may deviate from monolingual norms. Arguably, type of linguistic structure can be the underlying force for such variable performance among advanced HL speakers. A detailed analysis of the linguistic structures in these studies (e.g. whether or not certain structures are more difficult due to their position in the interface of syntax and external cognitive domains) is beyond the scope of this chapter; we refer readers to other work (e.g. Montrul & Polinsky, 2011; Sorace, 2011, 2012).

**Pedagogical interventions**

As argued before, prior language experience is very different for L2Ls and HLLs, because it corresponds to their respective contexts of acquisition, i.e. the classroom versus the home and the community. As a consequence, L2Ls present superior written performance and score higher in traditional classroom tests that measure explicit, conscious knowledge of the language, while HLLs show better oral production, richer lexical and syntactic variety, and higher cultural knowledge. A new awareness of these differences has pushed for separate classes for heritage speakers in recent years, as researchers have argued that curricula designed for L2 learners oftentimes are inappropriate for heritage language learners (e.g. Carreira & Kagan, 2011; Peyton, Carreira, Wang, & Wiley, 2008; Valdés, Fishman, Chávez, & Pérez, 2006). New courses, textbooks, and even programs have been developed for this population, especially for Spanish HLLs (e.g. Carreira & Geoffrion-Vinci, 2007; Kagan, Akishina, & Robin, 2002; Potowski, 2017; Roca, 2012), although there are proposals for accommodating different needs within the same classroom too, with separate objectives frequently concentrating on literacy skills and explicit grammar for HLLs (e.g. Colombi & Alarcón, 1997; Valdés et al., 2006). The situation currently is that most institutions do not offer especially tailored classes for heritage speakers, and instead opt for mixed classrooms of HL and L2 learners (Lynch, 2008).

While a number of resources have been developed over the years to meet the linguistic needs of HLLs in the classroom, little research exists as to how HLLs respond to pedagogical interventions. Recently, Polinsky (2016) argued rightly for more communication between language scientists and educators to maximize the ways in which HLLs can benefit from instruction, although the discussion was limited to HLLs’ linguistic competence and performance. We concur that findings from the field of formal linguistics provide vital information regarding linguistic structures that are vulnerable during heritage bilingual acquisition. And this information can certainly guide practitioners in deciding the areas of language that need special attention in an instructed setting. Yet, limiting the focus to only target linguistic structures that may or may not constitute the grammars of HLLs
is insufficient to ascertain the most optimal conditions that promote HL performance and development. For the past decades, research in the field of instructed L2 acquisition, which has contributed mostly to L2/foreign language teaching, has found that linguistic structures themselves are only one part of the puzzle and that other variables are also at stake. Certainly, much discussion has been devoted to investigating the properties of linguistic structures (e.g. saliency, redundancy, semantic opacity) that can hamper the form-meaning connections that needed to be made in an L2 (DeKeyser, 2005). But research has also documented that L2Ls exhibit internal constraints when processing L2 input. For example, Ellis and Sagarrá (2010) found that Chinese L2 learners of Latin have difficulty processing inflectional morphology when other cues (e.g. adverbial) are simultaneously available in the input. The learners’ early experience with the L1, the authors conclude, blocks the acquisition of new L2 cues as a result of depleted attentional resources that impede the processing of less salient cues, such as inflectional morphemes. Similarly, it is not far-fetched to suggest that the overwhelmingly superior exposure to the majority language, the HLs’ L2, blocks the acquisition of features of the L1 and/or promotes its attrition.

In addition to learners’ internal constraints, the role of external manipulations of the learning environment as well as learners’ individual differences can differentially affect the rate of L2 development. An external condition that has been widely studied in the field of instructed L2 acquisition is the use of tasks (Long, 1985). Essentially, tasks can be defined as language activities that imitate real-world events that require learners to solve a problem, to make decisions, and/or to provide justifications for behavior, among others. Different variations of tasks exist as design features can be manipulated, such as whether a task requires the participation of two individuals or whether a pair of learners share the same information to solve a problem. Long (2015) summarizes empirical studies, pointing to how differences in task-type can lead to a variation of L2 performance and outcomes. For example, open tasks that do not require one correct answer lead to more complex and accurate language whereas closed tasks (one correct answer) promote more negotiation of meaning and greater fluency.

Additionally, learner factors (e.g. motivation, language aptitude) can shape how learners benefit from pedagogical interventions. Serafini and Sanz (2016) investigated whether learners’ individual differences in working memory capacity could explain variability in L2 development at different proficiency levels. One key finding was that the role of working memory capacity facilitated learning at lower stages of L2 development. That is, beginning but not advanced learners with higher ability to process and store information demonstrated greater gains in L2 development. In sum, we have provided a brief snapshot of major factors that altogether conspire to facilitate or to prevent adult language learning, including linguistic complexity, internal processing constraints, external conditions, and individual differences.

Given these dynamics in instructed language learning, much research is needed to understand how these factors interact with one another in adult learners who decide to (re)learn their HL in an instructed setting. Importantly, we must consider
additional difficulties for HLLs, such as the conflict between the nature of their prior language experience, acquired almost exclusively in a naturalistic setting for most HL speakers, and the nature of the new, formal context. That is, the HLLs functional knowledge of the HL, which is mostly characterized as implicit, aural, and variable, has the potential to clash with the explicit, written, and standardized language that characterizes the language classroom. What pedagogical techniques better fit these learners? There are also the affective and psychosocial factors associated with the fact that HLLs are speakers of a contact, low-status variety of the language they are relearning, which is itself a minoritized language with low status in the larger society. Which variety to choose for use in the classroom? Finally, how do (if at all) pedagogical interventions override an entrenched linguistic system resulting from a prolonged experience of living with two grammars?

A few studies have probed into how HLLs respond to a number of pedagogical interventions originally designed for L2Ls. Preliminary evidence from laboratory and classroom studies suggest that HLLs do benefit from a number of pedagogical interventions, including explicitness and positioning of corrective feedback and processing instruction in promoting the development of Korean past tense, Spanish dative object marker, and the Spanish present and past subjunctive (Henshaw, 2015; Kang, 2010; Montrul & Bowles, 2008; Potowski, Jegerski, & Morgan-Short, 2009; Torres, 2018). One emerging observation, however, is that while HLLs demonstrated learning gains, these were not as high as for L2Ls; also, HLs’ performance differed significantly from that of monolingual native speakers, though one of the HLL groups in Torres (2018) gained almost as much as the L2L group. More research is needed to examine whether this observation holds true, and to find explanations for these differences, one being the assessment methods, which favor L2Ls by probing explicit knowledge.

Another strand investigates language performance in task-based interactions, mostly of HL-L2 dyads. These studies have reported that HL learners overall offer more assistance to their L2 peers, especially in the area of vocabulary (Blake & Zyzik, 2003; Bowles, Adams, & Toth, 2014), and that manipulating task modality can alter HL-L2 classroom interactions, with the L2 group providing their HL partners with more feedback on written tasks (Bowles, 2011b). Gass and Lewis’s (2007) study on Italian L2 and HL learners’ perception of interactional feedback showed that HLLs were less aware than L2Ls of morphosyntactic feedback. A similar observation is found in Torres (2018); his HLL participants reported being more focused on the content of the pedagogical task than on the feedback (i.e. recast).

Overall but with notable exceptions, the literature above suggests that HLLs respond differently to a number of pedagogical interventions, which in turn seemingly affects their language performance and development. However, the HLLs early bilingual experience does not always confer an advantage in (re)learning the HL in an instructed context. These observations remain inconclusive and as such are an indication that investigating the role of prior language experience is a
worthwhile endeavor. The outcomes of such research will enable researchers and practitioners alike to make sound decisions regarding instruction and curriculum design. As argued, though, in Torres et al. (2017), educators ought to have agency as to what findings are relevant to their local context, and whether results are replicated in classroom environments across different age groups, a task that potentially can be accomplished through classroom action research.

None of the studies reviewed above reported on distinct findings for advanced proficient HLLs, as every study that has assessed participants’ proficiency is limited to intermediate learners (Bowles, 2011b; Henshaw, 2015; Kang, 2010; Montrul & Bowles, 2008; Torres, 2018). Blake and Zyzik (2003) made a distinction among their HL participants between those raised mostly in a Spanish-speaking country and those raised in the United States. They also included self-reports in their design, but the authors made no claims about their groups due to their low number of participants. Kang (2010) implemented a picture description test and ACTFL benchmarks to measure her Korean HL participants’ proficiency. The remaining studies (i.e. Bowles, 2011b; Henshaw, 2015; Montrul & Bowles, 2008; and Torres, 2018) all utilized Montrul’s (2004) adaptation of the DELE. Henshaw reported including a few marginally advanced proficient learners (score = 40), but does not provide any separate analyses for them. Knowing that even advanced HL speakers lack certain linguistic structures and that their grammars are not impervious to restructuring patterns (e.g. Montrul & Bowles, 2009; Pascual y Cabo, 2016) demands that we all take steps to solve the current dearth of pedagogical research by focusing on identifying the best approaches for this population.

**HL speakers as additional language (Ln) learners**

Millions of students in foreign language classes around the world, including HL speakers, decide to learn yet a third language. However, we do not know much about how their prior language experience helps them in learning an L3 or Ln. How much experience is necessary? What is the role of biliteracy? And does any potential advantage depend on how the L3 is taught? Several studies have shown that bilinguals outperform monolinguals at learning an additional language (Brohy, 2001; Cenoz & Valencia, 1994; Sanz, 2000). Moreover, higher control of two languages in bilingual children and adolescents has been found to be beneficial for overall proficiency in an L3 (Lasagabaster, 2000; Muñoz, 2000; Sagasta, 2003) and for specific aspects of the L3 such as vocabulary (Keshavarz & Astaneh, 2004) and grammar (Sanz, 2007). Studies showing benefits for bilingualism for L3 learning included biliterate bilinguals, as they were conducted in bilingual areas with bilingual educational systems, like Quebec and Catalonia; in fact, other studies showed the importance of socioeducational conditions. Thomas (1988) found that college students raised in a bilingual home with classroom language experience had an advantage over those without an overall college L3 learning experience. Stafford, Sanz, and Bowden’s (2010) comparison of early and late bilinguals
learning Latin as an L3 through exposure to highly explicit instructional conditions showed that late bilinguals retained more of what they had learned than early bilinguals. These results were unexpected, as the authors saw early L2 learners as experienced language learners and users; they initially assumed that an earlier age of arrival to the L2 community meant a higher level of bilingualism. The authors argued instead that early exposure to the L2 (English) in an English-dominant community likely led to severely reduced opportunities for input and practice and consequently to arrested development and to attrition characteristic of HL speakers. This conclusion shows that the socioeducational situation needs to be factored in on any study on the advantages of prior language experience for L3 learning, rather than the assumption be made that all bilinguals, including HL speakers, are ‘expert’ language learners capable, for example, of using a wider variety of strategies and techniques and doing it more often than more ‘novice’ language learners (McLaughlin, 1995; McLaughlin & Nayak 1989). In a nutshell then, it is not enough to be bilingual, a native bilingual, or an advanced language learner to reap the benefits of prior language experience for further language learning. Literacy plays a key role (Sanz, 2007).

Conclusion and future research

We have discussed a number of issues that support the notion that HL speakers’ prior language experience places them in a rather unique position along a continuum of bilingualism; in fact, HL speakers’ grammars and experience support the need to view bilingualism as a continuous (not categorical) variable in communities of research and practice. Differences in proficiency levels among HL bilinguals are as important as in L2Ls; however, close examination of advanced proficiency in this chapter has been challenging given the limited number of empirical studies that included and/or reported on advanced HL speakers. This research limitation is no different from what L2 research is now finally overcoming, it is certainly exacerbated for HLL research as it is a field in its infancy. The scarce data do imply that advanced HL grammars fluctuate between alignments with and deviations from monolingual and other bilingual grammars; it remains unclear whether fluctuations are marked by linguistic structures or by proficiency measures. What is certain is that we need additional studies that provide a clear picture of HL learning across the entire range of proficiency levels to understand the interplay between teaching, learning, and cognition in this population.

Based on the literature reviewed here, we suggest a number of ideas for a research agenda on HL development. Future studies ought to provide, at minimum, one objective and one subjective measure of proficiency; they should report how the study’s findings are modulated by proficiency level, as we have seen in a few of the previously mentioned studies. Objective and subjective measures could also be combined to provide an overall continuous score for proficiency. Ortega’s (2000) Elicited Imitation Task promises to facilitate the necessary cross-linguistic
comparisons, as it is available in different languages. Also, researchers should include language dominance as an independent variable, given the distinction between language proficiency and language dominance; Birdsong and colleagues’ (2012) Bilingual Language Profile test can be a useful tool to accomplish this goal.

Regarding studies with pedagogical interventions, we echo Bowles’s (in press) research agenda and its proposal to implement a number of methodological approaches, including descriptive, experimental, quantitative, and qualitative. Crucially, the inclusion of online research techniques (e.g. eye-tracking, verbalizations) can elucidate further how HLLs process input to obtain a nuanced analysis of the effects of instruction that can explain observed learning gains or lack thereof. Additionally, this research agenda should be extended to include HL bilinguals’ experience in learning additional languages, and the role of literacy in the equation. Lastly, future research areas should consider re-exposure to the HL in naturalistic settings where the HL is the majority language, as is the case of HL speakers studying abroad and immigrant returnees. Overall, gathering this information is paramount to formulate sound theoretical claims about the interaction between HL bilinguals’ prior language experience, sociocognitive differences, and learning context in a way that can explain current proficiency and predict future development with an eye to the optimization of pedagogical decisions.

REFERENCES


Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition:


Meeting the Demands of Globalization: One Goal of ISLA Research

PAUL A. MALOVRH AND NINA MORENO
University of South Carolina

Introduction

Unlike many of the chapters in the current volume, the present one considers advanced proficiency from the perspective of instructed learning contexts and discusses development toward advanced stages by addressing basic language instruction and its implementation in university-level curricula from the bottom up. The professional demand for advanced proficiency in foreign languages is evident in the increasing requirements for professional licensure in various fields as well as the recent surge in language curricula targeting higher levels of oral proficiency as an outcome (see Brown & Brown, 2017; Hyltenstam, 2017, for an overview). As students are increasingly required to demonstrate advanced-level proficiency in order to obtain credentials in various fields, foreign language programs are in the position to explain to assessment coordinators, department chairs, and college deans why their success rate is low, particularly in regard to oral proficiency. Swender’s (2003) report is a good point of departure for language specialists to explain the general lack of success among adult second language (L2) learners. Her study revealed that less than 50% of undergraduate students, all of whom began study of their respective foreign language in their first year of university study, were able to reach the advanced stage of oral proficiency in a four-year period. To researchers in second language acquisition (SLA), this is no surprise. The fact that L2 acquisition among adult learners is characterized as having a low success rate has been one of the main findings of SLA research over
Advanced Proficiency and Performance: Multiple Dimensions and Contexts

the past 40 years (VanPatten & Williams, 2015), as has the fact that reaching advanced levels of proficiency involves many factors, such as aptitude, prior experience with language, motivation, and age of onset, to name just a few (Doughty, 2011). Yet, to language program administrators and/or university officials who are not specialists in applied linguistics or SLA, such facts are of little consolation when they affect a variety of assessment metrics, such as the amount of time needed to graduate, drop-out rates, and program enrollment. Similarly, parents and students who are paying the tuition bill are incredulous over the disconnect between an impressive grade point average in a foreign language and a low ability to communicate in it. A logical, yet often erroneous, conclusion is that someone did not do their job, or the system is failing.

The problem is not limited to academia; it is also a societal one, observable in the marketplace and current geopolitical phenomena (Brecht, Rivers, Robinson, & Davidson, 2017; Long, 2017), and academe remains slow to respond. The Modern Language Association (MLA) tried to initiate a change in academic culture, beginning with its 2007 Report, “Foreign Languages and Higher Education: New Structures for a Changing World,” in which an official call was put out for a restructuring of language programs in the United States in order to address the growing demand for language users. The MLA 2012 Presidential Address followed up on the call by encouraging language programs to rely more heavily on a relationship between theory and practice in order to develop curricula that would more effectively exploit natural learning mechanisms in the classroom; such a relationship would clearly require that experts in SLA and/or applied linguistics be a part of cultivating such curricula. Yet, as VanPatten’s (2015) White Paper showed, the presence of applied linguists and/or SLA experts in language departments is minimal. Not surprising, then, was a January 2017 issue of Inside Higher Ed, which reported a recent survey that found that out of the 134 professors and administrators who responded, more than half had read the MLA 2007 report, while only 39% had actually made efforts to make program changes (Redden, 2017).

Despite the slow change in academic culture, change is nonetheless beginning to take shape, and is apparent in various recent publications in the field of instructed second language acquisition (ISLA), a subfield of SLA, such as Leow and Cerezo’s (2016) call for a reconceptualization of ISLA’s vision, Loewen’s (2015) Introduction to Instructed Second Language Acquisition, Long’s (2017) proposed redefinitions of ISLA, and VanPatten’s (2017) question, “What is the goal of ISLA? [given the unlikelihood of L2 learners reaching native-like status]” (p. 53). Such research makes clear the relevance of ISLA in language programs’ pursuit to keep up with the changes imposed by globalization, and sets the stage for scholars in the field to better understand the potential and limitations of formal instruction. The impetus for the present chapter is the need in higher education to more actively address the social needs for advanced-level language use by effectively implementing curricula fostering the development of communicative language skills. Indeed, there have been successful programs aimed at facilitating higher proficiency levels, such as Georgetown University’s German Program, the Foreign Service Institute’s School of Language Studies, and the Language Flagship
Meeting the Demands of Globalization: One Goal of ISLA Research

Program (see Jackson, 2017, for reviews of such programs). The present chapter, however, is focused on the sort of change needed at any given institution of higher education where change is expected, regardless of its status as a research institution. The chapter is not a treatise on how to become an advanced speaker, nor is it a recipe for creating the perfect pedagogical approach. Rather, it aims to respond to VanPatten’s question by demonstrating one particular goal of ISLA research and its implications for language pedagogy and curricular design. The chapter begins with an overview of the demands for advanced language proficiency in private and public sectors, then discusses current trends in ISLA research, within a cognitive processing framework of SLA, involving the identification and implementation of best practices in teaching. It then provides an overview of one example of curricular restructuring, in the form of course design, aimed at facilitating more successful proficiency development.

Societal demands for advanced proficiency and academia’s response

The need for highly proficient foreign language users is a global reality. From the perspective of the United States, recent data from the MLA show that 80.38% of the population over the age of five speak English at home, while 19.62% speak a language other than English. In addition, the US Census Bureau’s 2014 American Community Survey reported that the number of Spanish-speaking people rose from 28 million in 2000 to 37 million in 2010, revealing an increase of 32% in just 10 years. Foreign Language (FL) teachers in more than 20 states of the United States need to demonstrate advanced-level oral proficiency in an American Council on the Teaching of Foreign Languages (ACTFL) Oral Proficiency Interview (OPI) in order to receive professional licensure. Government bureaus and agencies have specific language requirements and even provide incentive programs to attract people with superior FL abilities. The Central Intelligence Agency (CIA), for example, created a Language Hiring Bonus Program, in which job applicants who possess “superior” skills in certain foreign languages may receive up to $35,000 as a hiring bonus. The FBI and the State Department both require that job applicants for positions involving the use of foreign language demonstrate “general professional proficiency” by receiving a rating of “3” on the Defense Language Proficiency Test (DLPT). In the private sectors, descriptions of language ability are dispersed throughout job announcements. The US Department of Labor uses terminology such as bilingual, near-native, and fluent, to describe growing trends in the job market involving interpreters and translators of foreign languages.

Evidence of the growing importance of language proficiency abounds in an international context, as well. A recent index by the British Council, for instance, placed Spanish as the number one language to learn for the future welfare of industry and commerce in the United Kingdom. The Australian Department of Defence provides an allowance, “to encourage and assist members to maintain proficiency ... for the performance of linguistic duties.”
Such professional expectations have affected university language programs greatly; in addition to preparing students to achieve academic outcomes, they also need to prepare them to meet professional expectations. As Brecht et al. (2017) note, the demands for higher levels of foreign language proficiency extended from predominately public sector domains to the private sector starting in the late 1990s. Language-dependent industry is currently growing 8–10% per year (approximately three to four times faster than the US economy), and the sort of work requiring FL use now includes skilled trades and technicians. Brecht et al. (2017) conclude that the United States is no longer competing with other industrialized nations when it comes to producing qualified FL users, and the result is seen in recent trends to seek out new methods of addressing the development of proficiency as well as how to assess advanced-level performance. Human Resources departments in the private sector, for example, are now utilizing metrics such as 360-degree feedback, in which performance is evaluated according to additional measures such as peer review and native-speaker feedback, with particular attention given to how a language user’s abilities could affect target-language clients (McAloon, 2017). Despite such development, Malovrh (forthcoming) notes a disconnect between such socioeconomic data and the culture of university language programs; course curricula for a language major usually include one business language course, yet fail to recognize the demand for language skills in the private sector with more breadth and depth. More collaboration between language departments and business schools could easily overcome such a disconnect.

The phenomena described above, all of which can be traced to the changes induced by globalization, are the driving force for restructuring in language program design. Given the limitations of formal instruction cited in Swender (2003), the relatively low success rate of adult L2 acquisition, and the variety of factors influencing whether or not an individual may or may not reach higher levels of proficiency, academia has traditionally relied on study-abroad opportunities as the main solution to facilitating development at the undergraduate level. Yet, the study-abroad option favors those students who can afford such options, or whose life circumstances allow them, while disadvantaging the rest, insofar as it relates to linguistic development. Classroom-only students are less likely to analyze and develop complex morphological structures in their oral production than study-abroad learners (Malovrh, 2014; Malovrh & Lee, 2013). What alternatives are there for those who are not able to study abroad?

Most research investigating the need for advanced language development has addressed the problem from the perspective of language program curricula. That is, many believe it is necessary to integrate language courses with content courses. Zyzik and Polio (2008), for example, found that advanced-level content courses (i.e. courses that are considered advanced because they are beyond the basic-language program sequence, but actually have no correlation with proficiency level) consisted of student–teacher interaction in which the expert language user provided little, if any, corrective feedback regarding the target language to his/her students. Other researchers have addressed the problem through the development
of specific courses. Weyers (2010), for example, reported the results of an intervention in which students spent a semester in a course designed to develop metacognitive speaking strategies, with the aim of increasing learners’ likelihood of meeting oral proficiency standards of the National Council for Accreditation of Teacher Education (NCATE). He found that 71% of his participants who completed the course progressed at least one sublevel on the ACTFL oral proficiency scale, and that one-third of those who did not obtain the advanced-low rating before the course were able to do so after completing it. Could development be expedited even further through a curriculum-wide institution of such metacognitive training? Such questions need to be addressed empirically through careful design in an ISLA context, and results of such research need to be incorporated in the curricular design of language programs if advanced proficiency remains a curricular goal. As Byrnes (Chapter 8, Advanced-Level Grammatical Development in Instructed SLA) concluded in her call for future research, there remains a need for the following:

Serious engagement with (and, therefore, taking seriously) key dynamics of language programs, where the existence and substance of an educational philosophy regarding long-term learning, as manifested in differentiated curricular thinking and pedagogies in, are critical if educationally useable findings are to result. (p. 149)

ISLA in a foreign language curriculum: Goals for fostering change

One approach to addressing the need for academia to produce more advanced-level language users involves the field of ISLA. We return to VanPatten’s (2017) question, referring to the goal of ISLA, given the unlikelihood of adult L2 learners ever reaching native-like status in the target language. A response we offer, and central to the present chapter, is that one goal of ISLA is to examine the potential and limitations of instruction and its effect on the learning process, to the extent that formal instruction is better able to produce actors of foreign language use in a globalized world. While we know that there will always be variability regarding the successful outcomes among learners, recent research also informs us that more needs to be done in the context of ISLA to better understand the effects of learning contexts and processes. Specifically, Leow and Cerezo (2016) called for a reconceptualization of ISLA’s vision in terms of (i) a new definition of learning context, (ii) the use of online processing measures to better understand internal learning processes, and (iii) a better understanding of how to contextualize learning context and process in language curricula. We posit that curricular change aiming to address advanced proficiency, particularly as it relates to advanced oral proficiency, needs to be addressed from the ground up, beginning in basic-language instruction.

While SLA addresses general theoretical tenets underlying L1 and L2 acquisition / learning processes that can include naturalistic settings, ISLA’s research focuses
on the L2 development that occurs in a formal learning setting. Loewen (2015) provided a comprehensive definition of ISLA:

a theoretically and empirically based field of academic inquiry that aims to understand how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one’s own. (p.2)

The two main underlying constructs in Loewen’s definition involve the formal (instructional) conditions under which learning is intended to take place and how the learning process in such conditions may be manipulated through intervention. We proceed with a brief overview of ISLA research in terms of goals for instituting curricular change according to learning context, learning process, and their implementation in language curricula.

**Goal #1: (Re)define learning context**

Curriculum designers and instructors are tasked with creating class time to provide the students with input and opportunities for practiced output, and to provide appropriate corrective feedback (DeKeyser & Prieto Botana, 2013). Traditional notions of classroom instruction make such a task particularly daunting, given the limited amount of class time available in most university courses. Yet, the formal setting is not limited to the physical locus of a foreign language classroom, since the environment is not critical to the impact of the interventions by the instructor or the instructional materials (Loewen, 2015). The contexts of ISLA research include, in addition to the more traditional face-to-face classes, hybrid/blended and 100% online class formats, where, following Allen and Seaman’s (2015) distinction, a traditional course does not deliver any content online, while a hybrid/blended course does so in 30–79% of its content. The extant literature on hybrid models, while still rather scarce, has shown that these models can enhance not only students’ L2 development, but also their engagement with the content area, peer collaboration, and accountability, even in more advanced-level courses (cf. Kraemer, 2008).

These different contexts lend themselves well not only to the use of technology as the means to deliver input using experimental materials via computer-assisted language learning (CALL) tasks, but also to flipped instructional designs in which input, as well as tasks promoting the analysis and internalization of input, may be delivered to the students at home via computer, which will in turn create more class time for productive language use and corrective feedback. In other words, the use of technology should not only be measured in terms of *quantity* in the classroom, but also in terms of *quality*, i.e. what the role of technology is in the classroom. Based on Cerezo’s (2010) work, Cerezo, Moreno, and Leow (2016) presented a classification of CALL technologies that describes *tutors* as the technologies in a classroom with which some of our L2 textbooks are working now in what we have come to know as ‘computer-graded’ homework assignments. They refer to a
pedagogical tasks in which learners’ responses are graded as ‘right’ or ‘wrong’ when matched against a database of acceptable answers. But technology can also be used as a medium when it allows learners the “agency to perform some kind of operation, including information retrieval, processing, production, or dissemination” (Cerezo, Moreno, & Leow, 2016, p. 245).

More and more publishers are incorporating elements in their online ancillary materials that enable students to use technology as a medium—a step forward from claiming to be ‘hybrid’ simply for offering the more traditional paper-and-pen workbook in an online format. Even in fully online classes, students are now able to interact with each other synchronously and asynchronously. The resulting effects, in terms of L2 development and the creation of a cohesive class community, remain to be seen and are still a work in progress, but now we at least have a selection of materials that justify the use of technology in more principled terms.

Goal #2: Institutionalize the relationship between theory and practice

What is most remarkable about ISLA, however, and what distinguishes it from SLA, is that the former can act as a connecting bridge between research and practice, an issue that R. Ellis (2010) and other scholars have come back to address in the past few years. What, then, is the role of research in the L2 classroom? Hatch (1978) warned us of the mismatch between the questions researchers pose for themselves and the questions that interest instructors, and made the point that perhaps “our research is not misguided but our application may well be” (p. 137). Others (e.g., Ellis, 2010; Ortega, 2011) have seen the opportunity for the rise of a synergistic relationship in which the information is not delivered from the top down (from researchers to teachers) but rather flows horizontally in a reciprocal and symbiotic relationship in which each field brings their expertise to the classroom. In other words, it is not up to the researchers to mandate or prescribe to teachers how to teach, but they can offer knowledge gathered from the studies conducted, which then, when selected critically, holds the potential to strengthen teachers’ effectiveness. We propose that the same be done with the new or newer realm of technology as administrative demands are steering foreign language programs into more technology-friendly curricula. The field of Task-Based Language Teaching (TBLT) has already adopted Doughty and Long’s (2003) 10 methodological principles (MPs) (developed further in Long, 2009; see also Long, 2014). The critical next steps would be to ensure that programs that have embraced a change in delivery (i) do so armed with the knowledge of how such changes can happen while maintaining sound ISLA tenets, and (ii) consistently and continuously train new and seasoned instructors as well as graduate teaching assistants in successfully extrapolating these 10 MPs into the new learning environment. Current trends in research do not always make it to graduate teaching assistants’ methodology courses (Moreno, 2014), and so it should be a priority to provide relevant findings in research to novice instructors in order to make positive changes in the classroom. Doing so requires that we find innovative ways to deliver such a
message, as Baralt and Morcillo Gómez (2017) have done for online TBLT training, to cite a recent example.

**Goal #3: Incorporating models of (I)SLA in processing-based pedagogy**

There is an abundance of variables that have been investigated in ISLA as well as methods to measure their impact in the L2 development process. Depending on the theoretical framework that led the studies, some variables and some methodological tools lend themselves better than others to answer research questions that also had pedagogical repercussions. Leow (2015) and Leow and Cerezo (2016) have identified the various theoretical underpinnings of ISLA and matched them with the cognitive processes and variables that were addressed in each. Here, we begin by revisiting the model of acquisition that follows the learning path from input to output.

The model was first presented in the late 1980s (Gass, 1988, p. 200); in it the portion of ambient speech that a learner picked up became apperceived input. A part of this input, in turn, was further processed (comprehended input), and this then became the learner’s intake. Intake was integrated into the learner’s interlanguage, and could then be produced by him/her as output.

In later years, Gass (1997; Gass & Selinker, 2008) updated the model of second language acquisition to depict the “dynamic and interactive nature of acquisition” (Gass & Selinker, 2008, p. 491). While the five main components of the original model were kept, the updated model portrayed how universals played a role in the process and included learners’ personality and affect. A more simplified model has since been used (VanPatten, 1996), which presents the basic processes in a linear mode, even though, as VanPatten (2004) and Leow (2015) explain, it should not be misinterpreted as a claim that L2 acquisition take place in a linear manner while lacking interaction between these processes and products. In VanPatten’s model, we see that learners have the opportunity to process input as they pay attention to part of the input. The subset of the input that is attended to, stored in working memory, and further processed, is called intake. Some of the intake may not make it to any further stages of processing. The language that does get processed further, during the restructuring phase, will be incorporated into the learner’s L2 system and stored so that s/he can access it and use it as output. Output represents a learner’s knowledge of the L2, even though we now know that all of his/her knowledge is not always manifested in the output.

Leow (2015, p. 17) proposed a more fine-grained model in which he added the notions of products and processes. In his model of the L2 learning process, Leow addresses or revisits several cognitive processes mentioned in the previous models, based on the premise that at least minimal attention to, or noticing of, the target structure is needed if any learning is to take place. It becomes apparent from the earlier models up to the most recent one that best practices in the classroom should focus on how instructional intervention can activate learners’ cognitive processes that will foster attention on the target structure. Furthermore, instructors and
curriculum designers need to be aware of processing-oriented instruction based on models of language processing, and emphasize the process of converting input to (target-like) intake. We propose that a flipped course design is amenable to providing sufficient time for students to convert input to intake (before arriving to the classroom), and to practice using (speaking and writing) in the target language during class, when they are able to receive immediate corrective feedback.

**Goal #4: Situating noticing in the instructional design**

The notion of attention is not new in the fields of SLA and ISLA. Based on what had been found in the field of cognitive psychology, McLaughlin, Rossman, and McLeod (1983) posited that humans were limited-capacity processors of information, and as such the field of SLA was to find ways to incorporate that notion in its research and in L2 pedagogy. Not long after, Schmidt and Frota’s (1986) account of Schmidt’s experience learning Portuguese was the beginning of what would become known as the noticing hypothesis (Schmidt, 1990, 1995, 2001). It posits that L2 acquisition does not occur without at least the learner’s awareness at the level of noticing, which would entail the “conscious registration” of a structure. The higher level of awareness, at the level of understanding, involves the learner’s recognition of an underlying rule to the structure (Schmidt, 1995). Robinson (1995) and Tomlin and Villa (1994) presented similar models. Robinson added the notion of storage in the short-term memory to his definition of noticing, while Tomlin and Villa claimed that awareness was not necessary for learning to take place but that attention (categorized in three different levels) was.

According to Loewen (2015), noticing is “one of the most important theoretical supports for focus on form” (p. 60). The role of attending to negative feedback is at the core of Long’s (1996) Interaction Hypothesis. He and others (see, for example, Doughty, 2003) have underscored the importance of when that focus occurred, claiming that the focus on form has to occur in a meaning-focused environment to be effective. Their claims strengthened the belief that immediate feedback on an ill-produced utterance would be more effective than when delivered with delay. Twenty years later, it appears that still much of ISLA’s research agenda on focus on form revolves around two issues: (i) What are the best methods to bring about noticing? and, (ii) Does noticing truly result in L2 development? (Loewen, 2015, p. 61). Some of the pedagogical questions we may ask ourselves regarding the first issue are: How explicit should attention to form be? If our curriculum is communicative and learner-centered, it cannot be overly explicit lest it interfere with communicative interactions in class; at the same time, it cannot be so implicit that it not be noticed, and thus lend itself to ambiguity, as is the case with teachers’ most favored feedback technique—recasts (e.g. Oliver & Mackey, 2003; Philp, 2003).

This brings up the issue of attending to form during a communicative task in which meaning is key. In that vein, VanPatten’s (1994, 2004) processing instruction (PI) approach addressed the issue of noticing as well, although it is not at the core of his approach or his input processing theory. In VanPatten (1990), he examined the simultaneous allocation of attention to both form and meaning when processing
aural input, and his study was the beginning of a long series of partial and conceptual replications that have looked at attention to form and meaning using different modalities and languages. Some made design changes to the original study in order to increase the replication’s internal validity (cf. Greenslade, Bouden, & Sanz, 1999; Leow, Hsieh, & Moreno, 2008; Morgan-Short, Heil, Botero-Moriaty, & Ebert, 2012; Wong, 2001). No conclusive results could be reached, in part due to a lack of uniformity in research design. To remedy that, a multi-site replication (Morgan-Short et al., 2017) was conducted, in which both aural and written modalities were employed, and the protocol at each site strictly controlled so that clearer results could be drawn. Preliminary results from the multi-site study indicate that attention to form did not interfere with comprehension among participants who scored low in the comprehension test (which was the case for the population of this multi-site study); these results contradict those obtained in the original VanPatten (1990) study, and corroborate Leow et al.’s (2008) and Morgan-Short et al.’s (2012) replications, keeping in mind the limitations of the experimental materials of the multi-site replication, which include a low reliability index for the reading / listening comprehension tests (cf. Morgan-Short et al. 2017 for complete explanation and more details on results).

What does all of this mean for an L2 classroom? Bringing attention to the target structure is not as clear-cut as we may think. If all that was needed was to make the target items more salient, then simple input enhancement would be sufficient, yet we know that this technique does not always bring about the desired results (Leow, Egi, Nuevo, & Tsai, 2003). It has also been shown in the literature that explicit feedback does not always work, at least not among beginner-level learners (Hsieh, 2008; Moreno, 2007; Morgan-Short & Sanz, 2004), and there is still uncertainty about whether detracting attention from form by focusing on meaning has a significant impact on comprehension in either modality (aural and written). It suggests that converting input to intake is not only about registering the target item but about how it is processed further, or rather, how deeply it is processed. In the next section, we consider depth of processing.

**Goal #5: Controlling for depth of processing in ISLA**

As the previous section showed, the opportunity to notice the target structure can be facilitated in many ways by the instructor or the instructional materials in the classroom. Whether students take this opportunity and to what extent is not always as easy to determine, however.

In the 1990s and early 2000s, the first attempts of ISLA research to direct learners’ attention to L2 data and to measure it consisted of creating simplifications (e.g. Leow, 1993) or of comparing whether the modality of the input had an impact (e.g. Leow, 1995; Wong, 2001). Results did not seem to support simplifications in the L2 input as a way to help with L2 development. In terms of modality, while Wong’s results seemed to suggest that modality could affect learners’ input processing, Leow (1995), a replication of Leow (1993), found no significant results between the written modality of the 1993 study and the 1995 study.
In Schmidt’s (1990, and elsewhere) noticing hypothesis, noticing the L2 data allows the learner to be aware of the L2. This awareness can occur at two levels: at the level of noticing, and at the level of understanding. The former is a simple registration of the L2 form while the latter requires that the learner form a hypothesis of the rule. Tomlin and Villa, on the other hand, argue that learning can take place without awareness as long as attention has been drawn to the L2. In their model, attention can occur at three levels: alertness, orientation, and detection. Alertness only signals readiness to attend to the L2 data and orientation is the stage at which attention is directed toward the target structure; it is only at detection that input can be converted into intake. The strand of ISLA interested in the cognitive processes involved in the input→intake stage has operationalized attention and awareness via concurrent online elicitation measures such as online verbal reports (also known as think-aloud protocols; TAs), eye-tracking devices, or reaction times; other researchers have also used offline retrospective measures such as offline oral and stimulated recall (SR) protocols. None of these is without problems or critiques. While concurrent procedures such as TAs have been criticized for reactivity, that is, interfering (positively or even negatively) with the L2 processing as learners reinforce the data when they are verbalizing their thought process, offline retrospective procedures such as SRs are said to have a veridicality problem since what learners claim to have been thinking during the experimental task may or may not be what they were actually thinking. It should be noted that several researchers (e.g. Bowles, 2010; Leow & Morgan-Short, 2004) have found no evidence of reactivity in L2 learners’ development; moreover, more studies that use TAs as their data collection method have begun to include a control group that is not required to produce TAs, so as to keep the risk of reactivity in check.

A more accurate account and measure of how input is being processed was needed—one that would allow to unobtrusively determine how deeply engaged the learner is with the input s/he is processing. **Depth of processing** is a notion that has already received considerable attention in cognitive psychology while being inducted in SLA as well. It is defined as follows:

**Depth of processing** is the relative amount of cognitive effort, level of analysis, and elaboration of intake, together with the usage of prior knowledge, hypothesis testing, and rule formation employed in decoding and encoding some grammatical or lexical item in the input. (Leow, 2015, p. 204)

What it means for L2 learning and teaching is that, for instance, providing learners with explicit rules to follow, as opposed to having them read a text and hypothesize about a grammatical rule on their own, would be less effective, because the former requires less depth of processing than the latter. The notion that the deeper the level of processing, the more learning takes place also presents a challenge in how we view common teaching practices such as the PPP (present-practice-produce) design. This popular and widespread practice has instructors present input and explicit instruction of grammar; learners then practice it in class, before they produce something—a deductive approach to learning. There are other more inductive
approaches (e.g. PACE model; Adair-Hauck & Donato, 2002) that follow a more implicit approach, in which instructors aid the students in figuring out the underlying rule of the input they have been presented. These approaches demand more cognitive effort, a greater level of analysis, and the use of prior knowledge from students as they form a hypothesis and make form-meaning connections. This approach has also been extrapolated to the online learning context with positive effects on learning processes and outcomes (Cerezo, Caras, & Leow, 2016). From a pedagogical perspective, the challenge is to find a balance, creating a learning path that includes tasks demanding greater depths of processing while also providing enough opportunities for practice and sufficient effective corrective feedback.

In the next section (The Carolinas), we will present how we piloted a flipped/hybrid model to do just that and found positive, albeit not significant, effects of this format on beginner learners’ L2 development in all four skills. Such challenges necessitate a rethinking of instructional context, as it relates to location and delivery, which may include technology.

**Goal #6: Control for judicious use of technology**

Leow and Cerezo (2016) clearly state that one must first understand the cognitive processes used by L2 learners before instructional interventions are selected, lest we fall into the trap of putting the cart before the horse. Cerezo, Moreno, and Leow (2016) present a comprehensive account of how current CALL and computer-mediated communication (CMC) tools can abide by ISLA’s theoretical tenets and how, if carefully designed, CALL technologies can follow Long’s (2014) 10 methodological principles for task-based language teaching (TBLT).

Cerezo, Caras, & Leow (2016) present two types of online tasks: less and more controlled tasks. Among the less controlled environments, which need the involvement of a human interlocutor, it has been found, for instance, that CMC helps learners attend to corrective feedback (Baralt, 2013) or that a TBLT-based language program that had students playing a video game and watching authentic videos increased the learners’ Spanish proficiency (González-Lloret & Nielson, 2015). The number of studies that have utilized more controlled activities (cf. Cerezo, 2012, for a comprehensive review) where the link between the tasks and L2 development has been clearly established has increased. The tasks that learners engaged in required them to complete meaningful drills through which they recognized or produced the target structure (e.g. Bowles, 2008; Hsieh, 2008; Leow, 2015; Moreno, 2007). What makes these tasks effective are three characteristics, according to Cerezo, Caras, and Leow (2016): (i) they adhere to task-essentialness, a trait in a task that requires that learners pay attention to the target structure in order to successfully complete the task (Loschky & Bley-Vroman, 1993); (ii) they offer feedback that allows learners to formulate hypotheses about the target structure; and (iii) they are self-paced and hold learners accountable for their learning progress (Cerezo, Caras, & Leow, 2016, p. 253).

Even though online materials in the realm of ISLA have come a long way since the first “drill and kill” e-tutors, and we now have a wide range of online materials...
available, it is still not always easy to trace the learners’ L2 development back to the use of technology directly (Cerezo, Moreno, & Leow, 2016). In order to ensure internal validity in studies that have looked at more controlled tasks, depth of processing has begun to be measured as a way to operationalize for attention on the target structure. Operationalizing attention and processing of the target structures via levels of depth of processing have provided a more detailed picture of what differentiates students who retain new information better from those who do not. It has also offered the opportunity to revisit studies where non-significance between groups could not be explicated. Hsieh (2008) compared a group that received a form of face-to-face instruction (Computer FTF) and a group that received computer-assisted instruction (CAI). Hsieh found that the CFTF group had reported more instances of awareness of the target structure at the level of understanding, yet no significant differences between the two groups could be found. When Hsieh, Moreno, and Leow (2016) reexamined the data by coding the collected TA protocols through the filter of depth of processing, a more detailed picture was painted. Unlike the two levels of awareness—noticing and understanding—the levels of depth of processing do not only gauge the learner’s attention to the structure, but also look for other manifestations of deeper cognitive effort, where, it must be underscored, the highest levels of depth of processing can mean coming up with the wrong hypothesis or rule for the target structure. Hsieh et al. (2016) found that even though the CFTF group reported more instances of awareness at the level of understanding, which used to be equated with more learning, the same group also exhibited fewer instances of higher levels of depth of processing. The impact of this study then lies not only in what the results were in regards to computerized versus face-to-face instruction, but, just as importantly, in regards to how the field of ISLA and ISLA research had been measuring learners’ internal processing of new structures up until then.

Technology is now more accessible but its availability should not blur the focus of our lens when deciding what we want to teach via technology and how we want to measure L2 development. Technology in the classroom should be used as long and only insofar as it allows us to apply sound ISLA tenets; when using technology as a research tool, it should lend itself to tap into learners’ internal cognitive processes involved in attending to and retaining new information. Depth of processing, as an innovative framework, can be used along with effective introspective data collection measures, such as TA protocols, to determine how cognitively demanding a task really is. Higher levels of depth of processing should also be our aim in the online and in-class tasks we design for our learners as that is what learners will store in their long-term memory.

**One example of proficiency-based curricular change:**

**“The Carolina Project”**

Moreno and Malovrh (2016) reported the results of a pilot study aimed to measure the effects of a flipped and blended basic language course design compared to a traditional face-to-face course design. As part of a larger initiative to restructure
language curricula, they called the pilot, “The Carolina Project.” With the considerations of what ISLA research tells us and what the best pedagogical implications that can be drawn are, and taking into account that more foreign language programs are not only encouraged to jump on the technology bandwagon, but are required to do so in order to satisfy administrative demands, the project instituted the ISLA goals outlined above.

The endeavor was not only an executive exercise of adhering to administrative demands to maximize classroom space, shorten face-to-face instruction time, and increase enrollments by offering student-friendly scheduling via hybrid and even online courses, but was also to make changes in a foreign language program that would be theoretically sound and pedagogically advantageous, in order to implement a proficiency-based language curriculum. The program changes began with the pilot, in which three different sections of the lowest tier of the basic Spanish course were compared. Since then, all basic (for true and false beginners) courses have adopted the hybrid/flipped format carefully designed to follow ISLA-friendly practices. In this section, we summarize the results of Moreno and Malovrh (2016).

The pilot study was framed in Leow’s (2015) framework of processing and used a blended and flipped course design to implement the goals outlined above. By using technology as a medium, a flipped course design makes it possible to provide L2 learners with input, as well as the appropriate tasks requiring higher levels of cognitive effort and active engagement with the target language, thus increasing the depth of processing, as input is converted to intake. By situating the initial learning process in a computerized delivery format for the student to complete before coming to class, they were allowed more time to internalize language, more opportunities for guided induction and hypothesis testing, and more time for linguistic analysis (Cerezo, Moreno, & Leow, 2016). Thus, more classroom time was reserved exclusively for practiced output and corrective feedback. The following hypotheses guided our study:

1. Great depth of processing will be achieved by using homework as a learning tool before classroom work begins;
2. More class time will be used for productive language use and corrective feedback;
3. Greater uniformity among large multi-section courses will be achieved by having an instructional design grounded in a specific cognitive framework of learning and centralized in one computerized delivery system;
4. Additional class time will be created for learning other content, such as culture and literary analysis; and,
5. The blended and flipped course design will allow for a reconceptualization of learning outcomes in terms of proficiency-based language development in the four skills of reading, writing, speaking, and listening.

All participants in the study completed pre- and post-tests measuring listening, reading, writing, and speaking skills, and were divided into three groups. Over
the course of one semester, learners enrolled in two experimental groups, using a
flipped-blended course design, met three days per week for face-to-face instruction,
and received input and activities requiring the analysis of input before class, as
computerized homework. A third (control) group met four days per week in class,
and received homework in the form of practice activities, following a traditional
course design of present-practice-produce. Multiple analyses of variance
(MANOVs) and post hoc tests revealed that all three groups were homogeneous
at the pre-test stage, in all four communicative skills. After 13 weeks of instruction,
post-test analyses showed that one of the experimental groups outperformed the
control group in speaking and writing, and the other experimental group outper-
formed the control group in writing. Moreno and Malovrh concluded that the
increase in practice of productive language during face-to-face class time, which
resulted from the delivery of input and input-analysis activities online before
class, led to more practice of language production along with more corrective
feedback, which would explain the greater development of productive skills across
the two experimental groups. Furthermore, they hypothesized that the flipped-
blended design engendered a greater depth of processing, in terms of linguistic
analysis and internalization, than the traditional design, since learners were
required to work with input at home and to hypothesize about its appropriate use,
before coming to class. Analyses, which will either refute or corroborate that hypo-
thesis, are ongoing.

Conclusions and future research

The present chapter posited that one goal of ISLA research is to meet the demands
of globalization on advanced-level foreign language proficiency. To that end, it
provided a brief review of social and economic phenomena affecting the market-
place and the demand for foreign language use, and then provided an overview of
current challenges of ISLA research. The chapter asserted six specific goals for
foreign language programs to consider as they attempt to restructure language-
learning curricula with the aim of achieving higher proficiency-based assessment
metrics. Among them are the need for a reconceptualization of what learning con-
text is, the need for FL courses to be directly related to models of language
processing established in SLA research, the need for more institutionalization of
the relationship between theory and practice, the need for an understanding of
noticing and depth of processing as constructs in the learning process and the
ability to implement them in input-related learning tasks, and the need to use tech-
nology judiciously, in order to establish a realistic and successful learning context.
The chapter concluded with an example of a recent attempt at program-wide
restructuring, in which instructional designs were partially controlled using a
blended-flipped design, and in which change in proficiency was measured
according to the four skills.

Much remains to be understood with regard to the effect implementation of
proficiency-based change in FL curricula. We assert that future research in the field of
ISLA needs to focus on effects of instructional design on language processing using online measures, in accordance with recent research, such as Leow and Cerezo (2016). How does learning context change, if at all, the way L2 learners process linguistic input? And how does such change, if any, affect retention of L2 knowledge? Future research should also examine the effect of teacher characteristics on the L2 learning process. How can teacher training, for example, alter or inhibit the positive effect of proficiency-based curricula? Finally, we believe the online context of language learning is still in its infancy, and that much more needs to be understood regarding its potential and limitations, in particular as they relate to practiced output. Can L2 learners practice, and develop oral proficiency, in a virtual L2 community?

NOTES

7 See Lepsinger and Lucia (1997) for a description of 360-degree feedback.
8 In this context, term blended refers to a course in which 25% of instruction is delivered via computer.
9 Further data and analyses, as part of the ongoing project, will involve online measurements of noticing and depth of processing.

REFERENCES


Hsieh, H.-C., Moreno, N., & Leow, R. P. (2016). Awareness, type of medium, and


Meeting the Demands of Globalization: One Goal of ISLA Research


12 Task Condition Effects on Advanced-Level Foreign Language Performance

GAVIN BUI,¹ PETER SKEHAN,² AND ZHAN WANG³

¹Hang Seng Management College, Hong Kong
²St. Mary’s University, UK
³South University of Science and Technology, China

Introduction

Language teaching since the 1970s has, in many areas of the world, seen a move away from a focus on structure and form toward a concern to base instruction around meaning-oriented activities in the language classroom. The assumption has been that communicating meanings can catalyze more effective second language development. In the earlier days of this move, the term communicative language teaching was adequate in capturing what happened. But since the mid-1980s the approach has changed somewhat and now is better described as a task-based approach to language instruction. The premise is that giving tasks to language learners to transact will be more effective in driving forward interlanguage development. This approach has had strong pedagogic dimensions (see, e.g. Van den Branden, 2006). But it also, very interestingly, has natural connections with a burgeoning research literature (Long, 2014), and so there is some degree of harmony between the two areas. Each emphasizes its own set of questions and priorities but the degree of overlap between them is stimulating for each area.

These developments are the context for the present chapter. The task-based research literature has grown massively in the last 20 or 30 years, and it has provided insights not simply into factors which might influence pedagogy but also into the nature of performance and its subdimensions (Skehan, 2014a).
these connect with models of first (Levelt, 1989) and second language speaking (Kormos, 2006). What we will set out to do is clarify what the major findings are from this literature, and then explore what can be said about advanced-level language learners as they do tasks and how tasks might have an impact on language development at this level.

To set the scene, four areas emerge from this literature. First, there is the role of task characteristics. Researchers have explored issues such as task familiarity, task structure, the number of elements in a task, the sorts of reasoning demands involved, and so on. We will only touch on this research as appropriate. In the second area, the conditions under which tasks are done (which is the main focus for this chapter), the major issues have been whether there has been an opportunity for planning pre-task; and the time conditions under which the task has been done. There have also been studies looking at the effects of task repetition and whether there is any sort of post-task that has an effect on the earlier performance. Third, there have been more theoretical proposals to try to account for findings in task research. Two particular viewpoints will be mentioned. Robinson (2011) has proposed the Cognition Hypothesis, which suggests that the most important driver for task performance is task complexity: more complex tasks drive greater accuracy and complexity in performance. Task complexity, in turn, is influenced by what Robinson terms resource-directing factors (such as number of elements, reasoning demands, etc.), in contrast to resource-dispersing factors, such as planning or task structure, which have more general procedural effects on performance. Skehan (2014b), in contrast, through the Limited Attentional Capacity approach, suggests that limitations in working memory and attentional resources are fundamental to task performance. He suggests that such limitations make it more difficult for speakers to attain higher levels in all subdimensions of performance (e.g. complexity and accuracy) simultaneously, and that part of the role of task research is to discover judicious combinations of task characteristics and performance conditions that enable second language speakers to overcome such limitations and to achieve higher-level performance (Skehan, 2016). These two approaches have influenced the field of task-based research extensively, and also have implications for advanced-level learners.

Already there are hints of the fourth issue that arises from this research—the nature of task performance itself. Researchers have generally explored performance in terms of complexity (structural, and more recently lexical), accuracy, and fluency, often referred to collectively as CALF (Bui & Skehan, 2018; Housen, Kuiken, & Vedder, 2012). The dimensions have been shown to have some independence from one another (so one can achieve more highly in one or two, but not the others). In addition, it has been argued that there is an implied developmental sequence, going from complexity through accuracy to fluency, with this reflecting increasing degrees of control. This in turn links with the general nature of task-based research. One strand is concerned to establish selective relationships with different aspects of performance (e.g. which tasks and task conditions lead to, for example, higher accuracy), and the other explores how development can be
fostered more quickly, so that new language (indexed through greater complexity) can then be automatized, first without error and then through more rapid use.

This has been quite a long account of the wider nature of task-based research, but it presents the foundation for exploring influences on more advanced learners. Insights come partly from studying such learners in their own terms, but partly also through comparing advanced-level performance with performance at other levels, particularly intermediate-level learners and even native speakers. This is an important research strategy, yet it has to be acknowledged that it is thrust upon us to some degree. Participants in past task-related studies have typically been of intermediate levels. Unfortunately, the narrowness in proficiency level in most studies does not help us to understand the potential interaction of proficiency level with other variables. There is an urgent need to bring in research participants with a range of proficiency levels in task-based language teaching (TBLT) research, especially those with higher proficiency levels (Bui, 2014; Skehan, 2009).

The chapter is organized in four major literature-based sections. These cover the role of pre-task planning, the findings with on-line planning, the role of task repetition, and the insights provided by examining native speaker performance. The sections are followed by suggestions for the future.

Pre-task (strategic) planning with high-proficiency learners

This section will focus on pre-task planning, generally operationalized as offering learners planning time (of different lengths) prior to a task. While 10 minutes of planning time is most commonly adopted in the literature, other lengths such as one minute (Wigglesworth, 1997), three minutes (Elder & Iwashita, 2005), five minutes (Tavakoli & Skehan, 2005), or a combination of times (Li, Chen, & Sun 2015) are also possible. To locate such pre-task planning within a wider context, Bui (2014) showed that pre-task planning belongs to the task-external readiness dimension which offers learners extra opportunity for preparation, in contrast to task-internal readiness such as content familiarity (Bui & Huang, 2018; Qiu & Lo, 2017) where latent ‘preparedness’ is brought about by learners themselves without ad hoc manipulation from an instructor. This section will cover qualitative and quantitative approaches to pre-task planning in turn.

Behavior of advanced learners in pre-task planning

Ortega (2005) provides one of the pioneering views into learner pre-task planning behavior through retrospective interviews with low-intermediate and advanced students. The results were taxonomized drawing on O’Malley and Chamot (1990) and Oxford (1990). This study reported that learners rely heavily on retrieval and rehearsal strategies, with the advanced group using these two strategies relatively more evenly, and the low-intermediate group employing more retrieval strategies, suggesting that the advanced learners assumed retrieval could look after itself
during actual performance. A range of monitoring strategies were also found among the participants, though the advanced group were more effective in this regard. Participants thought about how to make the content more accessible, how they could retrieve simpler vocabulary, and interestingly, how they could avoid more advanced grammar. Some of the advanced participants were also reported to have taken the audience into account during the planning period.

Pang and Skehan (2014) partially replicated Ortega’s study, but with a different coding system (based on Levelt, 1989), and with actual performance data, enabling planning reports to be linked with CALF measures. Again there were two groups, but at low- and high-intermediate proficiencies. These two groups showed largely similar behavior during their planning time: planning small details, general lexical retrieval, some awareness of the listener, and so on. That being said, the findings did show a slight preference on the part of the higher-proficiency group to ‘think of ideas beyond the picture,’ ‘connect the pictures to develop a story,’ and retrieve more connective words. This in turn shows that the participants with more proficiency spent more effort in the ‘Conceptualizer’ stage in Levelt’s model of speaking, while the participants with lower proficiency paid more attention to the ‘Formulation’ stage, and lexical and grammar planning. It was also found that planning may contain limitations and even pitfalls. Subsequent performance may be worse if planning is over-ambitious, or too focused on detailed language, as many learners had problems remembering what they had planned in the actual performance.

A more recent study by Bui and Teng (forthcoming) investigated intermediate (B2 in the Common European Framework of Reference for Languages system) and high-proficiency (A1, A2) learners’ planning behaviors for familiar and unfamiliar topics. In addition to confirming earlier research, this study further revealed that planning brought about affective advantages. Even advanced learners feel more confident and relaxed in speaking when given time to prepare. At the same time, data showed that with familiar topics more proficient students benefitted less than the intermediate group, especially in fluency (Bui, 2014).

To sum up, advanced learners seem more effective in planning the content and going beyond the linguistic level. They are less worried about how to convey messages well, and can spare more attention on what to say, how to tailor messages, and even how to say things interestingly (Ortega, 2005; Pang & Skehan, 2014). When speaking on unfamiliar topics, planning is still an important facilitative factor for advanced learners but such effects attenuate when it comes to familiar topics (Bui, 2014).

**Effects of pre-task planning on CALF for advanced learners**

So far we have only touched on Levelt’s (1989) model of first language speaking, but now we need to outline it more systematically. It proposes three stages in speaking. The first, Conceptualization, is concerned with generating and organizing the ideas to be expressed. This is input to the Formulator, which is concerned with lemma retrieval and syntax building, as well as preliminary phonological
work. The final stage, the Articulator, is concerned with the detail of converting ideas and morphosyntax into actual sound. The central issue is that, in native language speech, the three stages work in parallel, modular, encapsulated fashion. Monitoring can occur but only at designated points in this flow of communication. Vital in this speaking process is the way the Formulator stage draws on the mental lexicon. In native language speech, this lexicon is large, rich, and organized enough to handle the real-time demands made upon by the pre-verbal message from the Conceptualizer, and so parallel, modular functioning can be maintained. In the second language case there are problems because the second language mental lexicon is likely to be smaller, less rich, and more poorly organized, with the consequence that the lemma-retrieval demands create processing difficulties and parallel processing cannot be maintained.

The model, applied to the second language case, is fundamental for understanding what happens as proficiency grows. The key is the nature and accessibility of the second language mental lexicon. As this grows, so speaking (and writing, almost certainly) comes to resemble native language performance, not least in becoming more parallel in nature, in contrast to the more serial mode of production at lower levels of proficiency. In that respect, the task and task condition literature becomes an important window into the nature of the intermediate to advanced transition and the way underlying resources are exploited to communicate in real time. It is against this background that we will now discuss the range of extant studies of the impact of planning on performance.

In Wigglesworth (1997), a study of advanced learners, only one minute of planning time raised complexity and accuracy, and this was more evident in the case of tasks with a high cognitive load. Learners of lower proficiency did not appear to benefit from (this relatively short) planning time. Yuan and Ellis (2003) included 42 advanced-level undergraduates (English majors who scores grades A and B+ in the oral component of the Higher Education Bureau Examination in China). In oral performance, neither fluency or lexical complexity (measured by lexical variety) were significantly improved for the planners, relative to the non-planners. Nor did accuracy show a planning advantage. However, there was an advantage from planning, though, in relation to syntactic complexity. In Ortega (1999), advanced-level students improved with planning for syntactic complexity and fluency, as measured by mean number of words per utterance and speech rate. However, as with Yuan and Ellis (2003), there was no significant difference for lexical variety. Regarding accuracy, the results were mixed in that the measurement of noun modifier target-like use showed improvement while the comparable measure for article use did not. There is reasonable consistency in these studies.

In contrast, Gu (2007) compared a high-proficiency group and an intermediate-proficiency group of Chinese L1 English as a foreign language (EFL) learners. The high-proficiency group benefitted from planning time regarding accuracy and fluency, while it was reported that the low-proficiency group had trouble balancing processing demands with fluency and accuracy, even in the planning time condition. Retrospective interviews indicated that the participants had not achieved readiness for a task (see Bui, 2014), particularly at lower proficiency levels.
The previous studies have focused on oral language performance, but there have also been studies of the effects of planning on writing by advanced-level students. Participants in Johnson, Mercado, and Acevedo (2012) were 968 Spanish L1 learners of English. It was found that pre-task planning had no impact on the lexical complexity or the grammatical complexity of L2 writing. The authors argued that speaking and writing are two different production modes, and that pre-task planning in the written mode yields no observable effect on fluency, grammatical complexity, and lexical complexity, because its effect cannot be teased apart from that of on-line planning. Ellis and Yuan (2004), following up their earlier work on oral production, studied the effect of planning on writing, in the same groups of participants. Contrary to oral task effects, pre-task planning was shown to increase the fluency with which learners produced their work. In addition, there were marked improvements in syntactic complexity and variety. Accuracy, however, showed no differences.

Although the range of pre-task planning studies with advanced learners is not enormous, we can nonetheless offer some tentative generalizations:

- There seems to be more of an effect of pre-task planning on oral rather than written performance (Johnson, Mercado & Acevedo, 2012).
- Syntactic complexity is the performance area most consistently affected by this form of planning, with the vast majority of studies showing this effect. Effects for lexical performance, fluency, and accuracy are not so evident, although they are each reported in a small number of studies (e.g. Yuan & Ellis, 2003).
- Especially regarding the syntactic complexity effect, advanced learners seem to benefit more from planning than do lower-proficiency speakers, an interesting generalization given that most research studies have been conducted with intermediate-level learners (Wigglesworth, 1997).
- It comes across that pre-task planning seems to be more Conceptualizer-oriented with these advanced learners. More Formulator-linked performance features are less affected by planning at this level (Bui, 2014).

**On-line planning with advanced learners**

*The theoretical underpinnings of on-line planning*

Defining on-line planning  According to Ellis’s (2005) framework of task planning, the difference between on-line planning and pre-task planning concerns when the planning time is available. Pre-task planning is conducted before speakers perform a task, but within-task planning is conducted on-line while learners are performing a task. Within-task planning can be further divided into pressured and unpressured on-line planning, with the former usually involving a limited time for on-line planning, and the latter involving unlimited time to perform a task. It is clear that the key concept in on-line planning is the issue of time pressure, and it is unpressured on-line planning that will concern us here.
Time pressure and on-line planning: General issues  

Time pressures are more likely to impede L2 speaking than other modalities. This is because speaking requires highly automatic skills. In comparison to L1 speakers, L2 speakers generally have more time pressure at both the conceptualization and formulation stages (Wang, 2014), due to their limited linguistic knowledge and resources as well as their slower processing speed. Even advanced L2 learners, although equipped with considerable knowledge, are still less automated and are slower in encoding processes. As a result, the construct of on-line planning proposed by Yuan and Ellis (2003) may provide opportunities for learners to have additional time for processing to help overcome processing and resource limitations and then focus on form, bringing about greater accuracy in speech performance.

The operationalization and effects of on-line planning

We will consider two general operationalizations of on-line planning: unpressured on-line planning alone and then hybrid on-line planning where this form of planning is combined with pre-task planning.

Unpressured on-line planning  

On-line planning requires more relaxed and less pressured processing conditions. In practice, this has been achieved in two ways: simply giving learners considerable time to complete a task, or using stimulus material, such as videos, the rate of which has been slowed. The former is relatively unstandardized in nature, since speakers can choose whatever length of time suits them. The latter is more controlled, since the (slowed) speed of the video has to be adhered to. The first approach has been used more frequently, as in Ahmadian and Tavakoli (2011), Ellis and Yuan (2004), Hsu (2017), and Yuan and Ellis (2003). The generalization that has emerged from these studies is that unpressured time conditions are associated with higher accuracy in performance, fairly consistently, and occasionally there are increases in structural complexity.

The second operationalization of on-line planning is less common, and there is only Wang (2014) to discuss. In this study, a video (Mr. Bean) was slowed (but piloting established that this was not noticed by participants) and so the video narration was slowed in a controlled manner. More time was available, but all participants conformed to the same time conditions (unlike the alternative operationalization in the previously viewed studies). In this method, the external speed of the video may shape the control of processing and provide moment-by-moment on-line planning time for the speakers. However, with this more standardized operationalization, Wang (2014) did not find any significant improvement in CALF measures, including accuracy, in the narrated performance, so some slight re-evaluation of the effects of on-line planning may be needed.

Hybrid on-line planning  

In general, researchers have explored either pre-task planning or on-line planning. But an alternative approach, what might be termed hybrid on-line planning, would be to combine the two of them within one study. For example, Ellis (1987) asked participants to tell a story based on a set of pictures...
but after they had been given unlimited time to write about the story, to compare with telling the story with no such planning opportunity. Hsu (2017) also provided L2 speakers with 10 minutes of pre-writing opportunity as rehearsal planning in addition to unlimited task completion time for the actual performance. Wang (2014) combined task content familiarity and pressured on-line planning in the operationalization of a ‘Watched on-line planning’ condition by allowing the speakers to watch the video first and then narrate the story played at a slowed speed in an on-line planning condition. Hulstijn and Hulstijn (1984) investigated another type of hybrid on-line planning, which combines unpressured on-line planning (operationalized as ‘take as much time as you wish’) with focus of attention (on either grammatical correctness or on information).

Results from these studies are fairly consistent. Ellis (1987) found that the hybrid of pre-task and on-line planning led to higher L2 accuracy regarding regular past tense (-ed) although no significant effect was found for irregular past tense measures, a distinction between rule-based and lexical-based language. The accuracy effect from on-line planning was also discovered by Hsu (2017) and Wang (2014), who, in addition to raised speech accuracy, found that the hybrid on-line planning (i.e. “Rehearsal Online Planning” in Hsu, 2017, and “Watched Online Planning” in Wang, 2014) enhanced both speech accuracy and complexity, a case that language ‘trade-off’ (Skehan, 2014a) can be mediated by effective task conditions. Hulstijn and Hulstijn (1984) found that the focus on grammar conditions helped produce significantly higher accuracy of two grammatical elements while the two on-line planning conditions did not show significant influence on accuracy. Meanwhile and interestingly, the focus-on-form groups took more time in processing the speech than the focus-on-information groups, somehow an indicator of careful on-line planning undertaken during the focus-on-form process. It would seem from these studies that unpressured on-line planning opportunities constitute a necessary but maybe not sufficient condition—they need to be supplemented by some additional, usually pre-task activity, and this delivers consistent accuracy and occasionally complexity improvement.

**Summary of on-line planning effects and high-proficiency learners**

From the above review of on-line planning conditions, we can see that the participants in these studies are all intermediate-to-advanced learners (except for Ellis, 1987, with intermediate learners). Regarding the effects of these on-line planning conditions, there are three patterns identified.

- Unlimited (but unstandardized) time for on-line planning enhanced speech accuracy (Yuan & Ellis, 2003).
- More controlled time conditions for on-line planning investigated in Wang (2014) did not lead to an increase in speech accuracy.
- Hybrid on-line planning conditions when combining with other instructional elements (either by means of focus on form, pre-writing, pre-watching,
rehearsal, or repetition) resulted in higher complexity and accuracy in participants’ oral or written performance—an example of task conditions being able to break through the performance trade-offs between complexity and accuracy (Skehan, 2015).

Task repetition effects with advanced learners

Defining task repetition

Task repetition, by definition, means a second language speaker repeats a task that has already been performed. This basic account then leads on to a consideration of a number of related factors. The interval between the original and the repeated task may vary. In addition, the repeated task may be essentially similar to the original, but it is possible that small changes might occur. There is also the possibility that a task may be repeated more than once. More radically, there is the possibility of a repeated task type, rather than an exact task.

Justifications for task repetition

Fundamentally, with task repetition the original performance releases attentional resources to enable a focus on form in later performances (Skehan, 2014b) because the previous engagement of the task can serve as a useful preparation (Ellis, 2005). This preparation can be in three areas, each linked to the Levelt model of speaking (1989). First, there are the benefits that derive from the conceptualization that takes place in the first performance. This clarifies the ideas that are to be expressed in the second performance, meaning that much conceptualization work is done, and in the second performance, additional conceptualization could embellish what has already been achieved, or more probably, release attentional resources for other stages of speaking. Second, during the first performance, lemma access will take place, but some of the time there will only be enough time for this to be partially done (as when some lemma information, probably the most superficial, is retrieved but rich information is not accessed). Then the repeated performance can capitalize on the earlier partial lemma retrieval, and enable deeper and more comprehensive formulation processes (Skehan, 2014b). Finally, there is the feedback opportunity from ‘monitoring’ (Levelt, 1989) during the learner’s initial task enactment (Wang, 2014). According to Levelt’s first language speaking blueprint, speech monitoring can operate at three locations: the pre-verbal message, the inner speech plan, and the overt speech plan (Levelt, 1989). When producing the first task performance, a speaker already involves his or her speech comprehension system, which will be perceived simultaneously by the speaker and the interlocutor (Samuda & Bygate, 2008). This speech comprehension system naturally triggers the speech monitoring system, resulting in the parsing of the speaker’s speech when performing the task for the first time. The monitoring opportunity provides feedback to the speaker and may elicit self-correction during his or her speech, so
that, when the task is repeated, released attentional resources can attend to
linguistic form, meaning, or the mapping between form and meaning (Wang,
2014). As a result, accuracy in speech production, as well as speech complexity and
fluency, will be enhanced in later task performances, as cited in the literature
(Ahmadian & Tavakoli, 2011; Fukuta, 2016; Lynch & Maclean, 2000).

Operationalizations of repetition

There are three broad operationalizations of repetition: immediate repetition,
repetition after a time interval (e.g. two to three days, or one week, etc.), or multiple
task repetitions. We will consider each in turn.

Immediate task repetition

A typical operationalization of an immediate task is to ask learners to perform a
meaning-focused task again to a different interlocutor. Lynch and McLean (2000)
asked L2 learners to make a presentation of a poster to different interlocutors six
times in an English for academic purpose class. They found that the intermediate
proficiency learners benefited from the task repetition with improved speech accu-
racy and fluency. Wang (2014) asked participants to perform a task to an imagined
listener immediately after their first task performance, and found that learners
enhanced their speech complexity, accuracy, and fluency in the second speech
performance and this with large-effect sizes.

Less typical as an operationalization of immediate repetition is Lambert,
Kormos, and Minn (2016), who investigated the effects on CALF of repeating the
task six times to different interlocutors. The results reveal that, regardless of the
learner’s proficiency (i.e. high, mid, and low), immediate task repetition effectively
enhanced speech fluency on various measures—although different measures
began showing significant changes at different rounds of task repetition. Speech
rate, interpreted as a general indicator of performance, improved until the fifth
replication. Clause final filled pauses, interpreted as Conceptualizer-related,
improved up to the second, while mid-clause filled pauses continued to decrease
up till the fourth performance, and this was interpreted as Formulator-linked.
Self-repairs, taken to reflect monitoring, only started to change after the fourth
performance. Lambert et al. (2016) relate the different repetition cycles to different
Leveltian stages. A later section considers other multiple repetition studies.

Task repetition after a time interval

Another set of empirical studies investigated intervalled task repetition (usually
having intervals of two–three days or one week). For example, Bygate (1996)
found that an intermediate learner improved speech fluency and accuracy by
repeating the same task after a three-day interval. Ahmadian and Tavakoli (2011)
and Fukuta (2016) investigated the effects of task repetition in one week on the
CALF measures, but had different results. With intermediate learners, Ahmadian
and Tavakoli (2011) found that task repetition resulted in higher speech complexity, accuracy, and fluency, whereas Fukuta (2016), working with upper-intermediate learners, found speech accuracy and lexical effects (measured by the Guiraud Index).

**Multiple task repetitions with time intervals**

We have already covered the Lambert et al. (2016) study of multiple immediate repetitions, and now will explore multiple repetitions with time intervals. These repetitions usually occur over intervals longer than one week or as a regular training method in class. Bygate (2001) investigated the effects of multiple task repetition occasions (and task type familiarity as well) on L2 speech performance, and found that the effects from the first enactment of a task was durable as it enhanced the learners’ speech complexity and fluency when they conducted the same task over a 10-week interval. Gass, Mackey, Álvarez-Torres, and Fernández-García (1999) also investigated the repetition of the same task (and the repetition of task type) across three time points with a two–three-day interval each time, and found that third task performance outperformed the non-repetition group on general proficiency, accuracy of a Spanish structure ‘to be,’ morphosyntax, lexical density, and lexical sophistication (i.e. the number of difficult words used), but the effect did not generalize to a new task. De Jong and Perfetti (2011) investigated a 4-3-2 repetition task (i.e. repeating a topic for four, three, and two minutes) as the training method, and found that the group that repeated the same topic three times with increased time pressure had significantly higher fluency in a post-test than the group that spoke on three different topics each time for a total of three training sessions over two weeks of training as a control.

**Summary of task repetition effects with advanced learners**

Reflecting on the studies that have just been covered, three general patterns emerge:

- Immediate task repetition has robust effects on speech complexity, accuracy, and fluency (CAF), but it does not affect lexical performance. Earlier it was argued that this results from engagement of Conceptualizer operations (complexity), which ease subsequent Formulator operations, a stage also assisted by the capacity to exploit the partial lemma access from the first performance. We also saw that monitoring processes are helped through repetition and this specifically influences accuracy. Finally, and remaining with the stages of the Levelt model, Lambert et al.’s study of multiple immediate repetitions suggests that the Conceptualizer stage is influenced first, followed by the Formulator, and finally monitoring (Lambert et al., 2016).
- Task repetition after a time interval, typically two or three days or even a week, enhances speech accuracy and fluency (Ahmadian & Tavakoli, 2011).
• Mixed results have been found from multiple task repetitions across long intervals, although such studies often contrast task repetition (same content) and task type repetition, and this may cloud the consistency of the results that emerge. There are also issues here with specific operationalizations in different studies and also the different foci of these studies, which render comparisons difficult.

Perspectives from native speaker performance

We have drawn attention to the problem that the database for task performance with advanced learners is not great. As a result, we have also covered some intermediate learners and their performance to try to illuminate the advanced level. In fact, we can do the same thing but from a slightly different direction by exploring native language performance. The database here is not that great either, but a growing number of studies are relevant. In each case, researchers have gathered data from native speakers and from non-native speakers doing the same tasks so that a comparison can be made and (i) the nature of the difference in performance between the two groups can be more clearly understood, (ii) effects that derive from the task itself (and which are similar for both groups) and tasks which generate different results for the two groups can be revealed more clearly. Perhaps a final introductory point is that some studies have compared different native and non-native groups doing the same tasks, and others have been more concerned to examine the very same learners, in their L1 and L2, but where there may be more variation in the tasks they do.

Foster (2001) administered an identical task to (different) native and non-native speakers, and used the additional variable of pre-task planning time. Her focus was on the tendency to use formulaic language. She proposed that using formulaic language is associated with reductions in processing pressures and so the speaker is more able to use chunks of language that do not require any internal processing. She showed that unplanned native speakers used more formulaic language than planned native speakers, but that, in contrast, unplanned non-native speakers used less formulaic language than planned non-native speakers. It was as though the native speakers, when deprived of planning time, relied on the crutch of formulaic language (which came easily to them) whereas they used planning time in order to be more creative (and less formulaic) in their language. The non-native speakers, in contrast, when deprived of planning time, did not have the resources under more time pressure to access processing-easing chunks of language, but when they had time to prepare, they helped themselves even further by accessing less processing-heavy language. Skehan (2009) also examined this same dataset, as well as a parallel non-native speaker dataset (with the same tasks), but his focus was on lexical performance. He looked at three tasks, (personal information exchange, narrative, and decision-making) and two measures: lexical diversity (the extent to which speakers recycled the same words, or not) and lexical sophistication (the extent to which learners used less frequent words, as judged by the British National Corpus wordlist). He showed that similar patterns occurred for
both groups of speakers, i.e. which task produced the highest performance on each of these measures, and so on, was exactly the same, suggesting a strong task type effect. But he also showed that while the lexical sophistication score was mostly significantly higher for the native speakers, the difference was not so marked as it was for lexical diversity, where the native speakers clearly recycled words less than the non-native speakers. These measures seem to be more sensitive indicators of level of proficiency and how to distinguish intermediate from advanced from native speakers.

Tavakoli and Foster (2008) provide an important insight on usage-based perspectives with language. They researched three groups: intermediate Iranian learners of English in Iran; Iranians in London; and native speakers. The two groups of non-native speakers were equated for proficiency, so there were no striking differences in general performance on tasks (e.g. CAF). Where they did differ, however, was in the nature of the lexical choices that they made. The London-based Iranians were much more likely to make lexical choices that resembled those made by native speakers. There may not have been any differences in word frequency in the two groups, i.e. their lexical sophistication was not the issue. But they did seem to have been influenced by the speech around them so that they used more general and flexible lexical choices as though this was the aspect of the surrounding input that had had most impact.

Skehan and Shum (2014; and see also Skehan, Foster, & Shum, 2016) also compared native and non-native speakers. They report little difference between these groups as regards structural complexity or lexical sophistication when they completed video-based retelling tasks. Accuracy was not addressed, since it was assumed that native speakers might have lapses, but not make errors as the non-natives did. But there were three striking findings. First, there were considerable differences in fluency between the two groups. This manifested itself in three areas. The native speakers spoke very considerably faster than the non-natives. They also had much greater lengths of run, reflecting automaticity in performance. Most intriguingly, the pattern of pausing was different in the two groups. Native speakers pause, obviously, but tend to do so at the ends of clauses. Non-natives also do this but they show much greater pausing mid-clause, which is much less common for the native speakers. Second, the lexical diversity for the native speakers was much higher. Once again, native speakers showed a considerable capacity to recycle less, and packed more different words into their performance. Third, there was strong evidence of style for some aspects of performance. Pausing, repair, and lexical diversity showed high correlations across tasks—if someone (native speaker or non-native speaker) paused frequently, they tended to do so on all tasks. In contrast, complexity scores did not show such high correlations, and for the non-native speakers, nor did accuracy scores. These, in contrast, seemed more task dependent. The study in question contained different native (L1 English) and non-native (L1 Chinese) speakers in English. Two other studies (Derwing, Munro, Thomson, & Rossiter, 2009; De Jong, Nivja, Groenhout, Schoonen, & Hulstijn, 2015) researched the same individuals speaking in their L1s and their L2s. These studies too report strong style effects in the fluency domain.
The native speaker studies are interesting because they do enable some generalizations to be made which are insightful for the nature of advanced language learner use.

- Lexically it is clear that even advanced learners have challenges. Native speakers clearly command a wide range of formulaic language, and advanced learners need to get to grips with this stock of formulaic language to begin to sound native-like (Pawley & Syder, 1983). Studies such as that of Foster (2001) and Tavakoli and Foster (2008) show how tasks can help to clarify this challenge, and how task implementation conditions can foster the use of such language. Similar considerations apply to native-like choices which, on the basis of the Tavakoli and Foster (2008) study, require exposure to considerable quantities of native speaker input, something, perhaps, which is more feasible in this digital age.

- There are also important issues regarding access to the second language mental lexicon. The clear differences between native speakers and all second language speakers doing tasks with regard to lexical diversity suggest that drawing on the mental lexicon to avoid repetition, compared to relying on the same words, is telling. This may simply be a reflex of proficiency and size and accessibility of mental lexicon, but it certainly serves to characterize native speakers. More broadly it appears that automatization, and having a mental lexicon that can withstand the lexical demands of the pre-verbal message and thereby avoid the need to pause mid-clause is very important (Skehan, Foster, & Shum, 2016; Skehan & Shum, 2014).

- Looking at the fairly standard performance measurements of CALF, one has to take note of the personal style influence on fluency. There are differences in level of fluency that distinguish native and non-native speakers but it appears that, cross-task, particular individuals are strongly driven by style in this performance area. The particular task does not seem to be very important. In contrast, structural complexity and accuracy seem less implicated in style, and vary as a function of different task (De Jong et al., 2015; Skehan, 2014b).

**Suggestions for the future**

Of course, the main suggestion has to be that there is a need for more research. Advanced learners have not fared well with task-based research, and there is an urgent need to explore to what extent the findings with other groups, including native speakers as well as intermediate learners, are consistent or differ in interesting ways from those with higher-level learners. There is a lot to be discovered about the details of planning (strategic and on-line), of repetition, of post-task effects, of task design features, and so on. This would generate a considerable program of research in itself. There is also the issue of trying to account for the path and the nature of development as one progresses from beginner, to intermediate, and to advanced. In this viewpoint, we need the research with advanced learners in order to see what differentiates them from other proficiency
levels and from native speakers. This would be helpful in understanding the nature of development. It would also be useful for assessment purposes in having a clearer view of what changes as proficiency changes.

These, though, are rather limited views about the future. More challenging are ways task-based research can contribute theoretically to wider issues in applied linguistics as the research focuses on advanced learners. The characteristics that are most likely to lead to this are the way measurement advances in the field (of complexity, structural and lexical, of accuracy, and of fluency) can help us understand what advanced proficiency is, and the way models of second language performance and speaking can be illuminated. We now have more sophisticated and multiple ways of measuring structural and lexical complexity, often in an automated manner (Skehan, 2014a). The same is true for accuracy, and even more so for the dimensions of fluency that are now more clearly understood (breakdown, repair, speed; Tavakoli & Skehan, 2005). Until now, the dominant approach to characterizing advanced proficiency has been assessment oriented, with rating scales of different atheoretical features of language knowledge and use. The indices used by task-based researchers can give us a different perspective on what makes an advanced learner advanced, both through the levels of performance on the different indices and also in terms of their interrelationship, and how slightly lower performance in one or two indices may be compensated for by higher performance in others.

But the most important challenges are those of a more theoretical nature, and two stand out. First, we can make progress in understanding what sort of model of speaking is appropriate to characterize advanced learners. At present the most common model of first language speaking used with second language speakers is the Levelt model (1989, 1999), a model which proposes stages within speaking which work, in parallel, in modular fashion. Intermediate second language speakers often function in a more serial manner, as problems at one stage of speaking, typically Formulation, derail the parallel modular system because of inadequacies in the second language mental lexicon (Kormos, 2006; Skehan, 2014a). This, in turn, has consequences for the detailed measures of performance and the way different task and task condition variables work. Advanced learners then become the key to showing us how the basic model of speaking, with a second language learner, can become more parallel and modular in nature. Advanced learners need to be studied to explore, with task measures, how performance is disrupted or how it is facilitated, and under what conditions, and to what degrees of advancedness, one can describe performance as parallel.

The second theoretical advance could be with a debate within the second language field—between different accounts of second language task-based performance. Two contrasting approaches have generated most research—the Cognition Hypothesis (Robinson, 2011) and the Limited Attentional Capacity approach (Skehan, 2014b). Skehan (2014b) argues that working memory and attention limitations mean that difficulties can occur in producing performance that is simultaneously higher for complexity, accuracy, and fluency (and that a target for research is to find tasks and task conditions which minimize the effect
and increase the chances that performance will be raised in most performance areas). Robinson (2011) argues that task complexity is the key to driving performance higher in multiple dimensions. As pointed out earlier, the majority of task studies have been conducted with intermediate-level learners. The two approaches differ in their view of attentional functioning, but it may well be that a key variable is proficiency level. The expandable view of attention (and the predicted jointly raised accuracy and complexity) may be more compatible with the less studied advanced learners, whereas the more constrained view of attention may be more appropriate with lower-level learners. This suggests that systematic studies which vary proficiency level, and which include both groups of learners, may make an important contribution to clarifying our understanding of the functioning of these two approaches.

NOTE

1 The co-authors shared their contributions to the chapter and the author order is purely alphabetical.

REFERENCES


Part III Advanced Phonology
13 Advanced-Level L2 Phonology

JOHN ARCHIBALD
University of Victoria, Canada

Introduction

The field of second language (L2) phonology investigates knowledge and skills in the domains of phonetics and phonology (Archibald, 1998). Phonology is concerned with the nature of mental representation and how those representations encode meaning contrasts, while phonetics is concerned with how individuals access those representations in real-time production and perception.

It is immediately obvious that foreign accents exist. Native listeners are very good at identifying non-native speech (and often the first language of the speaker) by phonetic properties of the speech (Scovel, 1988). In lay terms, we often hear references to ‘strong’ accents, or ‘heavily-accented’ speech. This may stem, in part, from the fact that the end state of L2 grammars is much more variable than the end state of first language (L1) grammars. Someone’s non-native German may be much ‘better’ than their non-native Japanese, for example. But what does it mean to be better? More accurate? More fluent? More native-like?

How should we define the construct of ‘advanced’ in the domain of L2 phonology? Given what we know about the separate components of linguistic competence, and the separate components of proficiency (or communicative competence), it is clear that there is no straightforward correlation between, say, a high International English Language Testing System (IELTS) score and a reduced foreign accent. There are many cases of highly proficient L2ers with strong accents (Henry Kissinger, Joseph Conrad). Conversely, we may be able to achieve by mimicry quite native-like speech on isolated phrases in a language we do not even understand. How would we identify the locus of advanced phonology?

In this chapter, I will argue that there is no single hallmark of ‘advanced’ L2 phonology. We will explore how phonological systems compare, and how L2
speech may be viewed through a lens of native-likeness or comprehensibility. We will look at the different components of phonological knowledge to see whether some are more difficult to acquire (and hence a diagnostic of more advanced systems). Pursuing the articulation of such a construct, however, has the benefit of revealing that, while L2 accents are ubiquitous, there are no phonological features or structures which are blocked from acquisition. Not all speakers may be advanced, but all can advance.

The locus of ‘advanced’ phonology

As a first step in formalizing the notion of advanced phonology, consider the common distinction between knowledge and skill. Knowledge is a relatively stable trait in an individual. You either have the knowledge (i.e. mental representation) of the English words *cat* or *quodlibet* or you do not. Native speakers of English know that /l/ and /r/ are contrastive in English, while native speakers of Japanese know that they are not in Japanese. Native speakers of French know that [p] and [b] are contrastive, while native speakers of Thai know that [pʰ], [p], and [b] are all contrastive. However, as with all aspects of human knowledge, our ability to access or implement that knowledge varies under certain real-world conditions. We, thus, invoke a construct such as proficiency or fluency to try to capture the notion of skill or ability in using a second language. Under this approach then, what does it mean to have ‘advanced’ phonology? Advanced knowledge or advanced skill?

Advanced knowledge

How could we conceptualize of grammar A being more advanced than grammar B? Comparing grammars can be fraught with unintended consequences, as we shall see, but in order to compare, we would have to invoke some sort of evaluation metric. ‘More native-like’ would be one possible measure. However, this is not as straightforward as it might seem. One possibility would be that we could look to a naturalistic order of acquisition (such as was found for, say, L2 German word order by Meisel, Clahsen, & Pienemann, 1981) and say that the later stages were more advanced. The developmental stages proposed were:

1. Canonical word order
   Die kinder spielen mim ball
   the children play with the ball
   The children play with the ball.

2. Adverb preposing
   Da kinder spielen
   the children play
   The children play.
3. Verb separation
   Alle kinder muß die pause machen
   all children must the break have
   All children must have a break.

4. Inversion
   Dann hat sie wieder die knoch gebringt
   then has she again the bone bringed
   Then, she has brought the bone again.

5. Verb end
   Ich trank das glas milch, während ich den brief schreib
   I drank the glass of milk while I the letter wrote
   I drank the glass of milk while I wrote the letter.

It is not straightforward to construct a phonological analogue of this type of sequence.

Another possibility would be to look at the two grammars within some sort of parametric model where the grammatical strings generated by various values of the parameter in question generated a nesting of grammars. An example of this type of model would be Broselow and Finer (1991), who look at the L2 acquisition of English onset clusters. Consider someone whose L1 does not allow any onset consonant clusters who is learning English as a second language. The target grammar allows many consonantal sequences which are generated by different values of a Minimal Sonority Distance parameter as shown below, based on the sonority scale of stop < fricative < nasal < liquid < glide < vowel:

So, if the L1 had the setting of 4 and the L2 had the setting of 2 then we might argue that as the learners moved from an MSD setting of 4 to 3 to 2, they were becoming more ‘advanced.’

*Functional load* Another way that I can conceive of a construct of *advanced knowledge* being operationalized is as follows. Imagine two non-native

<table>
<thead>
<tr>
<th>Table 13.1</th>
<th>Minimal Sonority Distance (MSD) and English consonant clusters.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSD</strong></td>
<td>4</td>
</tr>
<tr>
<td><strong>Clusters allowed</strong></td>
<td>pj, tw</td>
</tr>
<tr>
<td></td>
<td>pr, pl</td>
</tr>
<tr>
<td></td>
<td>fl, fr</td>
</tr>
<tr>
<td></td>
<td>sf, fn</td>
</tr>
</tbody>
</table>
speakers who were each able to successfully contrast 70% of the phonemes of the target language but who differed in their ability to successfully contrast the phonemes which were most important in the L2. Differing intelligibility paradigms might emerge. The critical element would be to define which phonemic contrasts were ‘most important.’ The only way that I can think of to do that would be via a metric of *functional load* (Martinet, 1933).

Imagine that a language has 500 word pairs that are contrasted by [i]/[ɪ] but only 200 word pairs which are contrasted by [n]/[ŋ]. In this language, then, the [i]/[ɪ] contrast would be said to carry greater functional load than the [n]/[ŋ] contrast. So, would a speaker who can reliably produce or perceive the former contrast be more advanced than a speaker who could produce or perceive the latter?

The feature competition model  The early work of Hancin-Bhatt (1994) addressed this question indirectly. She was not concerned with the notion of ‘advanced’ but she was concerned with which L2 contrasts were easiest to learn, and sought to determine if contrasts which carried a high functional load in the L2 were easier to acquire than contrasts which carried a low functional load. Admittedly, her empirical results were mixed.

Functional load was invoked to account for common substitution patterns. We might note that Greek L1 speakers may substitute [s] for English [ʃ] and Arabic L1 speakers may substitute [b] for [p]. Pre-theoretically, we might say that the speakers substitute the sound which is most similar from their L1. Hancin-Bhatt introduced the Feature Competition Model to explain how L2 sounds are mapped onto L1 categories. When perceiving the L2 input stream, the L2er will notice elements which are signalled by the most ‘prominent’ L1 features. Feature prominence is derived by a formula that divides the number of L1 phonemes for which the feature in question is specified1 by the total number of phonemes in the L1. Ignoring many details, she proposed feature-prominence scales for different languages (where > is read as ‘more prominent,’ and / is read as ‘equal prominence’):

- Hindi: [voice] > [strident] > [continuant]
- German: [strident]/[continuant] > [voice]
- Japanese: [strident] > [voice]/[continuant]
- Turkish: [voice] > [continuant]/[strident]

Ultimately, she found that functional load was *not* a robust explanatory for which contrasts were acquired early or easily.

However, there is work within a number of paradigms which suggests that the whole concept of grammars being more or less ‘advanced’ is problematic. Certainly work within a sociolinguistic framework (Labov, 1972) has shown that prestige and non-prestige grammars are equally complex. As Sawallis (1997) engagingly noted, “consonant deletion is phonology not lethargy.”
This is also reminiscent of what we see in discussion of heritage speaker grammars. As Kupisch and Rothman (2016) remind us, heritage speaker grammars are not incomplete, but rather different.

Ultimately, we must conclude that there is no clear way to define advanced L2 phonology. So, what about advanced phonological skill?

**Advanced skill**

In discussing skills, we need to make a further distinction between *production* and *perception*. As Stampe (1972) and Boersma (1998) have both reminded us, there are often competing pressures on production and perception, so the operationalization of *advanced* might look quite different depending on the perspective taken.

*‘Advanced’ production*  It strikes me that there would be two possible measures of advanced production by an L2er: (i) closeness to native-likeness, or (ii) successful communication.

*‘Advanced’ perception*  The measure of advanced perception by an L2er would have to be their ability to accurately recover the intended message of the speech of the native speaker (NS). This gets us a bit closer to formalizing what it might mean to be ‘advanced’ in production or perception, but there is one further distinction we need to introduce before we can fully define what ‘advanced’ might mean. As Munro and Derwing (1995) have made clear, we need to separate constructs like *intelligibility* from *comprehensibility*.

If L2 speech is intelligible, then the listener understands the intended message; if unintelligible then the listener does not understand the message. If L2 speech is more comprehensible then there is little effort required to process the signal and understand the message, while if it is less comprehensible then it takes more effort to recover the message.

To explicate these definitions, imagine the native listener as being engaged in a kind of lexical decision task. They are exposed to an input stimulus and have to retrieve the intended lexical item of the speaker from their own lexicon. The *accuracy* of matching the intended item is captured by *intelligibility*, while the amount of time or work (or *response* time) that the native listener has to engage in is captured by the construct of *comprehensibility*. Consider the following two hypothetical conversations:

Table 13.2  Deletion versus epenthesis repair strategies.

<table>
<thead>
<tr>
<th>Deletion/Substitution</th>
<th>Epenthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: I need to get a [kæʔ]</td>
<td>A: I need to get a [kæpə]</td>
</tr>
<tr>
<td>B: I’m more of a dog person.</td>
<td>B: A [kæpə]? Oh, a cap!</td>
</tr>
<tr>
<td>A: No, a <em>baseball</em> [kæʔ].</td>
<td></td>
</tr>
</tbody>
</table>
In the first conversation, there was an intelligibility problem. The listener accessed a lexical item different than the one intended by the speaker. In the second conversation, there was a comprehensibility problem. The listener had difficulty retrieving the intended lexical item, but ultimately did so.

Thus, we propose that more ‘advanced’ phonology would be both more intelligible and more comprehensible. But, if there were a hierarchy, comprehensible speech would be more advanced than intelligible speech insofar as it implies producing an utterance which is decoded with minimal effort. In this chapter, I will focus on the construct of comprehensibility.

The locus of the difference

Clearly, accentedness can (but does not necessarily) reduce both comprehensibility and intelligibility (Munro & Derwing, 1995). When the speech is not target-like, it can affect the recoverability of the message. But let us probe the nature of that term non-target-like.

In order to do so, it is necessary to distinguish the construct of phonology from phonetics in order to describe and explain why non-native speakers sometimes sound non-native-like, or why listeners sometimes have difficulty retrieving the intended lexical item.

Let me illustrate with an example from stress placement in L2 production. Take a language like English where stress placement is complex and can result in differing syllables being stressed, as in the following words:

First syllable: cinema
Second syllable: aróma
Third syllable: cigarétté

Languages vary in terms of which syllables are stressed, as indicated in the following patterns:

Czech: stress the first syllable
Polish: stress the penultimate syllable
French: stress the final syllable

If learners from these L1s were transferring their L1 stress properties, then they would sometimes get the English stress right, and sometimes wrong.

<table>
<thead>
<tr>
<th></th>
<th>Target initial</th>
<th>Target penultimate</th>
<th>Target final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech L1</td>
<td>cinéma</td>
<td>áróma</td>
<td>cigarette</td>
</tr>
<tr>
<td>Polish L1</td>
<td>cinéma</td>
<td>aróma</td>
<td>cigárette</td>
</tr>
<tr>
<td>French L1</td>
<td>cinémá</td>
<td>aromá</td>
<td>cigarétté</td>
</tr>
</tbody>
</table>
Which syllable is stressed is determined by properties of the phonological system. However, stress is also implemented phonetically in different ways in different languages. It is often realized by a combination of longer duration, higher pitch, and greater intensity. But languages vary in terms of whether unstressed vowels change in quality. In English, unstressed vowels are reduced to schwas while in Spanish they are not. So, an L2 speaker may be placing the stress on the correct syllable but implementing the prominence of the syllable via the L1 phonetic cues.

E.g. ‘photography’ implemented as [fotágrafí] as opposed to [fatógrafí].

In this case as well, the listener may have to work harder and perhaps compute and compare a couple of alternatives before recognizing the intended vocabulary item.

**Advanced with respect to what?**

So, ‘advanced’ knowledge would be more target-like phonology while ‘advanced’ skill would be more target-like phonetics.

Munro and Derwing’s use of the term *intelligibility*, however, brings to mind Levis’s (2005) distinction between the *Intelligibility Principle* and the *Nativeness Principle* when it comes to assessing L2 pronunciation. For Levis (and others), the goal of instructed pronunciation is to help the learner become more comprehensible—not to measure only by the bar of native-likeness. As Munro (2008, p. 194) notes, “native pronunciation in the L2 is not only uncommon but unnecessary.”

**Factors influencing phonological outcomes**

When we see variation in a particular learning outcome, there is, of course, a desire to seek out variables which might explain that variation. For example, when it comes to ultimate attainment, we might ask whether aptitude (Saito, 2017) or motivation (Saito, Dewaele, & Hanzawa, 2017) would lead to more advanced proficiency. The question which emerges here is: Does explicit instruction lead to more ‘advanced’ phonology?

**Effects of instruction**

Given that we have seen decided age effects when it comes to aspects of second language speech related to pronunciation (and particularly global accent ratings), it certainly is not surprising that we have seen an investigation of what factors can influence the global accent of L2 learners. For example, Abrahamsson and Hyltenstam (2008) noted that many near-native speakers with a late age of acquisition (AoA) had high language aptitude scores. But what of other factors? One factor that has received considerable attention is explicit instruction. Bongaerts, Mennen, and van der Slik (2000) argue that explicit phonetic instruction helps in the establishment of near-native pronunciation ability.
Thomson and Derwing (2015) provide a very thorough overview of the training studies. The broad structure of these types of studies is to take a measure of some aspect of pronunciation in a pre-test, provide some type of treatment to an experimental group, and then see if the performance of the experimental group differs from the control group. The following are some key points they raise:

- Many studies take native-like pronunciation as the measure of ‘success.’
- Segmental phenomena are studied more than suprasegmental phenomena.
- Few studies look at spontaneous speech production but rather focus on read-aloud methods.
- Some studies lack a control group, which makes it difficult to assign a causal role to the experimental treatment.
- Some studies show treatment effects lasting into a delayed post-test (Couper, 2006) while others do not (Ruellot, 2011).

Many of the studies indicate that the pronunciation instruction leads to improvement. Thomson and Derwing report that 82% report a significant improvement. However, they also note that this doesn’t necessarily tell us whether the treatment led to more comprehensible L2 speech. Thomson and Derwing (2015, p. 335) note, “we take the view that native-like pronunciation is an unrealistic goal, but that improved intelligibility and comprehensibility are achievable, and that L2 pronunciation research focusing on these speech dimensions is of more practical value than studies of accentedness alone.” They also draw from these studies that bringing about a global change in intelligibility or comprehensibility will require weeks or months of instruction, not hours or days. Many of the studies focus on a very limited number of phonetic features, and it may be unsurprising that the focusing of attention on a narrow item will bring about a change in behavior. The question remains as to whether these experimental results would transfer into a more ecologically valid task.

The role of attention

Related to the role of instruction in acquiring a new phonological contrast is the role of attention. There is a broad literature on attention in second language acquisition (SLA; see Robinson, 1995; Schmidt, 1990, 2001) but here I want to refer to the work of Guion and Pederson (2007), who look to see if students learn new sounds better when they attend to them. Working within Flege’s (1995) Speech Learning Model, they accept that adult L2 learners retain the ability to form novel phonetic categories, and that the ability to discern the phonetic differences between the L1 and L2 sound is a necessary precursor to novel category formation. Furthermore, they acknowledge that L2 categories may be cued by a variety of phonetic differences to which the L2 learner may have to become attuned. They looked at monolingual English subjects who were exposed to Hindi retroflex ([ʈʰ]) and dental ([ʈʰ]) stops. One group of learners was instructed to attend to the phonetic differences alone (sound-attending), while another group was instructed to attend to
sound–meaning correspondences (meaning-attending). They sought to determine whether greater attention to phonetic form leads to greater phonetic learning. Seventy-six monolingual English subjects were divided into the two groups. In addition to the contrast of interest ([ʈʰ]/[ʈʰ]), subjects were exposed to other contrasts:

[b]-[tʃ]; [k]-[ɡ]; [b]-[b]; and [k]-[kʰ]

After a pre-training discrimination test (to act as a baseline measure), there was a self-paced training session of 30–40 minutes in length. The sound-attending group was given the following instructions:

Your task is to listen carefully to the beginning of each word and try to learn the difference between the Hindi sounds.

The meaning-attending group was given the following instructions:

Your task is to try to learn the meaning of the Hindi words.

After the training a discrimination test was given. This was an AXB task with an ISI of 750 ms. The sound-attending group showed significant improvement on the target contrast. The group differences on other contrasts were not significant. Thus, they conclude that directing attention does have the ability to affect phonetic learning.

Work using high-variability phonetic training (Logan, Lively, & Pisoni, 1991) has also shown that exposing learners to multiple exemplars of words spoken by different speakers may improve their identification of non-native speech contrasts. It is argued that the variation between items focuses the attention of the listener on the critical phonetic cues.

**Working memory**

One factor which has received considerable attention is working memory (WM). Can WM tell us anything about advanced L2 speech? Darcy, Park, and Yang (2015) tested WM, attention control, and processing speed to see whether there were correlations with phonological processing in domains such as phonetic categorization, complex onsets, and word stress.

The following cognitive tasks were assigned which were assumed to draw on WM.

1. Simple Span Task. Forward digit, backward digit, forward non-word, and backward non-word recall tasks. Series length increased from 3 to 10 items.
2. Complex Span Task. Sentence repetition task.
3. Attention control. Q: word or non-word/male or female voice?
4. Processing speed. Speeded naming. Name the sizes, colours, and shapes of as many items as you can in 30 seconds.
Phonological processing

These WM profiles were invoked to see whether they could explain variation in the performance of a range of linguistic tasks.

(1) Segmental. Speeded segmental categorization with ABX.
(2) Suprasegmental (stress). Dupoux stress ‘deaf’ sequence recall task.
(3) Phonotactics. Speeded auditory Lexical Decision Task.

So, is there a correlation between behavior on the phonological tasks and general cognitive factors such as WM? Their hypothesis was that individuals with larger storage of a larger span or better quality processing will show what they call more robust phonological representations in the L2 which might result in “reduced inhibition from the L1 during processing” and hence more native-like performance on the phonological tasks.

However, there was not much correlation with WM, with r scores tending to be .3 or less. Attentional control did not correlate significantly with phonological measures. The only significant correlation with processing speed (r = -.427) was the stress task (which, as we will see in the stress deafness section, is problematic).

The main significance was between complex span (sentence repetition) and overall phonological score. They administered the three tasks for phonological knowledge: ABX segment discrimination, stress sequence repetition, and onset cluster lexical decision task. In the first stage of their analysis, they “computed a z-score for each participant for each task using the mean and standard deviation of the entire sample (N = 45).” A linear mixed effects model on these z-scores revealed there was no significant effect of task. In the second stage of their analysis, “an overall z-score was computed for each participant by averaging across the z-scores obtained for the test condition on each task to examine the effects for phonological development.” So, this ‘phonological score’ is an average of the results across three very diverse tasks.

Such tentative results leave much to be explored, and cannot be interpreted as supporting a claim that ‘advanced’ phonology is correlated with non-linguistic factors such as WM capacity.

In solidarity

As Scovel (1988) has pointed out, native listeners are very good at recognizing speakers who are not members of their social (read linguistic) group, which leads to another perspective. It has been well-documented in sociolinguistics (Labov, 1972) that accents can serve the function of marking group membership. There are times when speakers of non-prestige dialects are very committed to their variety because it marks their membership in a social group, a group they are proud to belong to. It is undeniable that there are political and social aspects to second language speech as well. In many cultures there are non-native speakers who are marginalized by the types of jobs they may take on, or their level of education or
literacy. Native speakers may pick up on certain linguistic cues or markers and may discriminate as a result. On the other hand, some speakers may choose in some venues to try to sound as native-like as possible, but in other venues, to maintain a recognizable L1 accent for solidarity reasons. In Norton’s (2000) terms, the non-native speaker must appropriate the L2 voice, and must be given license to speak by the mainstream culture. The relationships between these languages have a historical connection as well. Without going into details, imagine the range of tensions and attitudes one might find amongst Czech speakers of Russian, or Cree speakers of English, or English speakers of Japanese. My point is that each accent has its own political history. Due to such a relationship between speech and sociopolitical contexts, the consequences of trying to sound native-like in, say, Russian, would vary.

Global accent and comprehensibility

Jesney (2004) summarizes many factors which influence global accent. Many studies (e.g. Munro, Flege, & Mackay, 1996) clearly show age effects when it comes to global accent; early acquisition tends to result in more native-like speech. So, if global accent were our only measure then we would know that in order to achieve ‘advanced’ pronunciation then we would have to acquire the L2 at an early age. But, global accent is not the only measure. Let us contrast the perspectives of native-likeness and comprehensibility in the data from Munro, Flege, and Mackay (1996). When it came to judging whether the L2 subjects produced vowels in a native-like way, there were clear age-related differences. However, when the listeners were asked to identify the target vowel of the L2 speakers, the age effects disappeared.

The heritage speaker benchmark

Kupisch and Rothman (2016), while focusing on primarily morphosyntactic issues related to ultimate attainment in heritage speakers, raise a methodological point of interest to us here: who the heritage learners are compared to. In research following the native-likeness principle, the non-native speakers are compared to a monolingual control group. We would look at performance on a given measure and look to see whether it fell within the range of monolingual native speakers. Kupisch and Rothman look at heritage French and heritage Italian speakers in Germany. In other words, these are bilingual French/German or Italian/German speakers who grew up in minority-language households. Their French or Italian performance, however, is not compared to monolingual French or Italian speakers. Rather, their heritage performance is compared to the French or Italian of German heritage speakers living in France or Italy—in other words, bilingual German/French speakers who also grew up in minority-language households. And with this comparator group, they see that the heritage learners are performing in a way that is not significantly different from their bilingual control group.
The encapsulation of phonology

We should not expect a correlation between morphosyntactic ability and phonological ability. These are separate modules and develop separately. Archibald (1993) showed that there was no correlation between overall proficiency and the number of stress errors made by L2 subjects.

Bartning (2009) is central to the construct of the “Advanced Learner Variety.” She proposes three stages which make up the advanced profile: (1) advanced low stage, (2) advanced medium stage, and (3) advanced high stage. These are based on production data from two corpora: (i) the InterFra corpus (instructed learners) and (ii) the Lund corpus of naturalistic Swedish learners. The stages are delineated by morphosyntactic criteria (finiteness, subject, verb agreement, negation, noun phrase morphology, subordination, etc.). There is nothing to build on in determining what an advanced learner variety of phonology might consist of.

Similarly, Hyltenstam (2016) looks at the “characteristic features of advanced L2 proficiency” but focuses on “the grammatical, lexical, and discourse features.” There is still nothing to draw on in the domain of phonology. Towell and Dewaele (2005) showed that there was no correlation between linguistic knowledge and fluency after a few terms of L2 study. They suggest that the grammaticality judgement task (GJT) data they used to measure knowledge does not measure procedural knowledge. From this we can see no reason to expect phonological knowledge to be correlated with phonological fluency.

Thomson (2015) explores the construct of fluency and how it relates to such things as accentedness, intelligibility, and comprehensibility. He notes that this is a surprisingly understudied area. He concludes a summary and re-analysis of Derwing, Munro, and Thomson (2008) and Isaacs and Thomson (2013) by saying “fluency is most related to comprehensibility, somewhat related to accentedness, and apparently least related to intelligibility.”

All of this confirms that the notions of ‘advanced variety’ do not generalize easily to phonology (or at least to speech).

Near-nativeness

In the domain of phonology, there have been a few studies (Birdsong, 2007; Colantoni & Steele, 2006) which have looked at near-native speakers. Abrahamsson and Hyltenstam (2009) noted that 41 of their subjects (out of 195) were judged native-like by 6/10 judges, however, scrutiny of the competence of these 41 revealed subtle yet measurable differences from the native speakers. Thus, perceived native-likeness is different than actual native-likeness, which, they argue, is found in none of their L2 learners. In further probing of this population, Abrahamsson and Hyltenstam (2008) argue that high language aptitude (DeKeyser, 2000) is a prerequisite for achieving near-native-like ability. All late learners (who were judged as near-native-like) had high aptitude scores,
while early learners (who were judged as near-native-like) had aptitude scores that were normally distributed.

Moyer (1999) looked at 24 highly proficient English learners of German. She tested their pronunciation in four tasks: word list reading, sentence reading, paragraph reading, and free speech. Only one of the subjects passed for a native speaker in all four speech modes.

Birdsong (1992), building on the work of Coppieters (1987), probed the abilities of 22 English learners of French with a late AoA (all greater than 18). They were identified as being subjects who were often perceived as being native-like. He looked at such features as: vowel duration, voice onset time, appropriate liaison, and global accent rating. Of the 22 subjects, there were three late learners who fell within the native-like range on most elements. There were no individuals who evidenced only scattered pockets of native-likeness. And there were no domains where no one achieved mastery. This led Birdsong to state his universal learnability hypothesis: anything is learnable (by somebody); nothing is not learnable. Birdsong and Gertken (2013) remind us that we shouldn’t be trapped by the ‘missing-the-target’ mindset first identified by Klein (1998): Non-native-likeness ≠ failure.

Language switching

Another property of advanced knowledge in more than one language is that speakers switch between the languages; a phenomenon known as code-switching. In this section I will look at the issue of how a bilingual’s phonologies interact. The most obvious fact is that the phonologies do, in fact, interact. It is now widely accepted that code-switching is not evidence of limited language proficiency but, rather, is a natural byproduct of the architecture of the multilingual mind.

One domain where we can see this clearly is in lexical access. A range of work (e.g. Dijkstra, Grainger, & van Heuven, 1999) has argued for non-selective access to the bilingual lexicon by demonstrating bilingual interaction effects in a monolingual task. For example, interlingual homographs behave differently in lexical decision tasks. Interlingual homographs are words which share an orthographic form in two languages. For example:

English/French: pain, chat
Dutch/English: angel, glad

A closer look at such forms, though, quickly reveals considerable variation in phonological overlap, as Table 13.4 (adapted from Dijksta et al., 1999) reveals.

In this chart, S stands for semantics, O for orthography, and P for phonology. Thus, the SOP forms share semantics, orthography, and phonology; the SP forms share semantics and phonology, but not orthography, and so on. Of particular interest to us here are the interlingual homographs and the interlingual homophones. In a lexical decision task in which stimuli were controlled for word length and frequency, it was noted that interlingual homographs facilitated response time, while interlingual
homophones inhibited response time. Nakayama and Archibald (2005) replicated these results with an eye-tracking experiment. The interlingual homographs had shorter fixation times while the interlingual homophones had longer fixation times.

The role of phonology is also evident in the literature on translation equivalents (e.g. Nakayama, Sears, Hino, & Lupker, 2013; Voga & Grainger, 2007), which shows that cognate translation equivalents produce stronger priming effects than do non-cognate translation equivalents. For example:

Cognate: /remoN/ ‘lemon’
Non-cognate: /josei/ ‘woman’

In these cases, however, when activating the same lexical root, phonological overlap facilitates recognition (see Archibald, 2016, for further discussion of these issues within the framework of Distributed Morphology).

Thus, it is a fact of advanced phonology that the phonologies of the two languages will be interacting, and the high-proficiency subject has to control this. Let us then turn to the topic of executive control and language switching.

Some recent magnetoencephalography (MEG) studies (Blanco-Elorrieta & Pylkänennen, 2015, 2016) reveal differences in the switching mechanism in production and comprehension tasks. They compared the activation of a particular language (e.g. Arabic or English in a bilingual) triggered by a linguistic cue (e.g. orthography) versus a cultural cue (e.g. traditional dress in a photo). A task would be something like:

Speak English when you see “bdg”.
Speak Arabic when you see “әүә”.
Speak English when you see a man wearing a suit.
Speak Arabic when you see a man wearing a thawb (robe).

The orthographic script condition was activated more automatically and required more executive control to suppress. Their work shows that there is not a ‘single’
language switch as different neurological systems are implicated in a production switch compared to a comprehension switch. Furthermore, their work shows that there is a close relationship between the regions of interest in executive control of language and general cognition in production tasks but not so in comprehension tasks. This type of work has the potential of informing the debate on emergentist versus essentialist (i.e. domain-specific) approaches.

‘Advanced’ prosody

In the domain of what advanced perception might look like, let us turn to an issue which has received considerable attention in the SLA literature. This is the phenomenon which has come to be known as stress deafness. This literature predicts that speakers of certain languages are doomed to never be able to figure out the stress properties of the language(s) they are learning. The implication is that certain L1s will have difficulty achieving advanced levels of stress proficiency. Ultimately, I will argue that both the theoretical underpinnings and the interpretation of these studies should be reassessed.

Stress deafness

In a series of studies, Dupoux and colleagues have looked at the processing of stress in L2ers. Their conclusions are that certain L1 groups have difficulty discriminating contrastive stress on a variety of tasks, even into advanced proficiency levels. They catalogue languages in the following way with respect to their ability to ‘hear’ English stress:

- Totally deaf: French, Finnish, Hungarian
- Partially deaf: Polish
- Not deaf: Spanish

Their analysis is that Spanish has what they call lexical stress, while French, Finnish, Hungarian, and Polish do not (Peperkamp & Dupoux, 2002).

Previous research In 1997, Dupoux, Pallier, Sebastián, and Mehler argued that French L1 speakers were impaired in their ability to discriminate stimuli which differed only in the position of stress. This phenomenon, which they dubbed stress deafness, was found in tasks which used high phonetic variability and memory load but not in cognitively less demanding tasks such as AX discrimination. Dupoux, Sebastián-Galles, Navarrete, and Peperkamp (2008) probed whether this was a perceptual problem or whether the monolingual French speakers might simply lack a “metalinguistic representation of contrastive stress,” that would impair them in memory tasks. They concluded that “stress ‘deafness’ is better interpreted as a lasting processing problem resulting from the impossibility for French speakers to encode contrastive stress in their phonological representations.”
**Tasks**  Let me begin by looking at some of the tasks which were used (omitting some detail). A phonetic discrimination task was adopted in which minimal-pair nonce forms, which differed by either a single phone or a single stress position, were presented to L2ers.

E.g. [muki]/[muti]; [númi]/[numí].

Accuracy was reported to see if the L2ers could discriminate between pseudo-words which differed only in stress.

A second task was a sequence recall task which involved subjects listening to a string of five nonce forms (of the type given above) and having to recall the correct order of the string.

E.g. [númi] [númi] [númi] [numí] [numí].

They argue that the French subjects are significantly worse on these tasks than the Spanish subjects. Based on the (2008) paper, they argue that the L2ers have a processing deficit not a metalinguistic deficit.

Peperkamp, Vendelin, and Dupoux (2010) look at subjects from the L1s Finnish, Hungarian, and Polish, as well as French. Here is what they say about the L1 stress patterns:

- Finnish: initial
- Hungarian: initial
- Polish: penultimate

Using a variant of the sequence recall task from Dupoux, Peperkamp, and Sebastián-Galles (2001), the results showed that the Spanish (which has, like English, ‘variable’ stress placement) subjects were significantly different from the other groups. The Polish group was significantly different from the French, Finnish, and Hungarian groups. From this, they conclude that speakers of languages with “predictable stress” (i.e. French, Finnish, and Hungarian) exhibit strong stress deafness. Polish subjects exhibit an intermediate pattern between those of the French subjects and the Spanish subjects.

White, Muradás-Taylor, and Hellmuth (2016) present some of their own findings but also report on Taylor and Hellmuth (2012). Furthermore, they cite Correia, Butler, Vigário, and Frota (2015), who argue that Portuguese L1 subjects exhibit stress ‘deafness.’ These papers demonstrate ‘deafness’ from naïve listeners whose L1s have non-predictable stress. Remember this is the trait that was supposed to predict performance *success* on the Dupoux tasks. Both the Portuguese and the English subjects had difficulty with the tasks when the stimuli had no unstressed vowel reduction but only cues like pitch or duration to signal stress. Taylor and Hellmuth (2012) looked at English speakers (who were predicted to exhibit stress deafness) because they say that English stress is only “partially predictable, and must therefore be encoded in lexical...
representation.” However, they note that English subjects have trouble with L2 stress tasks. Kijak (2009) showed that English learners of Polish perform only slightly better than French subjects on a task of stress identification. All of this suggests that performance on the stress-deafness suite of tasks is affected by the phonetic realization of stress given that English speakers rely on vowel quality as a signal of stress/unstressed (Cutler & Pasveer, 2006).

White et al. (2016), building on the foundation of Taylor and Hellmuth (2012), looked at subjects who were not naïve but who had acquired advanced proficiency in L2 Spanish (a language without unstressed vowel reduction) to see whether what they term “cue-dependent” stress deafness is persistent. These advanced learners made very few errors (approximately 4%) on Spanish stimuli which varied only in pitch, intensity, and duration (i.e. no reduction). This demonstrates that new cues can be learned (consistent with Ylinen et al., 2009), and that that cue-dependent stress deafness is non-persistent.

**Stress deafness re-constructed**  In my view, many of these researchers are confounding variable with unpredictable. English stress is assigned in consultation with various parameters (quantity-sensitivity, grammatical category, etc.), but this does not entail that it is lexically stored. Properly conceived (see Halle & Vergnaud, 1987; Idsardi, 1992), stress is predictable and hence computed. Furthermore, one of the main differences between languages like English and languages like French is that in English stress is computed in the lexical phonology, and is tied in with many word-level phonological rules of English that depend on stress in some way: vowel reduction, tensing, laxing, and so on. One cannot get far in understanding English phonology without taking account of stress. French stress, on the other hand, does not serve as an environment for phonological rules of French. It is this difference that may be influencing the behavior of the French L1 subjects.

**What about representation?**

But let us consider these results, and the whole construct of stress deafness in the context of the results of Archibald (1993), who demonstrated that Hungarian and Polish subjects can reset their L1 parameters in the acquisition of L2 English stress. They are doing this on the basis of naturalistic exposure to English phonetic cues to stress. English, of course, has many redundant cues to stress, $F_0$, duration, and vowel quality. So, perhaps the success of the learners has to do with the nature of the cue not the nature of the L1 system leading to stress deafness.

Another possible perspective is that if we accept the processing difficulties that Dupoux et al. demonstrate and the fact that the Hungarian and Polish subjects can reset their parameters, then it is clear that the resetting is not being triggered by the input properties. This is an argument against input-driven models of acquisition; the learning is not just reflecting the input. Basically, it shows that they can learn how to compute the L2 stress, and the processing difficulty invoked by the high-load tasks is completely orthogonal to that question.
Linguistic assumptions  Now, let us probe the linguistic assumptions underlying the causal claims of the stress-deafness hypothesis. It could be argued their results do show difficulty in processing but not necessarily deficient representations (as is indicated perhaps by the term deafness). As the work of Lardiere (1998) demonstrates, problems in surface processing of a feature may arise even when there is accurate underlying representation.

There are some concerns with the theoretical assumptions about linguistic stress. Linguistic theory definitely has a part to play here. The stress-deafness work seems to assume a pre-theoretical, diacritic notion of what they call ‘lexical stress.’ Three of the languages they refer to as exhibiting stress deafness (French, Finnish, and Hungarian) all exhibit certain fixed stress patterns. French stresses toward the right edge of phonological strings while Finnish and Hungarian stress toward the left edge of strings (specifically word-initially). Polish, which exhibits partial deafness, has primarily penultimate stress. Spanish, which does not show deafness, has, like English, variable stress placement.

The problem with the causal connection proposed in this literature is that the stress placement of all of these languages is the product of a set of complex factors or principles (see Dresher & Kaye, 1990), not a primitive diacritic. Archibald (1992, 1993) has looked at the production and perception abilities of native speakers of Hungarian, Spanish, and Polish. He has argued that stress is not a monotonic primitive of linguistic representation but rather a complex system of knowledge which L2ers must, and evidently, can acquire.

Pater (1997) argued that French L1 subjects are able to acquire English trochaic feet but not the quantity sensitivity of English. Furthermore, Tremblay (2008) showed that only those learners who had acquired English stress settings could use stress as a tool for lexical access. Once again, this shows the importance of distinguishing between representation and process. Ultimately, what all this shows us is that the studies reveal something about processing but are not a direct window onto representation.

Directions for future research

The notion of ‘advanced’ phonology is an intuitively appealing construct which deserves exploration just as much as any other aspect of the ‘advanced’ profile of L2ers. However, as we have seen, formalizing and operationalizing the construct proves to be elusive. Much of the work has been done within a native-likeness paradigm. As linguists we must resist any judgmental aspect of such a comparison. Comparing grammars is like comparing elements on the periodic table. Helium is not better than hydrogen; it’s different. Canadian English is not better than British English; it’s different. Our attempts to describe and explain interlanguage phonologies, then, should not be viewed as placement on a linear scale from lacking to perfect but rather should be seen within the standard cognitive science paradigm of attempting to understand the fundamental properties of a mental system—as Gregg (2006) would say, the result of “naturalizing the mind.” That being said, the
field would benefit from heeding the call of researchers like Levis, Thomson, Munro, and Derwing by conducting more L2 phonology research within a comprehensibility paradigm. Future research must bring together the approaches of theoretical phonology, psycholinguistic experimentation, and pedagogic pronunciation. This is starting to happen in conferences such as Pronunciation in Second Language Learning and Teaching (PSLLT), and in some of the recent work of Cardoso (e.g. Cardoso & Collins, 2016).

Conclusion

Somewhat paradoxically, then, we must conclude this chapter by noting that there is no single hallmark of ‘advanced’ L2 phonology. We have seen that the perspectives taken as shaped by terms such as phonology, phonetics, intelligible, comprehensible, native-like, knowledge, and skill all influence the picture that emerges. So, even in the field of L2 speech, where there is both a mental and a motoric component, where age effects are uncontroversial, and where L2 accents are a daily occurrence, we must remind ourselves that comparing entities as complex as phonological grammars being implemented in real-world, real-time contexts is not straightforward. However, the positive picture that emerges reveals that no matter what the age, what the feature, or what the language, all second language learners can, indeed, advance.

NOTE

1 Based on Radical Underspecification theory (Archangeli, 1984), but I will not elaborate upon that here.

REFERENCES


14 Markedness and Advanced Development

FRED R. ECKMAN
University of Wisconsin-Milwaukee

Introduction

The purpose of this chapter is to argue for the relevance of typological markedness in characterizing advanced development in second language (L2) phonology. As limitations of space prevent an exhaustive account of markedness and L2 phonological development, the discussion will focus on representative studies carried out across the past 40 years or so on a number of different phonological structures.

Markedness is a linguistic construct developed for the purpose of making generalizations about the occurrence and non-occurrence of various structures across the world’s languages. This concept has been applied in the area of second language acquisition (SLA) as a measure of relative learning difficulty in acquiring certain structures of the target language (TL). In this chapter the converse will be argued, with the central claim being that, if markedness is defensible as a measure of learning difficulty in SLA, then the case can be made that, without any additional assumptions, markedness is also a viable measure of advanced development in L2 acquisition in general, and in L2 phonology in particular.

The structure of the chapter is as follows. The first section will describe three constructs that are crucial to the discussion, setting the basis for the claim that typological markedness is an evaluative measure for advanced competency in L2 phonology. The next section will outline several relevant studies. This will be followed by discussion of some issues surrounding markedness as an explanatory construct, and lead into the major claim of the chapter, namely, that typological markedness can also serve as an objective measure of L2 phonological development.
Background

Three concepts that are crucial to the discussion are the nature of L2 phonology, the construct of interlanguage (IL), and the notion of typological markedness, each of which will be taken up in this section.

L2 phonology

Most schools of thought in modern linguistics focus their analyses on spoken, rather than written language, defending the position that spoken language takes precedence over written language. As evidence for this view, linguists cite the facts that children learn to speak before they learn to write, that many languages of the world today exist without writing systems, and that writing is a relatively recent development in the evolution of human cultures. Thus, for most linguists, the spoken language reflects a living system that is part of the natural world, just as are magnetism, the rotation of the earth, and the movement of tectonic plates. Writing, on the other hand, represents a subsequent, human invention that has been imposed on the language. Thus, spoken language constitutes the general domain of study for the vast majority of modern linguistic schools of thought.

Many of the current approaches to linguistics view language as a mental system, termed a grammar, and maintain that this system has been acquired by humans, enabling them to produce and comprehend utterances of a language. And although there may be relatively little consensus across linguistic theories as to what these grammatical systems look like, virtually all schools of thought assume that some such system has been acquired. Phonology, in whatever way a grammar is viewed, is the subpart of that system that underlies the pronunciation patterns of the language.

The assumption that language can be insightfully conceived as some sort of mental system likely stems from the impact of generative grammar (Chomsky, 1965) on the field of linguistics. And although there is much variation and significant controversy in how such grammars are described, it is nonetheless fair to say that a native speaker’s ability to produce and comprehend utterances of a language is underlain by this system.

Interlanguage

Within this context, the question arises as to what kind of impact this thinking has on the characterization of second languages (L2), which are not learned natively. The hypothetical construct in second language acquisition theory that stems from the early work in generative grammar, and that could be considered one of the most important constructs in the overall research program in SLA theory, is the postulation of a “learner language,” which was proposed independently by three different scholars, Corder (1971), Nemser (1971), and Selinker (1972). The crux of all three proposals is that L2 learners internalize their own version of the target language (TL) grammar, termed an interlanguage.
The arguments for second language learners internalizing a system that represents their own version of the TL are identical to the arguments put forth by linguists for the postulation of native language grammars, viz., the occurrence of linguistic patterns produced by speakers of those languages. The linguist’s explanation for such patterns is the postulation of an underlying system. And so it is with L2 learners who produce systematic, albeit not necessarily target-like utterances.

The most important and interesting aspect of an IL is that it may contain structures that are independent of both the native language (NL) and the TL. Thus, although the IL may include aspects that have been transferred from the NL, and may also exhibit patterns that are part of the TL, and have presumably been learned on the basis of TL input, the IL may also reveal patterns that are independent of both the NL and the TL. Such regularities must be attributed to some system other than the NL or TL, namely, to the system that the learner has internalized, the IL. The occurrence of these L2 patterns that do not stem from the NL or the TL constitutes evidence for the independence of the IL system.

There are two significant implications of the concept of IL for L2 phonological theory. The first is that the research program in L2 phonology changed from one of trying to predict learning difficulty to one of testing hypotheses about the nature of IL phonologies. The second is that L2 learning now becomes characterized as the acquisition of an IL grammar, rather than the learning of the target language.

The concept of IL represents a breakthrough in that it changed the research paradigm in SLA from one attempting to characterize learning difficulty in L2 acquisition, to one of characterizing the nature of IL systems. With the concept of IL, it was possible to provide answers to questions, which, beforehand, could not even be formulated.

The final construct to be discussed in this section is typological markedness.

**Typological markedness**

The idea behind typological markedness is that some of the structures and distinctions that languages implement are not equal, polar opposites, but instead one member of the distinction can be shown to be privileged in that it has wider distribution, both across languages and within a language. Assigning the term *unmarked* to this privileged member is a way of giving it special status, and indicating that it is considered to be, in some well-defined way, more basic, more natural, and perhaps simpler than the less widely occurring member of the opposition, which is designated as being *marked*. The thinking behind **typological markedness** stems originally from the 1930’s Prague School of Linguistics (Jakobson, 1941; Trubetzkoy, 1939), and was developed further by Greenberg (1966) in his characterization of linguistic universals. Typological markedness capitalizes on the fact that the occurrence and non-occurrence of these structures divide languages into types, and allows the statement of universals in cases where a given linguistic structure may not be present in every language.
Markedness and Advanced Development

For example, not all languages exhibit a voice contrast in obstruents. However, if they do evidence this distinction, the syllable positions in which such a contrast occurs differ systematically across languages. Thus, some languages, such as English, have a voice contrast in obstruents both in syllable onset position and in syllable coda position ([to] *toe* versus [do] *dough*, and [ot] *oat* versus [od] *ode*, respectively). Other languages, such as Mandarin, do not have this contrast in either onsets or codas. Still other languages, such as Japanese, have this contrast only in onsets, but not in codas. However, there is no known language that exhibits this contrast only in syllable codas. These facts can be represented in Table 14.1 above, which shows that, with respect to the distribution of a voice contrast in obstruents, only three of the four logically possible language types are attested.

Table 14.1 shows that the distribution of a voice contrast in onsets and codas is not equal across languages; rather, a voice contrast in onsets is privileged in that it occurs more widely. Thus, a voice contrast in onsets is considered unmarked relative to a voice contrast in codas, and a voice contrast in codas is deemed as marked relative to such a contrast in onsets. These facts support the universal, implicational generalization in (1).

(1) The presence of a voice contrast in obstruents in codas implies the presence of this contrast in onsets, but not vice versa.

The unmarked structure can therefore be viewed as the default, in that a language does not necessarily have to exhibit a voice contrast in obstruents, but if the language does, it will always evidence the contrast in onsets.

Therefore, typological markedness can be defined as in (2) below.

(2) A structure X is typologically marked relative to another structure, Y, (and Y is typologically unmarked relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X. (Gundel, Houlihan, & Sanders, 1986, p. 108)

Under this view, typological markedness is a relationship holding between linguistic structures or representations across the world’s languages, and

---

**Table 14.1** Categorization of language types regarding the distribution of a voice contrast.

<table>
<thead>
<tr>
<th>Type</th>
<th>Contrast in onsets</th>
<th>Contrast in codas</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>No</td>
<td>Chinese</td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
<td>No</td>
<td>Japanese</td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>English</td>
</tr>
<tr>
<td>*D</td>
<td>No</td>
<td>Yes</td>
<td>None</td>
</tr>
</tbody>
</table>

---

For example, not all languages exhibit a voice contrast in obstruents. However, if they do evidence this distinction, the syllable positions in which such a contrast occurs differ systematically across languages. Thus, some languages, such as English, have a voice contrast in obstruents both in syllable onset position and in syllable coda position ([to] *toe* versus [do] *dough*, and [ot] *oat* versus [od] *ode*, respectively). Other languages, such as Mandarin, do not have this contrast in either onsets or codas. Still other languages, such as Japanese, have this contrast only in onsets, but not in codas. However, there is no known language that exhibits this contrast only in syllable codas. These facts can be represented in Table 14.1 above, which shows that, with respect to the distribution of a voice contrast in obstruents, only three of the four logically possible language types are attested.

Table 14.1 shows that the distribution of a voice contrast in onsets and codas is not equal across languages; rather, a voice contrast in onsets is privileged in that it occurs more widely. Thus, a voice contrast in onsets is considered unmarked relative to a voice contrast in codas, and a voice contrast in codas is deemed as marked relative to such a contrast in onsets. These facts support the universal, implicational generalization in (1).

(1) The presence of a voice contrast in obstruents in codas implies the presence of this contrast in onsets, but not vice versa.

The unmarked structure can therefore be viewed as the default, in that a language does not necessarily have to exhibit a voice contrast in obstruents, but if the language does, it will always evidence the contrast in onsets.

Therefore, typological markedness can be defined as in (2) below.

(2) A structure X is typologically marked relative to another structure, Y, (and Y is typologically unmarked relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X. (Gundel, Houlihan, & Sanders, 1986, p. 108)

Under this view, typological markedness is a relationship holding between linguistic structures or representations across the world’s languages, and
markedness is taken to be a property of the construction or representation itself. A given structure is not marked or unmarked absolutely, but is marked or unmarked relative to some other structure. A voice contrast in codas is marked relative to a voice contrast in onsets, and a voice contrast in onsets is unmarked with respect to a voice contrast in codas.

The concept of typological markedness is applicable not just to phonological contrasts but also to other linguistic distinctions and representations. For example, consider the contrast between singular and plural nouns. Cross-linguistic evidence indicates that singular and plural markings are not simply polar opposites, but that one can be shown to be privileged. English overtly indicates the difference between the singular and plural of nouns, as in *cats*, where the suffix *–s* signifies the plural. Not all languages make an overt distinction in number between singular and plural nouns. Some languages have no inflection for either the singular or the plural, as these languages use the same form for each, as in Mandarin where the word *mao* ‘cat’ is the form for both the singular and the plural. Other languages, such as Zulu, have a prefix to indicate the singular of a noun, and a different prefix to indicate the plural, as in umu-ntu *person* (singular) and aba-ntu *person* (plural). However, there are apparently no languages that distinguish number by using an inflection only on the singular noun. Thus, not all languages make an overt distinction between singular and plural nouns, but if a language does mark such a distinction, the marking will necessarily always be on the plural, and may also possibly occur on the singular. The occurrence of a singular morpheme in a language therefore implies the existence of a plural morpheme, but not vice versa, making the occurrence of the plural morpheme unmarked relative to the presence of a singular morpheme, and the singular morpheme marked relative to the plural.

These facts are depicted in Table 14.2 and lead to the universal generalization stated in (3).

(3) If a language uses an overt inflection for the singular, it also uses an overt inflection for the plural. (Croft, 2003, p. 89)

Over the decades since the inception of markedness, a number of different approaches to, and definitions of, this construct have been proposed, including the presence of overt morphological marking, as in the example of the English plural. Markedness has also been represented in terms of the inclusion of certain features

<table>
<thead>
<tr>
<th>Language type</th>
<th>Singular marker</th>
<th>Plural marker</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>No</td>
<td>Chinese</td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Yes</td>
<td>English</td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
<td>Zulu</td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>No</td>
<td>None</td>
</tr>
</tbody>
</table>
in the phonetic inventory of languages, and has also been defined in terms of the distribution of sounds within a language, where it is the unmarked member of an opposition that occurs when the contrast is neutralized (eliminated). This is seen in a number of Western languages (e.g., Catalan, German, Polish, and Russian) where the contrast in voicing is neutralized word-finally, where only voiceless obstruents (the unmarked member) occur. In other frameworks, markedness has been determined based on the relative amount of evidence required for acquisition by child-learners, with the idea that less marked structures require less evidence than more marked structures. And finally, a contrast or structure’s frequency of occurrence across the world’s languages has also been a criterion. Despite these different characterizations of markedness, the likely intent of the proposers is that they all point to the same set of structures and distinctions as being marked or unmarked. (See Battistella, 1990, and Moravcsik & Wirth, 1986, for further discussion.) The last notion mentioned above, distribution among the languages of the world, where there is a unidirectional, implicational relationship between the occurrence of the members of the opposition, the definition for typological markedness stated in (3), is the notion of markedness that has been used most widely by linguists in L2 phonology, and is the focus of this discussion.

Markedness as an evaluative measure in L2 phonology

Battistella (1990) in the title of his book labels markedness as the evaluative superstructure of language. Given that, in addition to exhibiting a number of similar properties, languages can also differ from each other in systematic ways, it becomes incumbent on linguists to characterize these similarities and differences in an interesting and non-ethnocentric way. Typological markedness provides a concept for stating universal generalizations about these similarities and differences (Greenberg, 1966), and for characterizing the notions of relative naturalness and simplicity across languages, as well as within the domain of language acquisition.²

Within this context, the purpose of this section is threefold. First, it will outline the use of typological markedness in L2 phonology as part of two hypotheses, the Markedness Differential Hypothesis (MDH) and the Similarity Differential Rate Hypothesis (SDRH). The former used markedness as a measure of difficulty and the latter invoked it as a factor in rate of acquisition. Second, the claim that ILs obey the same typological markedness relations as native languages will be discussed; this postulation was part of a general attempt to understand ILs in terms of native languages by hypothesizing that the universal generalizations formulated for primary languages held also for ILs. The third point to be made in this section pertains to the use of markedness relations being applied to linguistic phenomena in addition to contrasts and representations. Unidirectional implicational relationships have been formulated, in effect as markedness relationships, also to
include the distribution of segments, phonological environments, and speaking contexts. The argument will be made that markedness as an evaluative measure can be used to characterize advanced development.

**Markedness as learning difficulty in L2 phonology**

There are two hypotheses that have made use of typological markedness as a measure of learning difficulty (or acquisition rate) in the area of L2 phonology, the MDH and the SDRH.

The pre-eminent theory in SLA up to the middle of the twentieth century was the Contrastive Analysis Hypothesis (CAH), which sought to explain L2 learning difficulty on the basis of the differences between the learner’s NL and the TL. The CAH can be characterized by the quotation in (4) taken from Lado (1957, p. 2).

> (4) We assume that the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult.

Although the content and claims of the CAH can be attributed to the thinking of Lado and a number of his contemporaries, the name of the hypothesis is due to Wardhaugh (1970).

Over the years, empirical tests showed that the CAH was problematic in that its predictions were borne out in some cases, but not in others. One attempt to address some of the problems with the CAH was to incorporate typological markedness into the hypothesis as a measure of relative difficulty. Such a move yielded the MDH, stated in (5).

> (5) The Markedness Differential Hypothesis (Eckman, 1977, p. 321) The areas of difficulty that a language learner will have can be predicted such that
> (a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult;
> (b) The relative degree of difficulty of the areas of difference of target language which are more marked than the native language will correspond to the relative degree of markedness;
> (c) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.

According to this proposal, any given TL structure would be predicted to be more difficult if it was both different from the corresponding NL structure and was also more marked than that structure.

Whereas the CAH attempted to explain L2 learning difficulty only on the basis of differences between the NL and TL, the claim behind the MDH is that NL-TL differences are necessary for such an explanation, but they are not sufficient.
Therefore, the claim argues, one must incorporate into the hypothesis the concept of typological markedness as a measure of relative difficulty. Because the MDH claims that typological markedness must be incorporated into the CAH as a measure of relative difficulty in SLA, the former can be viewed as an extension of the latter, and thus completely programmatic with the CAH.

Over the past few decades there have been a number of studies addressing the claims of the MDH, supporting the claim that typological markedness is a reliable predictor of difficulty, that there are cases where the directionality of difficulty between the NL and TL involved in a language-contact situation follows the predictions of the MDH, and that the relative degree of difficulty corresponds to the relative degree of markedness. Limitations of space will allow a brief summary of only a few of the studies that sought to test the MDH.

One testable implication that follows immediately from the MDH is that not all NL-TL differences will cause systematic difficulty; different structures where there is no markedness relationship involved are not predicted to be difficult. For example, two learners from different NL backgrounds, each acquiring the same TL, are predicted not to experience equal difficulty with the same TL structure if there are different degrees of markedness between the TL construction and the corresponding structures in the respective NLs.

Likewise, two speakers of different languages, each acquiring the other’s language, may not experience equal difficulty with the same areas of difference because of distinct levels of markedness. For example, if native speakers of L_a and L_b are each acquiring the other’s language, then they may not experience equal difficulty with a given area of difference between the two languages. If the structures of L_b to be acquired by the speakers of L_a are more marked relative to corresponding structures in L_a than those facing the speaker of L_b, then more difficulty is predicted for the speaker of L_a than for the speaker of L_b. None of these predictions necessarily follows from the CAH, because unlike the MDH, the CAH does not incorporate an independent measure of difficulty.

To take a concrete example, the difference between German and English with respect to voice contrasts in syllable codas has been shown to cause more difficulty for German speakers learning English than for English speakers learning German. This asymmetry is predicted by the MDH because the German learners have to acquire a relatively more marked structure, a voice contrast in codas, compared to what the English-speaking learners of German have to acquire. Final devoicing in IL grammars has also been looked at from a markedness standpoint by Major and Faudree (1996) and by Yavas (1994).

Anderson (1987) showed that the order of acquisition and the relative difficulty involved in the learning of English syllable structure by native speakers of Amoy Chinese, Mandarin Chinese, and Egyptian Arabic followed the relative degree of markedness of the syllable type. The results supported the MDH in that the performance of the Chinese-speaking subjects was less target-like than that of the Arabic-speaking subjects on coda clusters, and the difference in performance correlated with the degree of markedness.
A final example is a study by Carlisle (1991) that reported evidence showing that learners’ performance on complex onsets in English by native speakers of Spanish could be explained only by invoking the markedness relationships that exist among the structures in question. The findings showed that the subjects modified the complex onsets by inserting an epenthetic vowel, and that the likelihood of a given onset type being altered was a function of the relative degree of markedness of two factors: the cluster in question and the preceding sounds. (See Hancin-Bhatt & Bhatt, 1997, for an analysis of L2 syllable structure within an Optimality Theory framework.)

The second hypothesis that invoked typological markedness principles in its explanation of facts about L2 phonology is the SDRH proposed by Major and Kim (1996). This hypothesis claims that phonological structures of the TL that are dissimilar to those of the NL will be acquired at a faster rate than those structures that are similar. The SDRH was later subsumed under the Ontogeny Phylogeny Model (OPM). This framework claimed that the learner’s rate of acquisition, rather than difficulty of learning, should be the focus, and it postulated that there is an explanatory role for typological markedness.

This can be exemplified using Major’s hypothetical illustration. If L1x is a representation in the NL, and L2x, is a structure in the TL that is similar to L1x, and L2x, is a TL structure that is dissimilar to L1x, then the SDRH predicts that L2xd will be acquired faster than L2x,. Moreover, if a markedness relationship is involved, then the rate is affected: faster if less marked, more slowly if more marked. Major and Kim (1996) supported the SDRH based on findings of research in which they studied Korean learners of English acquiring /dʒ/, a similar sound, and /z/, a dissimilar sound for speakers of Korean. The results showed that both the beginning and the advanced learners did better on the similar sound than on the dissimilar sound. However, a comparison of the rate of acquisition showed that the dissimilar sound was actually acquired at a faster rate than the similar sound.

To sum up this subsection, we have seen that typological markedness has played a significant role in the explanation of various facts about L2 phonology. L2 phonological patterns have been shown to obey the implicational generalizations that have been postulated on the basis of L1 grammars.

Having pointed out that markedness has been characterized in various ways by different phonological theories, but before moving on to the discussion of markedness as an evaluative measure, it is important to note that the phonological theory which most intrinsically and essentially incorporates markedness is Optimality Theory (OT; Prince & Smolensky, 1993). However, since OT is not the focus of this chapter, the discussion of it will necessarily be brief.

Grammars which are rule-based and grammars formulated within OT characterize the well-formed (grammatical) pronunciation patterns of a language differently. Rule-based grammars begin with a lexical form and execute a derivation in which all applicable rules are applied, and the well-formed output is thereby specified. An utterance is shown to be ungrammatical if the derivation violates at least one of the rules. OT grammars, on the other hand, do not consist of a set of rules,
but instead are made up of a set of universal constraints. A good way to think of the constraints is as criteria for well-formedness. The constraints can be violated, as no utterance is able to satisfy all of the conditions for well-formedness. Therefore, the constraints in an OT grammar must be ranked, and grammaticality within OT is not characterized in terms of whether an utterance violates one of the constraints, but rather in terms of an optimization procedure in which the grammatical utterances of a language are those that conform to the highest ranked constraints in the grammar.

Within OT, the universal constraints are divided into two categories, faithfulness and markedness. Faithfulness constraints value the utterance not deviating from the underlying form, and markedness constraints correspond to ease of pronunciation. This division aligns roughly with the notions of contrast and ease of articulation (Gundel et al., 1986). The important point within the context of this chapter is that OT incorporates markedness as a basic tenet of the theory. For more details and insights into OT and L2 phonology please see Hancin-Bhatt (2008).

**Markedness as an evaluative measure of IL grammars**

The Structural Conformity Hypothesis (SCH; Eckman, 1991), stated as in (6), claimed that principles of typological markedness hold *a fortiori* for second languages by asserting that all universal generalizations that are true for primary languages hold also for ILs.

(6) The Structural Conformity Hypothesis: The universal generalizations that hold for primary languages hold also for interlanguages. (p. 24)

The primary motivation for the SCH was to solve a problem with the MDH, namely, an L2 pattern in which the structures in question obeyed typological markedness principles, but the constructions were not in an area of difference between the NL and TL. Because the pattern did not arise in an area of NL-TL difference, it was not explained by the MDH, even though such patterns would seem to fall under the spirit, if not the letter, of the hypothesis. One way to address this problem was to eliminate NL-TL differences as a criterion for invoking markedness to explain the facts. Essentially, then, the SCH is the result of stripping NL-TL differences from the statement of the MDH. If we can assume that a learner will perform better on less marked structures relative to more marked structures, then the MDH can be seen as a special case of the SCH, namely, the case in which universal generalizations are obeyed by the IL in question, and the structures for which the generalizations hold are ones in which the NL and TL differ. In effect, the claim behind the SCH is an assertion about IL grammars rather than a statement about learning difficulty.

The kind of evidence that has been adduced in support of the SCH is an IL pattern that is neither NL-like nor TL-like, but nevertheless obeys the kinds of universal patterns found in the world’s languages. This kind of data has been reported in Carlisle (1997, 1998), Eckman (1991), and Eckman and Iverson (1994),
to cite just a few. Each of these studies considered the case of consonant clusters in onsets or codas, where the TL allowed both a greater number of clusters and more marked clusters than did the NL.

In Eckman (1991) the data were obtained using several elicitation tasks, including a free-conversation interview, from 11 ESL learners, four of whom were native speakers of Japanese and four of whom were native speakers of Korean. The remaining three were native speakers of Cantonese. The speakers’ performance was analyzed to determine which clusters were part of a subject’s IL grammar. These results were then used to test the SCH against several universal generalizations about the co-occurrence of consonant cluster types in a language. Out of over 500 such tests, the hypothesis was shown to hold in all but five cases.

The studies by Carlisle (1997, 1998) also tested the occurrence of consonant clusters, but in the IL grammars of Spanish-speaking learners of English. Carlisle’s findings supported the SCH in each case, and had the additional advantage of showing the operation of the SCH without imposing a criterial threshold on the data.

Finally, Eckman and Iverson (1994) analyzed English complex codas as produced in free conversation by native speakers of Japanese, Korean, and Cantonese. None of these languages allows complex codas. The findings showed that the learners had more errors on the more marked codas, with the consequence that the respective IL grammars had the more marked cluster type only if they also exhibited the less marked type.3

A common thread running through these studies supporting the SCH is that the IL grammars contain cluster types that are more complex (more marked) than those allowed by the NL, but not as complex as those required by the TL. In this respect, the IL patterns fall between the NL and TL, but always in a way that is in conformity with the universal generalizations (markedness relationships).

**Markedness as an evaluative measure of phonological environments and speaking contexts**

The definition of typological markedness as defined in (2) above can be extended beyond phonological segments to include a number of other aspects of L2 acquisition and performance. Thus, for example, Gundel et al. (1986) discussed the relationship of phonological distribution and markedness, making the point that, if two segments in a language alternate, the segment with the wider distribution is the unmarked segment.

This can be illustrated with the contrast between voiced and voiceless obstruents in German. German has a voicing contrast in obstruents in word-initial and word-medial position; however, this contrast is neutralized in word-final position in German, where only voiceless obstruents occur. Therefore, word-final position for a voice contrast is the most marked environment relative to word-medial and word-initial position. The same reasoning can be applied to other segments and positions of occurrence. In English, the velar nasal /ŋ/ occurs only in syllable coda
position, as in /sɪŋ/ *sing*, and not in onset position. Other languages, such as Vietnamese, have /ŋ/ in both onset and coda position. In this case, syllable onset position for /ŋ/ is the more marked environment, as it is apparently true that languages that have velar nasals in onset position also allow these segments in coda position (Maddieson, 1984).

The idea behind typological markedness can also be applied to speech styles. Thus, sociolinguistic research has shown that speakers talk at a slower rate and maintain more contrasts in careful, more formal speech than they do in a more casual style (Trudgill, 2011). This is true also for L2 speakers (Weinberger, 1987). Native speakers of English are much more likely to produce the word *opts* as [aps] in casual speech, reducing the final tri-segmental consonant cluster [pts] to the bi-segmental [ps], where the latter is less marked than the former.

Correlated with the relative markedness of speech styles in L2 phonological research is the kind of task that is used to elicit the L2 data. Thus, the kind of data that is obtained by having participants read a word list, or name pictures, or read a text corresponds to careful speech style. Such data obtained using this kind of methodology will represent the participant’s most careful pronunciation. Other task types, which attempt to elicit more casual, less formal speech style, include free conversations involving the participants. Contrasts, consonant clusters, and certain segments are much more likely to be realized in a careful speech style than in a casual one. Moreover, if a learner can produce these contrasts in casual style, then these speakers can do so in the formal style, but not necessarily vice versa.

Finally, it has been observed (Gundel et al., 1986) that marked forms are used when an utterance has been misheard by an interlocutor, and is repeated by the speaker. Thus, for example, in the case above where a speaker utters the word *opts* as [aps], and is asked for clarification or repetition because the word has not been understood, this constitutes a natural situation in which the speaker will produce the more marked coda cluster and utter [apts].

In a brief interlude, the discussion now turns to some issues surrounding typological markedness as an explanatory principle, before proceeding to the use of these principles as a measure of advanced development.

**Issues concerning typological markedness as an explanatory principle**

The above discussion has made the claim that typological markedness principles can provide an explanation for various facts about L2 phonology. However, some researchers in SLA have taken the position that markedness, in general, and the SCH, in particular, are not viable explanatory principles, but are instead simply facts to be explained. The purpose of this section is to show that principles of typological markedness make reliable predictions about phonological structures, and therefore can form at least part of the basis for the characterization of advancedness in L2 phonology.
The position that typological markedness generalizations are better viewed as facts rather than explanatory principles has been taken in at least three distinct forms in the literature. The first instance is found in Archibald (1998, p. 150), who takes the position that markedness principles present a description of the facts rather than an explanation.

(7) “My general assessment of this sort of typological universals approach to second language acquisition is that it provides an interesting description of the phenomena to be explained. I’m less sure of their [sic] status as an explanation of the observed facts.”

The second appears in Gass and Selinker (2001) and states that, rather than explaining facts about IL grammars, markedness principles simply “push the question back one step.”

(8) “… first, one must understand why a universal is a universal. It is not sufficient to state that second languages obey natural language constraints because that is the way languages are. This only pushes the problem of explanation back one step.”

The third instance of this type of claim about markedness not having any explanatory force comes from White (2007, p. 241), who states that explanations for SLA phenomena must be “… internal, individual and in part innately specified cognitive processes.”

(9) “… I assume that SLA is in fact a … branch of cognitive science. … A discernible trend … has been for increasing numbers of researchers and theorists … to focus their attention on SLA as an internal, individual, in part innately specified process.”

It will be argued that all of these criticisms have missed the point that there are levels of explanations, where the levels correspond to the generality of the laws or principles invoked. Therefore, let us take a brief look at the nature of scientific explanation in order to show that the accounts offered for facts about L2 phonology by markedness principles, in general, and the SCH, in particular, do in fact constitute explanations.

Scientists explain facts by subsuming them under general laws, where the fact to be explained is shown to be a specific instance of a more general phenomenon (Hempel & Oppenheim, 1989). For example, scientists explain the facts that when a mercury thermometer is immersed in hot water there is a temporary drop in the mercury column which is then followed by a swift rise by making reference to the laws regarding the expansion of materials when heated. The increase in temperature at first affects only the glass tube of the thermometer, causing it to expand and provide a larger space for the mercury, whose surface drops. Heat conduction then causes the temperature of the mercury to rise, which then expands faster than the
glass tube because of its larger coefficient of expansion. The behavior of the mercury in the thermometer is shown to be a particular case of a more general phenomenon, namely, the fact that different materials expand at different rates when heated.

The same kind of explanation was given to explain why, for example, the L2 learners studied in Eckman (1991), cited above, evinced patterns of consonant clusters in onset and coda positions, where these clusters were not TL-like, nor were they licensed in the learners’ NL. The SCH was invoked as a covering law in this case, claiming that the observed IL patterns that adhered to markedness generalizations about consonant clusters were a particular instance of a more general phenomenon, namely, IL grammars obeying universal generalizations.

There is an important relevant point that was first raised by Hempel and Oppenheim (1989), and that bears crucially on the above statements in (7), (8), and (9): the question of “why” can also be raised with respect to the general laws that are invoked as explanations. These laws can come to be regarded as facts to be explained, and would be explained if one could subsume them under generalizations which are more comprehensive; that is to say, they would be explained if it were possible to deduce them from some more-encompassing laws or principles.

Thus, there are levels of explanation, where “level” can be defined as the relative generality of the principles used in the explanation (Sanders, 1974). Any generalization from which it would be possible to derive the different coefficients of expansion, or which would subsume the SCH, would constitute a higher-level explanation for those generalizations. All empirical generalizations and hypotheses are, at the same time, a means for explaining lower-level generalizations, and also the object of explanation for higher-level generalizations (Sanders, 1974).

Some linguists may refer to principles of markedness as descriptions of the facts rather than as an explanation. And based on the above discussion, these linguists would be partly correct and partly incorrect. They would be right in saying that these principles constitute a description of the facts in the sense that lower-level generalizations become facts for higher-level generalizations to explain. But these linguists would be incorrect in asserting that these principles are not explanations, because they are, contrary to some claims, law-like statements which make testable predictions. Thus, a linguist who can propose higher-order linguistic or cognitive principles from which markedness generalizations are derivable is justified in referring to the SCH or to markedness relationships as facts, and not as explanations. However, it is unsound scientific reasoning to reject markedness generalizations as explanations unless one can propose higher-level generalizations under which these principles can be subsumed. In the absence of more general principles, it is scientifically imprudent to reject these generalizations as merely descriptions or out of step with current theorizing, because in so doing one would be left with no explanation at all.

With respect to the quotation in (7), what Archibald has missed here is that typological universals are laws that subsume phenomena under a generalization, make predictions, and thus constitute an explanation. Markedness principles, as is the case with all generalizations, can themselves be the target of explanation, and
can be explained if one can offer a more-encompassing principle. A similar point can be made with respect to the claim in (8). If one does not accept a universal generalization as an explanation for L2 facts because such a generalization “pushes the problem of explanation back one step,” one would never be able to accept any generalization as an explanation, because all generalizations, all hypotheses, push the question back another step by raising further questions. Indeed, such questioning is necessary if our level of understanding is to deepen. And finally, the statement in (9) also fails to recognize levels of explanation. The postulation of defensible cognitive principles from which markedness relationships could be derived would indeed advance the explanatory process. Until such principles can be proposed and defended, it would be premature and imprudent to reject markedness generalizations that can be empirically supported.

To summarize this subsection, markedness relations, which can serve as explanatory principles, may themselves also be the target of explanation. However, this does not vitiate their standing as an explanation, as all such principles are, at the same time, explanations as well as facts to be explained.

Markedness as an evaluative measure of advanced development

This brings the discussion to the central point of this chapter, viz., that typological markedness, without additional assumptions, can be invoked as an empirical measure of advanced development in L2 phonology. Moreover, it is argued that typological markedness is distinct from other scales of language proficiency, such as the Common European Framework of Reference for Languages (CEFR) in that markedness is a relationship that holds between linguistic structures or representations, whereas other extant measures of language proficiency represent a scale that refers to various pragmatic goals that the speaker is able to achieve or perform.

In the above discussion it was claimed that principles of typological markedness can be a measure of relative difficulty in L2 phonology, and as such could be used to explain the difficulty that learners experience with specific consonant clusters and with certain positions of contrast, and the fact that learners are prone to reduce some structures to other structures. Following Greenberg (1966), if typological markedness is included as part of linguistic theory, then invoking markedness as a measure of difficulty in L2 phonology is independently motivated. Likewise, markedness is equally available and independently justified as a measure of advanced development. In the latter case, markedness is simply the opposite end of the scale for advanced development compared to its use as a measure of relative simplicity. Thus, learners with advanced proficiency would be characterized as having internalized relatively marked IL grammars, capable of producing relatively marked structures, in corresponding marked environments, and in marked speaking contexts.

It is important to note here that other well-known scales of language proficiency have focused more on pragmatic and communicative aspects of language proficiency. For example, the United States Department of State uses five proficiency designations,
from 1 (the lowest) to 5 (the highest). In this scale, advanced proficiency is defined as “able to use the language fluently and accurately on all levels pertinent to professional needs.” Other proficiency scales, including the Interlanguage Round Table (ILR), the American Council of Teachers of Foreign Language (ACTFL), and the CEFR, use similar descriptors. For example, CEFR, which defines six levels of proficiency, characterizes Advanced Knowledge speakers as follows: “They are good at maintaining correspondence in a grammatically well composed manner. Can speak with well-structured sentences, with no hesitation or too much thinking. Writing is done quite professionally, with the controlled and proper use of grammatical patterns.”

On the other hand, typological markedness as an indicator of advanced proficiency would be an objective, empirical measure that is independent of other scales of language proficiency, and that would focus on phonological structures and representations, making it possible to triangulate the notion of advanced-level language proficiency.

In sum, typological markedness is an independently motivated construct that functions as a structural measure of advanced proficiency.

**Future directions**

It is virtually impossible to predict what the research focus of L2 phonologists will be in the near future, but it is possible to speculate on three possible areas involving new, or new uses of, technology.

The use of functional magnetic resonance imaging (fMRI) techniques may make it possible in the future to obtain neurological evidence on lexical representations, as well as other hypothetical constructs. The challenge would arise in ensuring that the granularity of both kinds of evidence is aligned. Eye-tracking methods have also been employed in studies on L2 acquisition to investigate attention to various cues in the input. And finally, the use of ultrasound techniques will make it possible to measure the tongue positions and movements of L2 participants to investigate whether there are any correlations with typological markedness, in terms of both production and perception. It has been determined that very subtle differences in tongue position can cause large distinctions in acoustic measurements.

**NOTES**

1 The claim here is that, however ‘advanced’ levels of L2 acquisition are defined, they would also be characterized as evincing more, rather than fewer, marked structures.
2 An anonymous reviewer of this chapter makes the point that some linguists may have claimed that if a language exhibits lesser complexity or markedness in some area of the grammar, that language will exhibit greater complexity or markedness in some other area(s). To the best of my knowledge, this has never been shown. But see Gil, Trudgill, and Sampson (2009) for some discussion on language complexity.
3 See Broselow, Chen, and Wang (1998) for an OT account of this phenomenon.
REFERENCES


Introduction

Few second language acquisition (SLA) researchers would disagree with the fundamental idea that second language (L2) learners improve their linguistic performance as a function of increased experience, practice, and interaction in the target language. For example, Flege (2016) pointed out that L2 learners’ speech performance continues to develop, restructure, and change over an extensive period of time (e.g. more than 10 years of length of residence [LoR]) before it becomes stable and plateaued, provided that their main language of communication is the L2 rather than their first language (L1). The ultimate attainment of these experienced L2 learners is typically referred to as ‘advanced’ performance in the sense that it is significantly different from that of inexperienced L2 learners (e.g. LoR<1 year; Flege, Bohn, & Jang, 1997; McAllister, Flege, & Piske, 2002; Piske, McKay, & Flege, 2001). Among these experienced learners, it is believed that those with special individual difference profiles, such as an early age of acquisition (AoA; Abrahamsson & Hyltenstam, 2009), high-level aptitude (Granena & Long, 2013), and motivation (Moyer, 2015), can reach near-native-like proficiency. In the extensive literature on native-likeness, which has investigated various dimensions of experienced and advanced L2 learners’ linguistic performance, the acquisition of native-like phonology seems to be particularly difficult compared to the acquisition of native-like vocabulary and grammar (e.g. Granena & Long, 2013).
From a theoretical perspective, examining advanced L2 speech attainment—and distinguishing it from that of inexperienced L2 learners and native speakers—makes a substantial contribution to theory building in SLA. Unlike L1 acquisition, which generally results in native-like proficiency across all linguistic domains of language (pronunciation, vocabulary, and grammar), L2 learners’ mental representations of language are built on their already-existing L1 systems (Best & Tyler, 2007; Flege, 2016). Thus, L2 systems are essentially different from those of monolinguals (Abrahamsson & Hyltenstam, 2009). Certain researchers (e.g. Cook, 2002) have claimed that any L2 phenomenon needs to be examined within non-native speakers themselves rather than via comparisons to a native speaker model. Focusing on beginner, intermediate, and advanced L2 learners versus native speakers, the growing number of empirical studies have examined the relatively difficult aspect of L2 speech learning—segmental and suprasegmental acquisition. Thus, it is high time that we summarize what comprises the process and product of advanced L2 speech learning and the extent to which L2 learners ultimately improve their segmental and suprasegmental proficiency relative to monolingual native speakers.

This chapter presents an overview of three fundamental questions regarding advanced L2 segmental and suprasegmental acquisition. These are: (i) how L2 learners improve and attain advanced L2 speech abilities through extensive and intensive immersion in a target-language-speaking environment (i.e. process); (ii) the extent to which their attained competence can approximate that of monolingual native speakers (i.e. product); and (iii) what kinds of learners likely achieve such high-level L2 speech performance according to a range of learner-extrinsic factors (age of acquisition and testing, length of residence, quality and quantity of input) and learner-intrinsic factors (cognitive abilities, aptitude, motivation) (i.e. individual differences). In this chapter, segmentals refer to individual consonantal and vocalic sounds in a target language, whereas suprasegmentals are defined as “the structure that organizes sound” (Cutler, Dahan, & van Donselaar, 1997, p. 142) subsuming a wide range of L2 phonetic phenomenon beyond segmentals, such as word stress, sentence stress, intonation, rhythm, and fluency (including speed, breakdown, and repair).

**Process of advanced-level L2 speech attainment**

In this subsection, we will first review previous literature investigating the developmental patterns found in the early and later stages of L2 speech learning (beginner → intermediate → advanced). Specifically, we will take a close look at how L2 learners develop new phonetic categories at the segmental and suprasegmental levels, and access them in perception and production, as they increase their amount of exposure to the target language.
Segmental learning

With respect to the L1 acquisition literature, infants initially use computational strategies to detect prosodic patterns and then start recognizing words (Jusczyk & Hohne, 1997). Importantly, after words are learned as whole phonological units, “the resulting increased vocabulary could result in sufficient pressure to fill in finer phonetic detail in the lexical representations in order to avoid confusion between similar sounding, known words” (Werker & Tees, 1999, p. 523). This form of L1 phonetic development continues to take place up to adolescence (Walley & Flege, 1999). According to many L2 speech researchers, similar phenomena are observed in L2 phonetic development (e.g. Best & Tyler, 2007, for Perceptual Assimilation Model-L2; Bundgaard-Nielsen, Best, & Tyler, 2011, for Vocab Model; Flege, 1995, for Speech Learning Model; Walley, 2007, for Lexical Restructuring Model).

Once L2 learners become immersed in the target language community in naturalistic settings (e.g. study-abroad, immigration), they tend to prioritize the fast, efficient, and robust recognition of, in particular, frequent and fundamental words (Nation & Webb, 2011); and emphasize producing them via intelligible and comprehensible pronunciation forms (Levis, 2005). Achieving basic L2 speaking proficiency is instrumental to successful social interaction with other native and non-native speakers in the early phases of L2 speech learning. According to corpus studies on L2 English, for example, these frequent and fundamental words comprise approximately 3,000–4,000 word families which cover 95% of vocabulary use in various daily communication settings (Adolphs & Schmitt, 2003); learning these can thus ensure a minimum and adequate understanding of various kinds of aural input (van Zeeland & Schmitt, 2012). As an attestation to this claim, Munro and Derwing (2008) longitudinally tracked the vowel production development of late 44 Mandarin and Slavic learners over their first year of residence in Canada. The results showed that much of their improvement was observed within the first six months of their immersion, especially for new vowel sounds within frequent words (e.g. [i] in ‘bit’ but not in ‘pit’).

As L2 learners increase their relevant experience and proficiency in the L2, however, they are forced to attend to fine-grained phonemic discrimination and identification (e.g. [i] vs. [ɪ], [p] vs. [b], [s] vs. [l]). This is because these learners need to accurately comprehend and produce speech by drawing on a large lexicon containing many confusing minimal pairs (e.g. ‘beat’ vs. ‘bit’, ‘pit’ vs. ‘bit’, ‘read’ vs. ‘lead’). At this stage, while L2 learners are sensitive primarily to word-sized units of L2 phonological information, they concurrently become more capable of detecting new sounds in the L2 input at a phonetic level. This phonetic-level restructuring is believed to lead L2 learners to create new phonetic categories and to generalize the newly acquired phonetic knowledge from familiar to unfamiliar lexical contexts.

Turning our attention to recent L2 vocabulary research, there are many suggestions as to how much vocabulary is needed for L2 learners to trigger such
phonetic-level restructuring (i.e. a transition from word to sound learning). When it comes to L2 English, for example, 6,000–7,000 word families are reported to account for 98% of lexical use in various discourse (e.g. Webb & Rodgers, 2009) and promote high-level comprehension of L2 aural texts (van Zeeland & Schmitt, 2012). Whereas knowing the most frequent 3,000–4,000 word families serves as a “minimum” requirement for beginner-to-intermediate L2 learners, knowing 6,000–7,000 word families could be considered as an “ideal” goal for advanced L2 learners (Schmitt, Cobb, Horst, & Schmitt, 2017).

Recently, Bundgaard-Nielsen et al. (2011) conducted a longitudinal research project with 31 Japanese learners of English with various vocabulary sizes (ranging between approximately 4,000 and 9,000 word families). Their longitudinal design is crucial, as the findings would directly relate to the causal relationship between vocabulary growth and the development of target-like phonological behavior. According to the results, the participants’ L2 English vowel identification patterns did not significantly change over a period of a few months. Rather, their vowel perception was strongly predicted by vocabulary size both at the beginning and at end of the project. Notably, the participants with knowledge in the band of 6,000–7,000 word families demonstrated relatively superior L2 vowel perception. On the whole, the findings here in turn suggest that L2 learners’ increased segmental awareness likely takes place especially when their vocabulary size goes beyond the minimum threshold for daily conversations (3,000–4,000 word families) and reaches the lexical standard for advanced L2 learners (6,000–7,000 word families; for the results of the production tests, see also Bundgaard-Nielsen, Best, Kroos, & Tyler, 2012).

Other L2 speech studies have also shown that more experienced L2 learners can attain more robust segmental competence as they become more capable of simultaneously processing word- and sound-sized units of L2 input. Owing to their relatively high awareness of individual consonants/vowels, these learners’ segmental performance may be less susceptible to the influence of lexical context. For example, Imai, Walley, and Flege (2005) found that experienced Spanish learners’ word recognition in L2 English was consistent regardless of the different amount of lexical density (i.e. whether L2 words are phonologically similar or dissimilar to other words1), although inexperienced Spanish learners’ perception abilities were significantly affected by lexical density across test items. Similarly, Flege, Frieda, Walley, and Randazza (1998) showed that highly experienced Spanish learners’ voice onset time (VOT) production in L2 English was not affected by any of the crucial lexical factors that are hypothesized to determine the quality of L2 pronunciation performance (e.g. familiarity, abstractness; see also Bradlow & Pisoni, 1999).

In sum, these studies equally indicate the presence of a strong vocabulary-phonology link—a phenomenon also attested in L1 acquisition—in the context of late L2 speech learning. As Bundgaard-Nielsen et al. (2011) pointed out, it seems reasonable to assume that “even late SLA is analogous to L1 acquisition in that the learning mechanisms employed in L1 acquisition may be available to the late L2 learner” (p. 459).
Suprasegmental learning

Similar to L2 segmental learning, cross-sectional and longitudinal studies alike have suggested that L2 learners continue to enhance different aspects of suprasegmental production—prosody (word and sentence stress, intonation) and fluency (breakdown, speed, repair)—at different learning rates over a prolonged period of residence in a target-language-speaking environment. Using a longitudinal research design, Derwing, Munro, Thomson, and Rossiter (2009) examined late Chinese and Slavic learners’ suprasegmental (prosody, fluency) performance in L2 English over their first two years of immersion in Canada. Much improvement was observed for breakdown fluency (the number of filled and unfilled pauses), but only within the first year of residence in Canada. In contrast, the participants continued to enhance the prosodic (vowel durations) and speed fluency (articulation and speech rate) dimensions of their L2 speech over the duration of the project.

The ‘quick’ and ‘immediate’ improvement in L2 fluency suggested here is consistent with the findings of recent study-abroad research. For instance, Mora and Valls-Ferrer (2012) showed that Spanish college students demonstrated significant improvement in L2 English fluency as operationalized by breakdown measures (the number and duration of pauses) and speed measures (speech and articulation rate) over one academic semester in the United States. For similar results with L2 learners of Spanish and French, see Towell, Hawkins, and Bazergui (1996) and Segalowitz and Freed (2004), respectively.

Conversely, improvement in L2 prosody—word and sentence stress, intonation—could be considered as ‘gradual’ and ‘slow’ in nature. Saito (2015a) conducted cross-sectional research to examine the spontaneous speech of Japanese learners of English with various LoR profiles in Canada (1 month to 13 years). According to the results of linguistically trained raters’ subjective judgments, the prosodic (word stress, intonation) and temporal (speech rate) qualities of the participants’ speech was significantly predicted by their LoR. In the context of late Korean learners of English, Trofimovich and Baker’s (2006) cross-sectional study also demonstrated the significant predictive power of LoR for L2 suprasegmental learning; specifically, the participants’ duration ratio between stressed and unstressed syllables (crucial for the assignment of adequate word stress) was significantly associated with their LoR (1–10 years).

Taken together, the findings here indicate that L2 learners may quickly improve, in particular, the breakdown fluency of their L2 speech (the number of filled/unfilled pauses) in the early phases of L2 speech learning. This is arguably because the acquisition of these temporal features is assumed to be relevant to the surface-level processing of L2 speech production (i.e. monitoring). Yet, a great deal of L2 experience may be needed for L2 learners to enhance, refine, and internalize other areas of L2 suprasegmental performance (i.e. word stress, intonation, speed fluency)—features which are thought to impact the automatization of phonetic encoding and articulation processes during L2 speech production.
Perception and production

As L2 learners develop segmental and suprasegmental representations as a function of increased experience and proficiency, they develop increasingly fluid access to these systems in order to comprehend and produce an L2 in various processing modes (perception, controlled and spontaneous production). Slightly different proposals have been developed to explain the interrelationships between perception and production. For example, certain researchers have assumed that phonetic and phonological representations could be stored in the brain based on perception (how they hear temporal and spectral dimensions of new sounds; e.g. Flege, 1995, 2009, 2016). Others have stated that these representations have a more articulatory basis (how they produce new sounds by using the tongue, lips, and jaw; Best & Tyler, 2007). Although there is some evidence that L2 learners’ perception abilities may precede their production abilities (perception-first; e.g. Flege et al., 1997) or vice versa (e.g. Sheldon & Strange, 1982), these researchers have at least agreed with the fundamental idea that L2 learners simultaneously improve their perception and production abilities due to their interconnected nature (Bundgaard-Nielsen et al., 2012).

Whereas L2 learners’ processing of new L2 sounds in the perception phase is relatively automatic (Flege, 1993), it remains open to debate to what degree their productive use of new L2 sounds is linked to their developing L2 phonology system. According to Major’s (2008) Ontogeny Phylogeny Model, L2 learners initially tend to substitute their own L1 counterparts for new L2 sounds. With increasing awareness of new L2 sounds, however, interlanguage pronunciation performance begins to demonstrate some universal characteristics, regardless of L1 background. For instance, universally, L2 learners tend to make more pronunciation errors in free speech tasks than in formal word reading tasks. Rau, Chang, and Tarone (2009) demonstrated that Chinese learners of English mispronounced [θ] more frequently in a picture description task than in word and sentence reading tasks. Similarly, Saito and Wu (2014) also found that Cantonese learners of Mandarin Chinese produced more tone errors when their speech was elicited via word and sentence reading than when via picture description. This task effect in performance was believed to arise from the increased demands on linguistic processing, due to the lack of substantial planning time for the picture description tasks, compared to the more controlled reading tasks.

Such task effects (controlled production > spontaneous production) can be explained by the cognitive psychology literature, which posits that L2 learners produce output via a gradual transition from effortful to automatic use of newly acquired L2 knowledge (e.g. DeKeyser, 2007). Very importantly, this line of SLA research has also demonstrated that L2 learners can carefully produce language (and even speed up retrieval) without fully integrating linguistic knowledge into their mental representations (for a detailed discussion, see Segalowitz, 2003). In the context of L2 speech production, L2 learners can consciously activate relevant articulatory gestures in order to produce these sounds at a controlled-speech level...
(where they are given a sufficient amount of time to access their explicit knowledge).

The developmental patterns of the early and later phases of L2 speech learning are summarized in Table 15.1.

**Table 15.1** Summary of L2 speech learning (early phase → later phase).

<table>
<thead>
<tr>
<th></th>
<th>Characteristics of early phase</th>
<th>Characteristics of ultimate attainment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segmentals</strong></td>
<td>Intelligible pronunciation of frequent and familiar words</td>
<td>More native-like and refined pronunciation of segmentals regardless of lexical context</td>
</tr>
<tr>
<td><strong>Suprasegmentals</strong></td>
<td>Much improvement as measured via breakdown fluency measures (pause frequency)</td>
<td>Continuous improvement in prosody and speed fluency (articulation rate)</td>
</tr>
<tr>
<td><strong>Processing</strong></td>
<td>Perception may precede production or vice versa. Spontaneous production is lacking in the early stages.</td>
<td>Perception matches production. Spontaneous production ability balances out over time.</td>
</tr>
</tbody>
</table>

(e.g. word and sentence reading), where they are given a sufficient amount of time to access their explicit knowledge.

The developmental patterns of the early and later phases of L2 speech learning are summarized in Table 15.1.

**Product of advanced-level L2 speech attainment**

Many experienced learners (e.g. LoR > 10 years) ultimately reach relatively high-level L2 performance compared to inexperienced learners (e.g. LoR < 1 year). Yet, it remains controversial the extent to which such high-level L2 competence can be close to native speaker ability. In this subsection, we will review previous empirical studies comparing linguistic abilities of three crucial groups: (i) experienced and advanced L2 learners, (ii) inexperienced L2 learners, and (iii) monolingual native speakers. As such, I will aim to provide a rough index of the potential (vs. inexperienced L2 learners) and limitations (vs. native-speaker baseline) of advanced L2 speech learning.

**Segmental attainment**

In the early L2 speech literature, there is ample evidence that bilinguals’ speech performance is essentially and fundamentally different from that of monolinguals
(e.g. Sebastián-Gallés & Soto-Faraco, 1999). More recently, L2 researchers have extensively examined whether and to what degree experienced and advanced L2 learners can ultimately improve their speech performance to reach a “near-native-like” level. Near-native-likeness, defined as “levels of nonnativeness that are non-perceivable in normal, everyday language use” (Abrahamsson & Hyltenstam, 2009, p. 294), has been the focus of much empirical investigation. Notably, the concept of near-native-likeness is different from native-likeness, where non-native speakers have become not only perceptible, but also linguistically equivalent with monolingual native speakers.

One well-known theoretical model especially related to the early phase of L2 speech learning is Flege’s Speech Learning Model (SLM), which specifically highlights the “ultimate” attainment and “near-native-likeness” of L2 segmental acquisition (Flege, 1995, 2009, 2016). According to the SLM, when L2 learners identify a new L2 sound as perceptually distant from the closest L1 counterpart, they create a new phonetic category and strive to maintain its phonetic contrast relative to the L1 counterpart. Building on the SLM, Flege and his colleagues (e.g. McAllister et al., 2002) also proposed that such learning difficulty can be further predicted by the extent to which target acoustic (and articulatory) features are shared between the L1 and L2 phonetic inventories at a spectral (e.g. differences in formant frequencies) and temporal (e.g. differences in segmental and formant transition length) level.

Following this line of thought, the amount of learning difficulty can be predicted in conjunction with cross-linguistic L1/L2 distance:

- The incidence of near-native-likeness can be high for the adjustment of acoustic and articulatory phonetic features already present in both the L1 and the L2.
- Few learners can achieve the near-native-like proficiency required for establishing and acquiring complexly new features.

Since the fundamental concept of the SLM draws on the influence of the L1 on L2 development, it could also refer to the role of age of acquisition (the first intensive exposure to target language) in advanced-level SLA:

- Compared to late bilinguals whose L1 systems are fully developed and interfere with L2 speech learning to a great degree, early bilinguals, who are to start acquiring L2 while developing their L1, may exhibit more advanced L2 attainment, as the process and product of early SLA may be less susceptible to L1 influence.

Munro, Flege, and MacKay (1996) analyzed the English vowel production of 240 experienced Italian learners who widely varied in their AoA profiles in Canada (1–23 years). The results of listener judgments and acoustic analyses demonstrated that L2 English vowel production of many early learners (AoA < 7–8 years) was perceptibly indistinguishable from the monolingual native speaker baseline. Such near-native-like performance was not observed among late learners (AoA = 10–21 years).
A number of studies have delved into the role of L1 phonetic systems in the acquisition of L2 vowels by late learners (AoA > 16 years). Flege et al. (1997) explored the vowel perception and production of experienced (LoR = 7.3 years) and inexperienced (LoR = 0.7 years) L2 learners of English with three different L1 backgrounds (German, Korean, and Mandarin). Whereas the results noted that experienced L2 learners generally demonstrated better performance than inexperienced L2 learners, few achieved native-like performance for any vowels. Importantly, the experienced learners’ attainment was uniquely related to their L1 phonetic categories (e.g. German learners overused the durational cue to differentiate the English [ɛ]–[æ] distinction, as the temporal information is critical for the L1 counterparts (German [ɛ]) vs. [ɛː]).

Similarly, Schmid, Gilbers, and Nota (2014) acoustically examined late experienced Dutch learners’ vowel production of the phonemic contrast between [ɛ] (‘pet’) and [æ] (‘trap’). The results showed that a very small portion of the participants attained native-like vowel performance, suggesting that “late bilinguals, even at very advanced levels of proficiency, may have difficulties with vowel categories that are not instantiated by their L1” (p. 149).

To probe the detailed nature of L1 influence on late L2 vowel acquisition, McAllister et al. (2002) examined how experienced L2 learners of Swedish (LoR = 17–24 years) with three different backgrounds (English, Estonian, and Spanish) differentially perceived and produced vowel quantity. The results showed that, whereas Estonian learners’ performance attained near-native-like proficiency, English learners significantly differed from native Swedish controls and outperformed Spanish learners. McAllister et al. attributed the different ratio of ultimate attainment to the relative importance of vowel quantity in the participants’ L1 phonetic systems (Estonian > English > Spanish).

With respect to consonants, much research has exclusively focused on the acquisition of the non-native contrast between English [ɹ] and [l] by Japanese learners, as it is hypothesized to be one of the most difficult instances in L2 speech learning (for a review, see Bradlow, 2008). Since Japanese has only one alveolar tap [ɾ], which acoustically lies somewhere between English [ɹ], [l], and [d] (Hattori & Iverson, 2009), Japanese learners of English tend to make much effort to create new phonetic categories for both English sounds, resulting in much learning difficulty in perception and production.

Early L2 children likely demonstrate similar learning trajectories to those of monolingual children (phonetic development throughout childhood), and are able to attain near-native-like proficiency (Idemaru & Holt, 2013). In contrast, few late learners (AoA > 16 years) attain such high-level L2 speech performance. In the previous literature, even experienced late Japanese learners (LoR > 10 years) perceived the English [ɹ]–[l] contrast with 80–90% accuracy (e.g. Ingvalson, Holt, & McClelland, 2012) and produced highly intelligible English [ɹ] and [l] pronunciation forms (e.g. Flege, Takagi, & Mann, 1995), although their perception and production were significantly different from monolingual native speakers’ baseline (Flege, Takagi, & Mann, 1996).
Saito (2013; Saito & Brajot, 2013) acoustically analyzed English [ɹ] production in the context of 150+ Japanese learners with varied LoR in Canada (1 month to 40 years). It was found that most of the participants with more than 1 year of LoR experience had already demonstrated native-like proficiency in terms of two articulatory parameters of English [ɹ]— tongue retraction (i.e. lower second formant) and segmental/formant transition lengthening (>50 ms), both of which are available in the Japanese phonetic system (e.g. [y] vs. [w] for the retraction parameter; [i] vs. [i:] for the lengthening parameter). Several Japanese learners—especially those with extensive LoR experience (>6 years)—successfully acquired the new articulatory parameters: the labial, palatal, and pharyngeal constrictions (i.e. lower third formant). However, their production fell short of native speaker baselines.

Another well-researched topic in the previous native-likeness literature is the acquisition of VOT in word-initial stops (e.g. [p]-[b]). Unlike Japanese learners’ English [ɹ] and [l] acquisition, it is intriguing that many L2 learners (including even late L2 learners) have been reported to attain native-like proficiency in terms of this feature (e.g. Abrahamsson, 2012; Schmid et al, 2014; but see Chapter 17 for details of VOT in Advanced SLA). This is arguably because L2 learners tend to show less difficulty in acquiring new temporal (e.g. adjusting VOT between L1 and L2) compared to spectral (e.g. enhancing sensitivity to third formants to differentiate English [ɹ]-[l]) information (see also Bohn & Flege, 1992).

However, it is also important to point out that L2 learners do not necessarily master all areas of L2 temporal features according to the availability of the relevant phonetic cues in L1 inventories. Using acoustic analysis, Baker (2011) closely examined two acoustic features of late Korean learners’ ‘word-final’ stop production in English—vowel and closure duration. The results showed that most of the participants mastered native-like vowel duration—a phonetic cue that may be often used to differentiate short and long L1 Korean vowels. However, they did not reach native-like-level closure duration—a new phonetic cue that is not accessed in producing L1 Korean sounds.

**Suprasegmental attainment**

Though fewer in number, certain studies have examined the ultimate attainment and near-native-likeness of L2 suprasegmental acquisition. Guion (2005) examined highly experienced Korean learners’ (LoR>10 years) knowledge and production of L2 English word stress, finding that even early learners’ performance (AoA=1–6 years) was significantly different from a monolingual native speaker baseline. Guion argued that the lack of native-likeness can be ascribed to the cross-linguistic difference: Word-level prominence in L2 English does not exist in the L1 Korean phonetic system. In the context of a similar population (early English-Korean bilinguals), Trofimovich and Baker (2007) investigated a range of suprasegmental features (stress timing, intonation, speech rate, pause frequency and duration) in L2 English speech. The results showed (i) that early learners attained near-native-like proficiency in prosody-based suprasegmentals (stress timing and
intonation); and (ii) that the fluency-based suprasegmental measure (speech rate) was significantly different from a monolingual native speaker baseline, regardless of the participants’ AoA profiles.

As for fluency, Lahmann, Steinkrauss, and Schmid (2016) recently conducted a large-scale study with 102 experienced German learners of English (LoR > 20 years), investigating the acoustic characteristics of various dimensions of L2 fluency—mean number of unfilled pauses (breakdown), mean syllable duration (speed), and mean number of self-corrections (repair). Because Lahmann et al. (2016) did not include a native baseline group, the study does not allow us to speculate the native-likeness of the participants’ advanced fluency performance. Saito (2015b) demonstrated that even highly experienced late Japanese-English bilinguals’ (LoR > 10 years) perceived fluency was substantially different from a native speaker baseline.

The comprehensive review above suggests several broad, tentative conclusions about the ultimate attainment and near-native-likeness of L2 segmental and suprasegmental acquisition. These conclusions are summarized in Table 15.2.

### Individual differences in advanced-level L2 speech attainment

Second language speech learning is subject to a great deal of individual variability, indicating that certain (but not all) L2 learners can make the most of their relevant L2 experience and attain advanced-level performance. This subsection will provide a systematic review of those learner variables thought to play a crucial role in determining the incidence of advanced- and near-native-level L2 segmental and suprasegmental proficiency attainment. Such affecting factors include L2 learners’ age (age of acquisition and testing), experience (quality and quantity of input), cognition (attentional and inhibition control, L2 learning aptitude), and conation (learner motivation and willingness to communicate).

#### Age

Flege, Munro, and MacKay (1995) explored the pronunciation attainment (measured via perceived accentedness) of 240 long-term Italian residents in Canada (LoR > 20 years). The results found that much variance (i.e. 59%) in their foreign accentedness could be explained by age of acquisition—the first intensive exposure to target language. Such findings have been replicated in many other L1/L2 contexts (e.g. Abrahamsson & Hyltenstam, 2009; Flege, Yeni-Komshian, & Liu, 1999). L2 researchers have put forward a number of accounts for explaining such strong age effects on L2 pronunciation attainment.

First, given that L1 and L2 acquisition are assumed to take place in the same phonetic space in the brain, an earlier AoA would seem to indicate a weaker influence of the L1 on L2 speech learning (Flege, 1995). Second, early bilinguals tend to have many opportunities to use the L2, especially in the schooling context,
Table 15.2  Summary of the ultimate attainment and near-native-likeness of experienced L2 learners’ segmental and suprasegmental performance.

<table>
<thead>
<tr>
<th>Target features</th>
<th>Incidence of near-native-likeness</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmentals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readjusting acoustic and articulatory features present in the L1</td>
<td>All learners attain high-level proficiency given ample input.</td>
<td>• Tongue retraction in the acquisition of English /l/ by Japanese learners &lt;br&gt;• Durational differences in various L2 segmental acquisition (e.g. VOT in the acquisition of English stops by Spanish learners) &lt;br&gt;• Vowel duration for the acquisition of English word-final stops by Korean learners &lt;br&gt;• Spectral difference in L2 vowel acquisition &lt;br&gt;• Simultaneous labial, alveolar, and pharyngeal constrictions in the acquisition of English /l/ by L1 Japanese learners &lt;br&gt;• Consonant closure for the acquisition of English word-final stops by Korean learners</td>
</tr>
<tr>
<td>Acquiring acoustic and articulatory features absent in the L1</td>
<td>The acquisition of high-level proficiency requires early age of acquisition.</td>
<td></td>
</tr>
<tr>
<td>Suprasegmentals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prosody</td>
<td>The acquisition of high-level proficiency requires early age of acquisition.</td>
<td>• Word/sentence stress and intonation</td>
</tr>
<tr>
<td>Breakdown and repair fluency</td>
<td>All learners attain high-level proficiency given ample input.</td>
<td>• Pause and reformulation frequency</td>
</tr>
<tr>
<td>Speed fluency</td>
<td>Attaining high-level proficiency is extremely difficult regardless of age of acquisition.</td>
<td>• Articulation and speech rate</td>
</tr>
</tbody>
</table>
whereas certain late bilinguals can choose to interact with speakers of their L1 community rather than with native speakers of the target language (Flege & MacKay, 2004). Third, L2 learners gradually lose their access to the implicit language acquisition device by which they pick up an L2 through mere exposure in an automatic and effortless fashion (Abrahamsson, 2012). When L2 learners are exposed to the L2 after puberty, their language learning process could be characterized as rather explicit, effortful, and incomplete—a concept termed the Critical Period Hypothesis. Fourth, many cognitive functions (which are assumed to be instrumental to successful language acquisition) begin to decrease after the late 20s, such as working memory, attentional and inhibitory control, and speech sound processing (Birdsong, 2014).

Interestingly, such AoA effects have not been reported when it comes to L2 fluency attainment. Saito (2015b) found that late experienced Japanese learners of English (LoR > 10 years) successfully attained high-level perceived fluency regardless of their age of acquisition. Lahmann et al. (2016) further illustrated that the breakdown (the number of unfilled pauses) and repair (the number of reformulations) dimensions of L2 fluency could be related to L2 learners’ age of testing (when they took the tests) rather than age of acquisition (when they arrived in an L2-speaking country). The findings here suggest that humans gradually lose their ‘retrieval’ abilities for their linguistic competence as a function of increased age—a similar phenomenon in L1 acquisition, where aging and dysfluency are strongly correlated with each other (e.g. Burke & Shafto, 2004, 2008). In other words, whereas the quality of phonological representations (segmental and suprasegmental accuracy) is greatly determined by the timing of arrival in an L2 speaking environment (i.e. age effects), L2 learners’ speech production and monitoring abilities (reflecting on breakdown and repair fluency) correlate with the timing of speech data collection (i.e. aging effects).

**Quality and quantity of input**

According to Flege’s (2009) oft-cited definition, input is termed as “L2 vocal utterances the learner has heard and comprehended, including his own, regardless of whether these utterances have been produced correctly by L2 native speakers, or incorrectly by other non-native speakers of the L2” (p. 175). It has been shown that L2 learners likely exhibit continuous improvement in L2 speech performance in relation to the amount of input via interaction with native and other non-native speakers. However, little learning occurs when L2 learners lack such willingness to communicate, regardless of how long they stay in an L2 speaking environment (Derwing & Munro, 2013; Flege & MacKay, 2004).

However, what remains unclear and controversial is the role of different types of input in L2 speech learning (Flege, 2009). For instance, what kind of input is most optimal for speech learning? To what extent does such optimal input facilitate SLA? To date, L2 speech researchers have made tremendous efforts to track not only the quantity but also the quality of L2 input received over years of immersion. However, these studies have exclusively focused on immigrants and
measured their L2 experience in a retrospective way (e.g. self-reports of frequency of L1 and L2 use). As a remedy, other L2 speech researchers have conducted intensive phonetic training studies, controlling the quality and quantity of L2 input and giving precise descriptions of how different types of L2 input enhance the nature of phonetic and phonological representations.

High variability input Logan and his colleagues (e.g. Logan, Lively, & Pisoni, 1991) conducted a series of studies to test the effects of High Variability Phonetic Training (HVPT; i.e. intensive exposure to natural L2 tokens produced by many talkers) on the acquisition of the /ɹ/-/l/ contrast by Japanese learners of English. The results showed that (i) learners who received HVPT (30 minutes × 15 sessions = 7.5 hours) could identify the English [ɹ]-[l] contrast at post-test sessions more successfully than at pre-test session; (ii) the gains resulting from HVST were generalized to novel lexical contexts and new talkers; and (iii) improvement was sustainable for six months without any additional training. Finally, the extended HVPT (30 minutes × 45 sessions = 22.5 hours) allowed learners to achieve perfect generalization (Yamada, 1995) and transfer the learning to the production domains (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997). Similar findings have been reported in the case of L2 vowel acquisition (Thomson & Goswami, 2010) and L2 suprasegmental acquisition (e.g. Wang, Jongman, & Sereno, 2003).

Acoustically enhanced input Given that exaggerated, acoustically enhanced speech stimuli are hypothesized to be facilitative of discriminating non-native contrasts as shown in the L1 acquisition literature (Kuhl, 2000), other researchers have investigated the acoustic enhancement technique. In McCandliss, Fiez, Protopapas, Conway, and McClelland’s (2002) training study, Japanese learners first received synthesized tokens where the acoustic difference between English [ɹ] and English [l] was enhanced, and gradually moved on to listening to natural speech tokens. The results showed that acoustically enhanced input significantly helped the Japanese learners improve their English [ɹ]-[l] perception abilities.

Socially interactive input Many L1/L2 researchers have examined the role of social interaction in L2 speech learning. In the bilingual infant literature, Kuhl and her colleagues have empirically shown that learners tend to acquire new foreign sounds through reciprocal interaction with actual persons compared to mere exposure to televised or audio-only input. This is because live situations arouse motivation and attention toward provided speech information via eye gazing and gestures to objects of reference (Conboy & Kuhl, 2011). Further, sharing the communicative intentions of others helps learners code word and phonetic information from ongoing speech (Kuhl, Tsao, & Liu, 2003). In the case of L2 speech learning, recent studies have provided some evidence that adult learners who receive input through meaning-oriented interaction with teachers tend to show quicker and more robust improvement with even a limited amount of training time (several hours), compared to those who practice target sounds via computer in a
decontextualized manner (Lee & Lyster, 2016; Saito & Wu, 2014). Taken together, these studies (with both early and late bilinguals) suggest that social interaction might also play a key role in facilitating adult L2 speech learning, arguably because the presence of interlocutors provides learners with more immediate, engaging, and motivating conditions for learning compared to mere exposure to computerized treatment.

**Cognition**

A growing number of researchers have begun to investigate whether, to what degree, and why L2 learners with special cognitive abilities (e.g. working memory, processing speed, inhibition control) and language learning aptitude (e.g. associative and rote memory, phonemic decoding) can achieve high-level speech perception and production abilities. For example, it has been shown that certain learners with strong working memories and processing speed can potentially optimize the L2 learning experience, as they have the capacity to store and access a great deal of received input efficiently and effectively (Darcy, Park, & Yang, 2015). Possessing inhibitory control also allows L2 learners to selectively focus on the target input while ignoring other information not relevant to the successful completion of tasks (Darcy, Mora, & Daidone, 2016). In terms of language learning aptitude, Granena and Long (2013) found that late Chinese-Spanish bilinguals’ (AoA > 16 years) pronunciation performance (perceived foreign accent) was significantly related to their phonemic decoding test scores (i.e. connecting foreign sounds with corresponding symbols), whereas such an aptitude-proficiency correlation was not found for early bilinguals (AoA < 16 years).

**Motivation**

Finally, another well-researched variable in the context of advanced L2 speech attainment is learners’ motivation. Previous research has revealed a surprising amount of variation in the motivation-attainment link (for a review, see Piske et al., 2001), with some studies showing that L2 learners with highly advanced oral proficiency are likely to have a great deal of professional motivation (e.g. “to teach an L2 as university-level academic jobs”: Moyer, 1999), instrumental motivation (e.g. “to get a job and/or respect at work”: Flege, Munro, & MacKay, 1995), integrative motivation (e.g. “to have as many native speaking friends as possible”: Flege et al., 1999), and strong concern for L2 pronunciation accuracy (e.g. “to pronounce English without any L2 accents”: Bongaerts, van Summeren, Planken, & Schils, 1997). However, others have failed to find such significant predictive power of motivation for developing successful L2 pronunciation ability (e.g. Oyama, 1976; Purcell & Suter, 1980). The confounding finding signals that the relationship between motivation and L2 oral proficiency development is complex, and that research investigating this link needs to elaborate valid methods for quantifying motivation in various L1/L2 contexts (Piske et al., 2001).
Conclusion and future directions

Given the theoretical and practical relevance of the topic, this chapter set out to synthesize the extensive literature on advanced L2 segmental and suprasegmental acquisition. As suggested by many researchers (e.g. Flege, 2009), L2 learners continuously improve their segmental and suprasegmental competence and ultimately reach advanced-level proficiency as long as they regularly use the target language as their main medium of communication over long periods of time (e.g. LoR > 10 years). With respect to segmentals, interlanguage development can be characterized as a transition from vocabulary learning (processing lexical units of L2 input as a whole) to sound learning (high sensitivity to segmental aspects of L2 speech). With respect to suprasegmentals (prosody, fluency), L2 learners’ speech may first become fluent (i.e. a reduction in the number of pauses at the early phase of L2 speech learning). Subsequently, L2 learners may slowly enhance the prosodic (word stress, intonation) and temporal (speech rate) qualities of their speech over time, as the automatization of phonetic encoding and articulation requires a great amount of L2 experience. The final quality of their asymptotic performance can be greatly determined in accordance with the L1/L2 cross-linguistic distance (the readjustment of existing cues vs. the acquisition of new cues), learner-extrinsic factors (age of acquisition and testing; quantity and quality of input), and learner-intrinsic factors (cognitive abilities, aptitude, motivation).

To close, I would like to point out a range of issues worthy of future investigation. First, most of the studies reviewed in this chapter have elicited L2 speech performance via controlled speech tasks (e.g. word and sentence reading tasks). This is arguably because these measures allow researchers to focus on target sounds while controlling for other confounding factors (e.g. variation in lexicogrammatical use). Under such highly controlled speech task conditions, however, L2 learners can carefully monitor correct language use, which does not necessarily reflect real-life situations, where language is used in various social settings in multiple tasks under time pressure. To tap into the present state of L2 learners’ segmental/suprasegmental representations and processing abilities, it is important for future studies to adopt not only controlled but also spontaneous and conversational speaking tasks (cf. Piske, Flege, MacKay, & Meador, 2011).

Another promising direction for future research relates to the role of individual differences in advanced L2 speech learning. Although researchers have begun to highlight the relationship between cognition, conation, and ultimate attainment, the concept of individual differences has dramatically changed over the past 20 years. As reviewed above, for instance, the research findings have thus far supported the strong predictive validity of L2 learners’ cognitive abilities for their successful L2 speech learning in the long run, especially when it comes to late L2 learners who likely learn their L2 in an intentional, explicit, and analytical fashion. More research is needed to examine the importance of cognitive abilities and aptitude for implicit and incidental language learning in both the early and the late phase. Several test formats have been adopted for this purpose, such as serial
reaction time (e.g. Granena, 2013), semantic priming (e.g. Linck, Michael, Golonka, & Twist, 2013), and phonological sequences (e.g. Speciale, Ellis, & Bywater, 2004). Such future studies will in turn shed light on the cognitive correlates of successful L2 speech learning throughout one’s lifetime.

Although the current review did not find a strong motivation effect on L2 speech learning and attainment, this conclusion needs to be interpreted with caution. Over the past 20 years, many researchers have proposed different theoretical models to capture the multifaceted, complex, and dynamic nature of motivation (e.g. Dörnyei, Henry, & Muir, 2015, for the L2 Motivational Self System). Drawing on Norton’s (2000) notion of imagined communities, for example, Yashima (2002) developed the idea of “international posture” to explain the specific motivation profiles of Japanese learners of English who are primarily driven by their desire to participate in an imagined international community. Saito, Dewaele, and Hanzawa (in press) provided longitudinal evidence that L2 learners with context-specific motivation (learning English as long-term preparation for an uncertain future career in an imaginary international community) can improve the comprehensibility and intelligibility of their pronunciation forms (but not necessarily their accentedness). These motivated learners tend to have a high willingness to communicate, and to seek and increase the number of opportunities to interact with native (and non-native) speakers in the target language community (Derwing & Munro, 2013). It is presumed that the increased amount of L2 experience will, as a result, facilitate SLA in the long run (Flege, 2009).

NOTE

1 Neighborhood density (i.e. phonological similarity) between words can be determined by (i) substitution (e.g. ‘late’ vs. ‘rate’), (ii) deletion (e.g. ‘late’ vs. ‘ate’), and (iii) addition (e.g. ‘late’ vs. ‘plate’).

REFERENCES


O. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.


learning: In honour of James Emil Flege (pp. 315–330). Amsterdam: John Benjamins.


16 Connected Speech in Advanced-Level Phonology

BURCU GOKGOZ-KURT¹ AND D. ERIC HOLT²

¹ Dumlupinar University, Turkey
² University of South Carolina

Connected speech and the advanced L2 learner

Understanding the spoken language poses a challenge for most second language (L2) learners, especially because it requires L2 learners to determine the boundaries of word sequences uttered in context. As Stephen Pinker (1995, pp. 159–160, cited in Alameen & Levis, 2015, p.160) notes:

In speech sound waves, one word runs into the next seamlessly; there are no little silences between spoken words the way there are white spaces between written words. We simply hallucinate word boundaries when we reach the edge of a stretch of sound that matches some entry in our mental dictionary. This becomes apparent when we listen to speech in a foreign language: it is impossible to tell where one word ends and the next begins.

In this chapter we will flesh out the nature and difficulty for the learner of connected speech. In addition to describing connected speech processes and discussing notions of advancedness, we will look at issues of proficiency, learner context, perceptual saliency, and others, as well as particular challenges learners face, the effects of training, and the effects of individual differences; we likewise touch on pedagogical perspectives, and finally conclude with directions for future research.

Regarding the fluid nature of speech, a Spanish example will illustrate the points made by Pinker: Los alumnos están en el aula (‘the students are in the
Connected Speech in Advanced-Level Phonology

classroom’) is pronounced as \( \text{Lo-\textasciitilde} | \text{a-lum-no-\textasciitilde} | \text{es-ta-n}| \text{e-n}| \text{au-la} \), where no word’s boundaries line up with both beginning and end of syllables as pronounced. If a learner is expecting the beginning of a word to coincide with the beginning of a syllable, and the end of the word to coincide with the end of a syllable (even though a word may be made up of more than one syllable), this type of linked speech greatly degrades perception, and word identification may be hampered or thwarted. However, armed with the knowledge that Spanish prefers open/CV syllables and disfavors vowel-initial syllables when possible, the learner in perception can ‘undo’ the syllable linking to achieve comprehension.

However, however, it may seem to be produced by the interlocutor, deciphering these word boundaries in an L2 is not an easy task at all. This is mainly because word boundaries may disappear or their position may shift, usually as a result of the integration of words into each other during articulation, thus causing a loss of “phonetic identity” in words (Kohler, 1990). This process in which speakers “draw [the sounds] together” and make the word boundaries smooth is referred to as connected speech, e.g., in English, pronouncing ‘want to’ as ‘wanna’ [w\(\text{a\textasciitilde}n\text{a}\)], and ‘going to’ as ‘gonna’ [g\(\text{a\textasciitilde}n\text{a}\)] (Clarey & Dixson, 1963). These shifts in word boundaries and variability in phonetic forms are triggered by various factors, such as the ways of articulation, morphological, lexical, and syntactic properties, sentence stress, and more importantly, the speaking styles as stipulated by the communicative context (Kohler, 1990).

Given the subtlety in producing these processes and the changes they bring about, it should not be surprising that a majority of L2 learners find it challenging to decode the commonly used connected speech forms in an L2, which may cause them to have breakdowns in communication in real life. While in low-proficiency L2 learners this difficulty in decoding may be more likely to stem from gaps in interlanguage syntax, morphology, or lexicon, among others, they are less likely to be the cause in highly proficient L2 learners, who may be more concerned about discerning and addressing style differences in diverse communication situations (see Rubin, 1994). That being said, it should be noted that the proficiency level of an L2 speaker may serve as an important factor determining the type of connected speech forms they are ready to learn. (See more on this in the section ‘Training L2 learners to perceive and produce connected speech processes.’)

Regardless of proficiency level, though, depending on the learning context, there may always be a gap between the language used in the classroom and that used outside. Therefore, higher proficiency in an L2 may not necessarily mean more exposure to connected speech forms, especially in foreign language contexts (Ur, 1987). Learners who are usually exposed to fully articulated ‘teacher talk’ during their L2 learning experiences in classroom contexts find it frustrating when they cannot understand authentic conversations among native speakers and highly proficient non-native speakers of a language. This is especially true for those who learn L2 vocabulary and grammar in their home countries in a rather decontextualized way, including those who are highly proficient in written language but less so in spoken language. Hence, upon arrival at the host country, they usually have a “rude awakening” (Ur, 1987, p. 10) and claim that native speakers
talk “too fast” (Gilbert, 1995, p. 97). This may originally be attributed to the lack of perceptual saliency of connected speech forms, which requires L2 listeners to allocate more attentional resources to be able to recognize these forms in spoken language, and this ability improves as L2 learners develop higher levels of attention control as proficiency goes up (Segalowitz & Frenkiel-Fishman, 2005). Given the difficulty of noticing and deciphering these forms, their use by L2 speakers in spoken utterances seems to be as challenging as perceiving them, although views on this might vary depending on the theoretical framework and other individual differences discussed further below.

Despite such challenges in teaching and learning these forms, a knowledge of connected speech forms (in terms of a general understanding of relevant structures and processes) is crucial for maintaining effective communication skills in an L2 and fosters improved performance. This is mainly because it provides an L2 learner with the necessary skills to understand the subtleties of the authentic use of the target language in spoken language, which in turn helps learners to improve their ability to transition smoothly while listening to and speaking with various speakers in a diversity of contexts.

Describing connected speech processes

The term connected speech is used to refer to processes such as reductions, minimizations, or full eliminations (Brown & Kondo-Brown, 2006) occurring across word boundaries following certain language-specific phonotactic rules (Joyce, 2013). This type of speech has also been referred to as reductions, reduced forms of speech, sandhi variation, or weak forms (Brown & Kondo-Brown, 2006, p. 5; Ito, 2006, p. 17). Connected speech processes (CSPs) may involve changes, additions, or eliminations to sounds and sound sequences, which occur due to various linguistic as well as communicative and pragmatic factors. For example, in English, stress and intonation patterns play a significant role in determining which sounds or sound sequences are to be eliminated or modified. While function words, which usually do not bear stress, undergo deletion, content words and their stress-bearing syllables are not usually eliminated in connected speech processes. The word ‘and’ in the phrase ‘now and then’ is pronounced as [әn] because ‘and’ is a function word and thus it is pronounced in its weak form in connected speech. Similarly, in German, function words including certain types of pronouns, conjunctions, prepositions, and auxiliary verbs are reduced when unstressed. Kohler (1990) adds that as a result of this reduction, it is possible to see different phonetic realizations of the same word depending on the function of reduction. These changes in citation forms of words occur as a result of certain “temporal and articulatory” features of spontaneous speech, among other reasons (Hieke, 1987).

Speakers choose to speak using features of connected speech presumably to save time and energy. In speaking, there is the concept of efficiency, which essentially tolerates the maximum elision of language patterns in an effort to minimize the number of phonological units (Rost, 2002). This is also known as “the principle
of least effort” (Zipf, 1949), or “law of economy” (Clarey & Dixson, 1963), both of which explain why speakers are attracted to speaking with elisions, contractions, and assimilations in their conversations. Describing this as a balance between what the speaker and the listener do during communication, Kohler states that:

Word production is a compromise between articulatory economy for the speaker and acoustic distinctivity for the listener. Economy of effort in speech production is governed by a number of anatomical, physiological and temporal constraints in the speech producing apparatus that introduce directionality into reductions, such that they are not chaotic. Not just any changes, but only certain types are possible, which occur over and over again in the languages of the world and in historical sound change. (1990, p. 9)

So, according to Kohler, connected speech processes are mostly a result of articulatory factors which may make rule-formations describing such changes possible. As the linguistic constraints of the languages differ, “the manifestations and distributions” of these changes also vary across as well as within languages. Among within-language factors are “the common core of linguistic context and context of situation in the widest sense between speaker and hearer, ranging from world knowledge through culture and society to the individual discourse setting” (Kohler, 1990, p. 10). For example, in English, the primary function of the use of connected speech is, in fact, to maintain the rhythm of English by “compressing” unstressed sounds and syllables and making articulation easier (Shockey, 2008). Similar examples of “compressing” can be found in many other languages. For example, in Turkish, in which reductions are observed, the phrase ‘gideceğiz’ meaning ‘we’ll go/leave’ may optionally be pronounced as ‘gidicez’; while the former represents the citation form of the word, the latter is used as a variation of the future suffix, due mostly to a desire to save time and energy. That all of these alterations occur partly because of linguistic requirements but also mostly for communicative and pragmatic reasons poses a significant challenge for L2 listeners: the difficulty of keeping up with the message while listening to this “reduced” speech. In other words, the more reduced a message is, the more effort L2 listeners have to invest in order to perceive and process the spoken text.

One way to help L2 learners to improve their perception and production of connected speech forms is by constructing a comprehensive classification of these forms for languages, and by familiarizing learners with it. Alameen and Levis (2015) recently classified CSPs in English into six main categories. These are linking, deletion, insertion, reduction, multiple processes, and modification. Their definition of linking is limited to “situations in which the ending sound of one word joins the initial sound of the next, but only when there is no change in the character of the segments,” e.g. pronouncing the phrase ‘some of’ [sam əv] as if it were one word (p. 162). Deletion includes elisions, e.g. pronouncing ‘call him’ as [kəl ɪm] by eliding the initial [h], and by contractions they mean pronouncing ‘he will’ as ‘he’ll’. For insertion, they give the example of the cartoon character Popeye’s statement of ‘I am what I am’ realized as ‘I yam what I yam’, in which vowels are connected by
glides at word boundaries (p. 163). Reduction involves vowel reductions in unstressed syllables and some consonant reductions, e.g. the lack of release of /d/ in the phrase ‘bad boy’ (p. 163). Under the category multiple, they mention commonly used lexical chunks that show several changes simultaneously, e.g. phrases such as ‘want to’ pronounced as ‘wanna’ or ‘going to’ pronounced as ‘gonna.’ Finally, the category of modification involves four subcategories: assimilation (e.g. the assimilation of [n] to [m] before a bilabial stop in a phrase like ‘sun beam’); flapping (e.g. pronouncing the alveolar stop [t] as an alveolar flap in North American English in the phase ‘sit around’); glottalization (e.g. pronouncing the phrase ‘can’t make it’ as [kænʔmekɪt] as a result of the [t] before the nasal [m] being pronounced with a specific glottal articulation); and finally, palatalization, e.g. pronouncing ‘that you’ as [ðæʧʊɹ]. From a pedagogical perspective, having such a classification of connected speech is very valuable for linguists, language educators, and L2 textbook writers, as it may help guide them in presenting and teaching these forms in a systematic way.

Alameen and Levis’s classification describes connected speech processes in English, and many other languages also have similar processes in their phonology. In French, for example, a word-final voiceless consonant will be voiced when it is followed by a voiced segment as in the word ‘avec’ /avek/ being pronounced as [aveg] when it is followed by the word ‘vous’ /vu/, that is, /avek vu/ becoming [aveg vu] in its phonetic realization. This is an example of regressive voicing assimilation and is not normally found in English, and this leads French speakers of English to (mis)pronounce ‘nice voice’ /najs vois/ as [najz vojs] by transferring (negatively) the assimilation process found in French.

It should also be pointed out that the term connected speech is not usually used to describe processes occurring within words (Alameen & Levis, 2015). For example, the coalescent assimilation in the transformation of the word ‘face’ *[feɪs]~‘facial’ [feɪʃəl] is similar to the modification in pronouncing ‘that you’ as [ðæʧʊɹ]; however, while the former type of palatalization occurs within words, the latter occurs across word boundaries. In fact, since the processes described here are articulatorily easily explicable or phonetically grounded, and so are ‘natural,’ it is common for a language to evidence these in all applicable phonetic/phonological contexts. To draw from Spanish, phonological processes occur without regard to word boundaries. For example, there is place assimilation between a nasal consonant and a following obstruent, both across words (e.g. *un gato ‘a cat’ pronounced as [uŋ-ga-to]) and within a word (e.g. tango ‘tango (dance)’ as [tæŋ-go]). Likewise, spirantization of the voiced occlusives /b, d, g/ (whereby the plosives become fricatives) occurs not only within words (e.g. haba ‘bean’, hada ‘fairy’, haga ‘you do (subj.)’ with [β, ð, ɣ]), but also across words (e.g. mi balón ‘my ball’, mi dedo ‘my finger’, mi gol ‘my goal’). Similarly, for syllable structure, Spanish prefers open/CV syllables, such that an intervocalic consonant (VCV) is produced as the onset of the final syllable (V-CV), rather than the coda of the initial syllable (*VC-V); when a consonant-final word is uttered in connected speech before a vowel-initial word, the same principles of onset satisfaction and coda avoidance apply, as shown earlier with the example of Los alumnos están en el aula. With regard to
vowel-vowel contact, Spanish likewise ignores word boundaries for determining maintenance of hiatus or formation of a diphthong, with the high vowels /i, u/ only being realized as [i, u] when stressed (or the only vowel of the syllable), being realized as [j, w] otherwise: *siete* ‘seven’ as [sje-te], *bueno* ‘good’ as [bwe-no] within words, but also *si es* ‘if you are’ as [sjes] and *su ala* ‘your wing’ as [swa-la] across words. Thus, while the focus of this chapter is “connected speech phenomena,” this is merely to emphasize the between-word nature of various natural processes that can impede speech processing and word recognition, and degrade listening comprehension, as well as being features that contribute to accentedness and native-likeness.

**Notions about the term advanced**

The term *advanced* may be interpreted and has been employed in various ways. It is used to refer to *instructional levels* that correspond to (usually collegiate) year of study, such that *novice/beginning* = first year, *intermediate* = second year, and *advanced* = third year. Similar terms correspond to *proficiency level*, with the scale developed by the American Council on the Teaching of Foreign Languages (ACTFL) being well known and highly regarded; it is composed of four main levels, *Novice, Intermediate, Advanced*, and *Superior*, with the first three levels further divided into sublevels, *Low, Mid*, and *High* (see www.actfl.org). The ACTFL scale is used to determine a speaker’s level based on the administration of the Oral Proficiency Interview (OPI). Since it assesses proficiency, rather than achievement, the scale measures what a user can do with the language, rather than what the user knows about the language. High function in the target language can be important to determine, but it doesn’t necessarily correspond closely with mastery of phonology. Pronunciation/phonology per se is not really addressed or assessed by the OPI or ACTFL proficiency scale; that is, a speaker may be able to carry out high-level functions with appropriate vocabulary and grammar, yet speak with a heavy foreign/non-native accent. Similarly, a learner may have passive knowledge of L2 phonology/pronunciation, yet not be able to implement it, and still sound far from a native speaker. So, it may very well be the case that a student who is enrolled in or even has completed third-year language courses does not evidence ‘advanced’ proficiency. (Relatedly, the student may or may not have even achieved ‘advanced’ knowledge/competence for some particular target structure/feature.)

*Fluency* is often referred to in definitions of levels of proficiency, but it is also the case that a learner may be quite able to mobilize language for communication and deliver a message effectively yet still not master either segmental (individual sounds) or suprasegmental (syllable structure, pitch/intonation, between-word linking, etc.) features of the L2; likewise, a learner may be able to speak without hesitations using ‘advanced’ lexis and structures, even at a fast rate of speech, yet not be able to decode/understand a response delivered fluently.
Finally, we might also evaluate features of connected speech themselves along some scale of difficulty, such that a given phenomenon may be considered ‘basic’ (easiest to notice and/or acquire), ‘intermediate’ or ‘advanced’ (hardest to notice and/or acquire), with ease of noticing (perceptual salience) also not necessarily corresponding with ease of production.

One would think it would be reasonable to assume that increased noticing (passive knowledge) would lead to increased oral proficiency, but there are intervening factors of physical performance and phonetic implementation, e.g. a lifetime of fine-grained L1 motor habits that need to be adjusted and overcome to approach native-like norms. Likewise, it seems reasonable to hypothesize that increased noticing would lead to improved speech processing and consequently listening comprehension. However, these are undetermined hypotheses, especially outside of an ESL/EFL (English as a second/foreign language) context, and much research along these lines remains to be conceptualized and carried out, and we encourage the interested and curious reader to pursue these open issues.

Similarly, research remains to be done to determine what sorts of phonological units/structures/features/processes (e.g. number and/or complexity of individual sounds/phonemes/allophones, phonotactics/syllable types and adjustments to them, as well as prosodic features of stress, tone, intonation, and focus) in a learning context are ‘basic,’ ‘intermediate,’ or ‘advanced.’ Surely this will vary from one learning context to another (due to the particular L1 and L2), but likely there are universals to be uncovered. And not unrelatedly, it remains to be elucidated which features can be most easily taught and acquired in the first-, second-, and third-year classroom, and beyond, as well as the ‘yield’ of various CSPs toward improved ability to decipher/decode in listening comprehension, and exactly which features most raise native-speaker perceptions/ratings of unaccentedness/native-likeness.

These are many and difficult questions to answer. The remainder of this chapter discusses some of these issues, and at the end of the chapter, others are given as areas for further study.

Training L2 learners to perceive and produce connected speech processes

Motivation and challenges

Ability to perceive and produce connected speech forms has been shown to lead to better L2 communication (Matsuzawa, 2006; Underwood & Wallace, 2012). One of the earliest studies by Henrichsen (1984) compared native speakers and low-level and high-level learners in their comprehension of spoken English sentences in the absence and presence of connected speech. The findings indicated an interaction between proficiency level and the comprehension scores for the spoken input in the presence/absence of reduced forms. In other words, while advanced learners performed similarly to native English speakers in the absence of reduced forms,
their comprehension scores were much closer to low-proficiency L2 learners in the presence of such forms. The results show the importance of familiarizing L2 learners of all proficiency levels, including advanced learners, with CSPs for better L2 comprehension.

A comprehensive study by Joyce (2011) looked at the relationship between linguistic knowledge, psycholinguistic subskills, and L2 listening proficiency to investigate the factors that may help determine the L2 aural comprehension ability and processing. His findings suggested that knowledge of connected speech processes, phonological modification knowledge, as he calls it, was one of the two most important subskills—the other being syntactic knowledge—playing a role in L2 listening performance (p. 86). As an implication of his study, he encourages test designers to incorporate the ability to accurately process reduced forms as part of their goals in designing their testing tools, adding that this information could be used to adjust the difficulty level of a listening-test item as an indicator of proficiency (pp. 87–88). In fact, this is also in line with Kostin’s (2004) study which investigated the factors affecting the difficulty of the Test of English as a Foreign Language (TOEFL) dialogues and found sandhi variation as one of the phonological variables that most cause L2 listeners to experience difficulties in comprehension. Thus, it is clear that being acquainted with connected speech features is highly important in understanding native or highly proficient speakers of an L2 in a variety of contexts, including high-stakes testing conditions. Consequently, paying equal attention to all aspects of an L2 education is central to helping learners advance their foreign/second language skills.

These studies show that teaching and learning connected speech forms is crucial. Immersion in the L2 learning environment is one way to enable learners to familiarize themselves with features of connected speech by intensively experiencing the language in authentic oral contexts. However, learning by such exposure may require relatively extensive periods of time, which may not be feasible for most L2 learners. L2 instruction, on the other hand, may benefit L2 learners ideally by providing them with opportunities to help them notice the target structures (see Schmidt, 1990, 1995). One challenge for learners in noticing the various connected speech forms is the issue of ‘perceptual saliency.’ This refers to the features of speech that make “certain features of the input more comprehensible, and thus more liable to become intake” (Henrichsen, 1984, p. 106). Since perceptual saliency plays a crucial role in determining the ease of learning certain L2 features, Henrichsen (1984) sees various connected speech processes as a disadvantage since such forms are not salient by nature. Therefore, teaching these forms may help learners to notice and eventually learn them. However, there have been various challenges in teaching connected speech in L2 classrooms.

One is related to the ‘informal-only’ or ‘substandard’ view of connected speech. This view prevails among many teachers and listeners, as such processes have been claimed to occur in fast, colloquial, casual, informal, and relaxed speech (Brown & Kondo-Brown, 2006, p. 5; Trager, 1982; Weinstein, 1982). Nevertheless, such suppositions need to be reconsidered because connected speech may occur in all registers, including academic and formal settings (Brown, 1977, pp. 2–3; Brown &
Kondo-Brown, 2006, p. 5; Ito, 2006; Rogerson, 2006). In fact, Underwood (1989) explains how difficult it is to draw the boundaries of formality/informality as follows:

... for the language learner the division is not as neat ... It frequently happens, for example, that a lecturer delivering a very formal lecture from a prepared set of notes switches to informal language when making an aside or recounting an anecdote as an illustration of a point just made. Or a person involved in describing a complicated phenomenon to a friend over coffee may switch in and out of formal and informal styles depending on whether he/she is describing the phenomenon or commenting on it. Between the extremes, there is a range of formality/informality depending on the social setting, the relative ages and status of the speaker and listener, their attitudes to each other and the topic, the extent to which they share the same background knowledge, and so on. (p. 14)

In this vein, it does not seem reasonable to be too restrictive in making claims regarding the contexts in which connected speech is used, which means that depriving L2 learners from exploring the features of connected speech might not be the best choice to make in a language classroom. In fact, connected speech could potentially help L2 learners feel sociolinguistically more advantaged, and even when considered as a marker of informality, it might help learners determine any ‘switching’ between informal and formal use of language in spoken discourse (Underwood, 1989). Unfortunately, this aforementioned stereotypical view of connected speech as ‘substandard’ (Brown & Kondo-Brown, 2006, p. 5) or ‘informal-only’ prevails among many L2 teachers and listeners, which is, in fact, one of the reasons why teachers tend not to teach it and learners tend not to consider it a priority in their English learning experience (Brown, 1977, p. 3; Gilbert, 1995, p. 105). Instead, both teachers and students, of all L2s, need to be made aware of and appreciate that connected speech is indeed ‘everyday speech’ and should be considered the default for most circumstances.

In addition to its perceived ‘substandard’ status, another reason for the reluctance to practice connected speech phenomena in an L2 classroom may be teachers’ lack of knowledge of and attention to these forms and of appropriate methods and techniques to teach them (Ito, 2006). A majority of L2 teachers are not usually familiar with these structures, especially in foreign language contexts, and even if they are, the instruction may not be systematic enough for learners to make generalizations (Rogerson, 2006, p. 91). These challenges are exacerbated by the lack of materials and of sufficient time to devote to teaching connected speech in classrooms simply because they are not usually part of the curriculum (Brown & Hilferty, 1986/2006; Henrichsen, 1984; Joyce, 2013; Rogerson, 2006; Underwood & Wallace, 2012). All these result in an absence of focus on these forms in foreign and second language classrooms, despite their significant role in L2 communication.

However, it should be noted that these challenges will vary from language to language, and from instructor to instructor. With regard to the latter, teachers (or their curricular coordinators) may feel they have to make choices about what material and structures to cover in a given course, and often neglect matters of
pronunciation, which are often minimally present and ‘marginalized’ anyway (see Derwing & Munro, 2005, p. 382). Additionally, as shown by previous research, L2 instructors do not always have in-depth linguistics-oriented training or any such training at all (Breitkreutz, Derwing, & Rossiter, 2001; Derwing & Rossiter, 2002), instead being practiced in pedagogy, and are more aware of and comfortable with issues of vocabulary and grammar. Very frequently, the first comprehensive and systematic study of phonology, including CSPs, is a dedicated course in foreign language pronunciation that is typically taught after the third-year grammar sequence, or in beginning graduate-level coursework in linguistics. Students often comment that they wish they had been exposed to this knowledge earlier in their learning, and instructors often comment that even so-called advanced learners have fossilized in pronunciation, given the late exposure (if at all) to in-depth knowledge and practice, and the cumulative time of language learning while they were often left to develop unintelligible speech patterns, and were not aware of negative transfer from their L1. However, this will vary some according to the particular L2. It is very true of Spanish, in part because most learners of L2 Spanish often suffer from the misconception that Spanish is pronounced as it’s written, and written as it’s pronounced, both of which notions are false. In French, on the other hand, at least some CSPs (and other phonological phenomena) are taught from early stages because of the depth of French orthography, that is, because it is patently obvious from the start that French spelling does not correspond very closely/shallowly with the spoken form, and that if a learner pronounces French as it is written, it will be largely incomprehensible, and if learners expect to recognize words in spoken French, they must realize that the sound-spelling mapping is a very complex one.

Looking at the previous research (see Henrichsen, 1984; Joyce, 2011; Kostin, 2004) that reveals the essential role that connected speech plays in L2 comprehension and communication, it can be maintained that CSPs are worthy of study and practice, being central to successful communication.

**Effects of training and individual differences**

The teachability and the effects of instruction on connected speech comprehension and perception have been systematically investigated in several studies (Brown & Hilferty, 1986/2006; Crawford, 2005; Matsuzawa, 2006), and the findings showed an improvement in learners’ listening skills in connected speech. One of the pioneering studies looking at connected speech, by Brown & Hilferty (1986/2006), investigated the effects of four weeks of daily 10-minute instruction on connected speech in L2 English. Their findings suggested an improvement in Chinese university students’ connected speech abilities on a dictation and an integrated grammar test, but there was no improvement on the general listening comprehension test. A more recent study by Matsuzawa (2006) showed that instruction on reduced forms can in fact lead to improvement in L2 listening ability based on the results of a cloze-test. Similar studies are needed for other languages.
Limited previous research on the production aspect suggests that the production of connected speech features improves over time when such features are practiced in a traditional classroom context (Underwood & Wallace, 2012) as well as when using computer-assisted language learning (Yang, Lin, & Chung, 2009). Underwood and Wallace (2012), who looked at both production and comprehension, also found a significant improvement in Japanese learners’ connected speech comprehension and their self-confidence in conversational ability following 10 weekly instructional periods. Although the findings showed significant improvement in both production and comprehension, there was no correlation between learners’ ability to comprehend reduced forms in a listening test and their production in a spontaneous peer conversation. Similarly, Alameen (2014) looked at the effects of different instructional methods on the ability to perceive and produce linking as a connected speech phenomenon in L2 English. Her results showed no significant improvement on the perceptual ability based on the results of the dictation test; however, significant values were reached in the ability to produce linking. Another study, by Kennedy, Blanchet, and Trofimovich (2013), investigated gains by L2 learners of French in segmental production, stress, intonation, liaison, and enchaînement, in addition to learner awareness as measured by learning diaries. The findings indicated no post-instructional gains except for segmental production.

However, it should be noted that there seems to be a disagreement among researchers and language practitioners as to whether or not producing CSPs should be the focus in classrooms. Norris (1993, 1995) suggests that the purpose for learners should be the recognition of connected speech features in order to communicate well, rather than imitating native speakers’ use of connected speech features in learners’ own speech. Brown (1977) also explicitly disapproves of teaching students to “produce” these “assimilated” or “elided” forms because “sophisticated students who have been taught to be aware of these forms will introduce them into their own speech in a natural context when they feel able to control them” (pp. 156–157). However, she finds “the failure” to understand these forms as “disastrous for any student who wants to be able to cope with a native English situation” (p. 157).3 This being said, advanced L2 learners should be able to naturally make such forms a part of their utterances if they are taught how to perceive, and ideally produce, these forms in a classroom context. The literature mostly agrees on the prominence of teaching the perception and comprehension of connected speech features more than the production of them, mainly because the primary goal of pronunciation teaching is considered to be accuracy in perception and comprehension, followed by production.

Previous literature has shown that L2 learners mostly benefit from instruction in helping them make progress on improving their perception and production of connected speech forms. However, learners vary a great deal. There is always variation in their gains in an L2 classroom in terms of their pace of learning and ultimate achievement. Learner backgrounds and learner conditions, as well as other individual differences in cognitive abilities, have been shown to influence the learning process. According to Ellis (2004), some of the most commonly researched factors include language aptitude, learning styles, motivation, anxiety, personality,
proficiency, learner beliefs, learning strategies, and intelligence and memory, among others. There are many other factors, such as L1 transfer, linguistic background, and age of exposure or length of residence, that also affect the learning process.

In fact, studies looking at CSPs in relation to individual variables are very scarce. Some of these looked at study abroad, learner awareness, phonological and cognitive skills, and listening conditions. In order to identify the contribution of certain phonological skills to connected speech perception, Wong and his colleagues (2017a) gave intermediate-to-advanced L1 Chinese learners of L2 English a battery of tasks to test their spoken word discrimination, part-word recognition, phoneme awareness, receptive vocabulary, and phonological memory. The findings revealed that while receptive vocabulary and part-word recognition predicted connected speech perception, phonemic awareness and phonological memory only did so indirectly through the mediator of part-word recognition. The spoken word recognition task, on the other hand, was not found to be a factor in explaining the individual differences in the comprehension of the reduced forms in this study. In a related study, Wong et al. (2017b) this time looked at the effects of varying listening conditions (multi-talker babble noise, speech-shaped noise, factory noise, whispering, and sad emotional tones) on connected speech perception by Chinese L2 learners of English. Learners were found to have more difficulty recognizing connected speech forms under noisy (as opposed to noise-free) conditions, with multi-talker babble noise creating the greatest challenge.

A study by Kim (1995) investigated the role of attention in understanding speech at different speech rates and found that listeners paid less attention to speech when it was read at a faster rate. Therefore, according to him, listeners should be “encouraged to move from a more lexical mode […] to a more syntactic mode” (p. 78) because this way, they would be able to comprehend connected speech processes occurring across word boundaries, which would otherwise go unnoticed (Ito, 2006, p. 23). These findings indicate that training learners to ‘notice’ particular forms might prove helpful not only for L2 learning in general, but also specifically for learning connected speech forms (Kim, 1995). Another study, by Gokgoz-Kurt (2016), showed how attention control relates to improvement in connected speech perception, specifically, word-boundary palatalization in L2 English, following online training sessions. Results indicated that L2 learners of English benefited from online training and there was a significant relationship between learners’ attention control and phonological learning, which shows the crucial role attention control plays in learning connected speech.

In his small-scale study, Simões (1996) looked at the effects of study abroad on pronunciation of five L2 Spanish learners. Only two participants were found to have improved on several aspects of their pronunciation, with their use of linking between words being one of them. Kennedy and Blanchet’s (2014) study investigated how L2 French learners’ improvement in connected speech perception was related to their language awareness. Following various activities geared to practicing connected speech perception (e.g. linking) and raising their awareness, learners who focused more on “how to use that knowledge to extract meaning
from speech” rather than rehearsing knowledge improved more on perception. Finally, Ernestus, Kouwenhoven, and Van Mulken’s (2017) study examined the role of phonotactic constraints and native language in L2 listeners’ interpretation of reduced vowels in casual speech, and their findings show a direct effect of L1 phonotactic rules, but this effect decreases as the proficiency level increases.

The studies summarized so far contribute to our understanding of CSPs in various ways, but more studies are needed to explore individual differences as correlates of connected speech processes.

**Suggestions for future research**

Much work needs to be done to better understand the perception and production of connected speech at advanced levels of proficiency, as well as the efficacy of pedagogy, and there is much more we could learn from innovative research methods. While most research in connected speech has been carried out using similar methods and tools, various technological tools would be useful to continue to explore and develop, and various research topics merit much further attention. In what follows, some of these are fleshed out further, while others remain for other investigators to elaborate.

**Instruction and training**

The effects of instruction and training on CSPs have been investigated in various learning contexts; however, not many empirical studies have looked at the effects of proficiency level in learning these forms, or those which did so did not have any significant differences in learning outcomes, which can possibly be due to the low number of participants representing each proficiency group in the respective studies (e.g. Gokgoz-Kurt, 2016). Further studies might consider including a balanced number of people from a variety of proficiency levels as this may reveal to what extent each group of learners benefits from instruction. Empirical studies may test whether and how the classification of CSPs (see Alameen & Levis, 2015) may be useful for different proficiency groups. Some appropriate questions to ask are: What may be the best time to introduce CSPs in the L2 learning process? How can a further classification of different CSPs be made based on the readiness of different proficiency levels? What types of instructional methods are most beneficial?

Moreover, research investigating the perception and production of different aspects of pronunciation are not scarce, but there are very few studies looking at the interaction between perception and production of connected speech (e.g. Alameen, 2014), which calls for additional study. Considering proficiency level, is there a better way to make CSPs more accessible and comprehensible for different proficiency levels? Is there a preferred or more effective order when it comes to teaching connected speech perception and production to low- versus high-proficiency learners? More studies to investigate related questions would contribute much
to general understanding of CSPs in relation to proficiency levels and eventually help improve pedagogical methods and materials.

Another interesting avenue for research is the use of technology in learning and teaching CSPs. Several studies have used some form of technology in their investigation of various CSPs (e.g. Alameen, 2014; Gokgoz‐Kurt, 2016; Yang et al., 2009); however, technology has more to offer in teaching connected speech perception and production. In search of effective and practical ways to learn, teach, and assess CSPs, more studies are needed using acoustic analyses for visual feedback, speech recognition, and other exciting possibilities technology offers, besides those which compare computer-assisted versus traditional methods of teaching CSPs. Although all proficiency levels will benefit from the use of technology in learning connected speech, the reality is that advanced learners will most likely be more motivated to put time and effort into using technology to learn these forms compared to low-proficiency-level learners. This is because low-proficiency learners may prefer to use technology to understand ‘be going to’ as a pure grammatical point before, if not simultaneously, practicing ‘gonna’ [gʌnə]. Additionally, technology may be a good means to help learners to work autonomously in their connected speech learning process. This is especially true for advanced L2 learners, as shown by previous research in other aspects of pronunciation (Mantini, 1980). In addition to classroom learning studies, more studies investigating naturalistic L2 connected speech learning are also very much needed. That way, it might be possible to investigate the effects of instruction versus mere exposure to target forms, which may lead researchers to make stronger claims regarding the sources of improvement.

**Individual learner differences**

In fact, examining individual learner differences is what will give us insight into the interaction of proficiency level and the teaching and learning of CSPs. So far, studies exploring cognitive and affective factors in relation to connected speech learning have looked at attention (Gokgoz‐Kurt, 2016; Kim, 1995), learner awareness (Kennedy & Blanchet, 2014), L1 phonotactics and proficiency (Ernestus et al., 2017), listening conditions (Wong et al., 2017b), study abroad (Simões, 1996), and other factors such as spoken word discrimination, part-word recognition, phoneme awareness, receptive vocabulary, and phonological memory (Wong et al., 2017a). However, not all these studies found a considerable relationship or interaction between them, and more importantly none looked specifically at advanced learners or proficiency levels. We should continue to survey these factors in a deeper and extensive way, varying the learning contexts, the time and nature of instruction and/or contact, and learner profiles. What factors better predict connected speech learning in advanced L2 learners? To what extent does aptitude account for successful learning/failure in connected speech learning in advanced L2 learners given the same conditions? There are many other factors, such as learning strategies, cognitive styles, L1 transfer, age of exposure, or length of residence, which affect the learning process and constitute avenues for further
research. Also, advanced learners’ and instructors’ perceptions of (teaching) connected speech or their motivations or learning and teaching strategies may also be investigated.

**Research methods and tools**

In order to uncover the nature and interaction of connected speech with other factors, some of which were enumerated above, we need accurate and innovative methods and tools to measure connected speech. Dictation tests and cloze-tests have served as the most commonly used techniques, but they are not without shortcomings. Dictation tests are hard to design because the test designer needs to make sure that the lexical and syntactic structures that make up the test are not above the proficiency level of the test takers, as this would otherwise undermine the results of the test. In a cloze-test, on the other hand, test takers do not need to keep the whole sentence in mind as it is being dictated, as they only need to focus on one or two missing words. This means a greater burden is placed on working memory when taking a dictation test than when taking a cloze-test. In a cloze-test, since learners see the rest of the sentence and know how many words there are in each sentence, their likelihood of guessing the missing word(s) is higher (Joyce, 2013). Possibly to avoid such a problem, Matsuzawa used a different type of cloze-test, in which the number of words missing was indicated in each sentence with blanks—three blanks meant three words—and did this for both the target and the non-target words. This alternative may help to a certain extent to address the problem of guessing words from context; however, before administering such a test, respondents should be informed about what exactly corresponds to a ‘word’ or whether a contracted form is counted as one or two words. Yet another alternative method to test connected speech perception is having forced-choice tests in which similar sounding options are presented to learners to choose from (see Gokgoz-Kurt, 2016). Using such a test relatively decreases the burden on the working memory and eliminates the problem of having to write using correct spelling or of having to understand all the lexical and syntactic structures, making it more suitable for testing aural perception rather than listening comprehension. However, there are specific challenges to preparing a forced-choice test for the assessment of connected speech. Although they vary depending on the type of CSPs, among these challenges are coming up with similar sounding yet contrasting pairs of grammatical phrases while adhering to the language-specific phonological constraints such as stress or limiting the number of cues which inadvertently help learners guess the right option. Therefore, in order to have a more thorough understanding of connected speech phenomena, better techniques and tools need to be developed. A good starting point could be designing mixed-method studies using two or more types of tests simultaneously, and then comparing the results of each tool to see if they would yield similar findings. This may provide us with alternatives to assess connected speech in a more structured way, paving the way for CSPs to constitute a larger part of L2 curriculum and instruction. Previous studies have also suggested better ways to assess connected speech production.
Designing studies which use speech recognition technology to provide feedback on connected speech production is one of them (Alameen & Levis, 2015). However, since it is already hard to recognize connected speech usages in natural speech, developing some technology to identify them in an effective way seems to be a challenging yet interesting task to achieve.

Finally, a complete picture of L2 connected speech learning in advanced L2 learners is only possible through a better understanding of the role of instruction and individual learner differences. Hence, further studies should look at the predictors of better connected speech perception and production, which could help researchers gain deeper insight into the factors underlying L2 phonological acquisition, and what it takes to reach truly ‘advanced’ standing in second language acquisition.

Acknowledgment

The authors wish to recognize the volume editors and several of the authors’ colleagues for sharing their expertise and insights; they are not to be held responsible for any deficiencies in the present work. This work was partially supported by a SPARC Graduate Research Grant from the Office of the Vice President for Research at the University of South Carolina.

NOTES

1 A 1999 Tribune Media Group comic manipulates connected speech processes (CSPs) for humorous effect:

South Carolina US Senator Strom Thurmond to incoming US Supreme Court Chief Justice William Rehnquist:

DO YEWSOMLY SWEAH DUPHOLE DEECONSTOOSHIN ADEEYOONATTID STAYSUVAMECKA SAHEPYAGOD?

(‘Do you solemnly swear to uphold the Constitution of the United States of America, so help you God?’)

Rehnquist, baffled, responds: Whatever you say.

Here, segmental and syllabic reduction, palatalization, and other segmental differences occurring in connected speech (here, specifically a regional variety of Southern English) blur word boundaries and may cause confusion for the untrained listener.

2 Another prominent framework is that of the Council of Europe, whose Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR; Council of Europe, 2001) assigns learners to the categories of Proficient User (C2, C1), Independent User (B2, B1), and Basic User (A2, A1). Fluency is a criterion, but not pronunciation.

3 However, it should be noted that while some learners of L2 English do not need to follow native-speaker norms in learning English, there are others who are motivated to learn and speak native-like L2 English. In addition, although learning and teaching
connected speech forms may not be considered a part of ELF (English as a lingua franca, that is, English spoken as a common language among non-native speakers of English). Jenkins (2000, 2004) suggests that one should improve their ability to understand these forms if they are expecting to have considerable contact with native speakers.

REFERENCES


Gokgoz-Kurt, B. (2016). *Attention control and the effects of online training in improving connected speech perception by*


17 Voice Onset Time in Advanced SLA

ALFONSO MORALES-FRONT
Georgetown University

Introduction

Common sense would suggest that if the L1 and L2 both have a phoneme /d/, its acquisition should be unproblematic for L2 learners (L2rs). However, in many cases, the opposite happens to be the case. Take for instance English and Spanish. Both languages have /d/ but the phonetic realizations are very different. In Spanish, the sound is apical; in English, it is laminal. In Spanish, it is pre-voiced after a pause, fully voiced between vowels, and devoiced before a pause. In English, it is voiceless before a pause, voiced between vowels, and only partially devoiced before a pause. The Spanish stop becomes approximant between vowels and the English counterpart is flapped in the same context. To make matters more complicated for the L2r, /d/ needs to maintain a contrast with /t/ so where /d/ is pronounced voiceless in English, /t/ must have some additional feature (aspiration) to avoid the confusion of the two sounds. In Spanish, there is no confusion because /d/ is actually produced as a fricative or approximant between vowels. In light of these considerations, it is reasonable to wonder if the L2rs perceive those details and if they can acquire them. Previous research has given a positive answer to both of those questions but the when and how of this acquisition remains a matter of debate.

Common sense again would suggest that the L2r will start at an initial state heavily influenced by the L1, and will eventually progress to a final state matching the L2 target. However, it seems that reaching this ideal final state is not easy. Pronunciation is an area notorious for ceiling effects. Advanced learners can normally master syntax or morphology, but as the Conrad phenomenon suggests, pronunciation seems to be different. For the laryngeal contrasts, previous research reports linear progress up to the advanced level. However, during the advanced
stage, at the point where we would expect L2rs to finally narrow down the gap and move toward the complete mastering of the L2 patterns, we find that learning stalls.

The issues we review in this chapter revolve around ultimate attainment and this apparent lack of success. To understand the nature of the topic at hand we will first give a general introduction to the concept of voice onset time (VOT). To benefit from the arguments of the second part, it is paramount to have a basic understanding of what VOT is, and how it contributes to cross-linguistic voicing contrasts. We will then offer a review of scholarly debates surrounding theoretical issues and will consider two main options for the feature specifications of the laryngeal contrasts. Next, we proceed to offer a synthesis of research on the acquisition of VOT focusing on what different studies have reported about the advanced stages of acquisition. After reaching a conclusion about the nature and origin of the general limitations to VOT learning, we will turn to the instructional setting and outline a call for future research to address pedagogical intervention as it relates to VOT and advanced proficiency.

**Voice onset time**

There is a tendency among language instructors to think that VOT is a technical term for aspiration. Many textbooks mention the visual technique of producing stops against two tightly held pieces of paper, or a candle, to illustrate the puff or air that characterizes the production of some voiceless stops in languages such as English. This is a metonymic conceptualization in that it substitutes the underlying gestural activity for its effect. From an articulatory point of view, VOT is actually a measure of the timing difference between an oral and a glottal gesture, such as the timing between the precise moment when the lips open after a closure (oral gesture) to the instant when the vocal cords are set in position to start vibrating (glottal gesture). Acoustically, VOT corresponds to the timing difference between the perceived burst of energy that follows the stop release and the onset of perceived quasi-periodical vibration. Aspiration is the byproduct of the delayed closing of the glottis in preparation for voicing in the following sonorant. It is a byproduct of a long lag VOT (usually defined as 50 ms or more) that results in turbulent air rushing through the vocal cords before the right conditions for voicing are set in place.

VOT is one of the laryngeal attributes that together with voicing, aspiration, articulatory effort, duration of closure/explosion, intensity of the stop burst, duration of the following vowel, and $f_0$ transitions, contribute to the phonetic contrast of stops across languages. Many languages distinguish between a voiceless series of the stops represented orthographically as $<p, t, k>$ and a series of voiced counterparts represented as $<b, d, g>$. The consistent representation of the writing systems belies a number of phonetic differences that L2rs have to discover and master in order to improve their pronunciation of the target language.
Besides the two-way contrast between voiced and voiceless obstruents, there are languages with only one series of voiceless stops (e.g. Hawaiian, Mandarin), others with a three-way contrast (e.g. Korean, Eastern Armenian), and still others with a four-way contrast (e.g. Hindi, Marathi). For reasons of space, the number of languages involved, and the potential number of L2rs affected, we will limit our attention to two-way contrasts and will use Spanish and English as proxies for the two main types of VOT settings. Finally, we focus on VOT (rather than the other laryngeal features) following most previous research. After Lisker and Abramson (1964), it has become a common place to present VOT as the basis for voicing contrasts across languages. These authors argued that the other factors at play in laryngeal contrasts (i.e. voicing, aspiration, vowel duration, f0, to name a few) are either devoid of constant phonetic correlates or are effects of VOT that are not independently controlled. Although we do not necessarily agree with these arguments, their impact pervades most of the previous VOT literature, and it is for us a convenient way to narrow down in several ways the depth and scope of this chapter. Nonetheless, in those cases where it is pertinent to give a fuller picture, we will emphasize that VOT is actually not the only basis of laryngeal contrasts.

To understand VOT one needs to understand voicing first. Abstracting away from other details, voicing is the result of a sophisticated control of aperture and muscular tension that takes advantage of aerodynamic forces. In the case of stops, the glottal gesture responsible for voicing is coordinated with the oral gesture. The oral gesture is realized by approximating an active articulator to a passive one to the point of completely stopping the passage of air and creating a buildup of pressure behind the point of articulation. This pressure is incompatible with voicing because, for voicing, air passing through the cords is necessary. Voicing can only be maintained if there is a sufficient transglottal pressure drop. As speakers, we use different strategies in the articulation of true voiced stops to expand the oral cavity in an attempt to maintain the required pressure drop (Westbury, 1983). This is normally accomplished by lowering the larynx and expanding the pharyngeal cavity or by slightly opening the velum to allow some air leakage.

The oral and glottal gestures are normally juxtaposed such that the end of the stop coincides with the beginning of glottal pulsation for the vowel. VOT is the measure of the time between these two gestures. When gestures are synchronous, the measured VOT has 0 value. When the glottal gesture precedes the end of the oral gesture, we have negative VOT (lead), and if there is a delay between the two gestures, we have a positive VOT (lag).

Since voicing requires a precise approximation of the vocal cords with a delicate control of muscular tension coupled with the exact amount of subglottal pressure, it is not something that can be switched on and off instantaneously. It takes time and anticipation. The three types of coordination illustrated above are arguably the result of conscious manipulation of this timing.

Lisker and Abramson (1964) studied the word-initial stop contrasts of 11 languages. For languages with a two-way contrast, they distinguish two major types: languages that contrast VOT lead to short-lag (Spanish, French, Italian, Ukrainian, Russian, Dutch, etc.) and languages that contrast short-lag to a long-lag (English,
Languages with a two-way contrast choose one of those two parameters while a three-way is obtained by contrasting lead, short-lag, and long-lag phenomena. A four-way contrast adds to the mix the rare aspirated voiced stops, and so it is no longer just a matter of VOT, but also voicing.

Lisker and Abramson (1964) reported the values for utterance-initial VOT in Spanish and English, provided in Table 17.1.

It is worth mentioning that there are many factors that can affect the means reported by researchers and this results in different authors reporting different numbers. Flege and Eefting (1987), for instance, reported 18 ms, 22 ms, 39 ms for Spanish p, t, k using the same Spanish dialect (Puerto Rican) as Lisker and Abramson (1964). In spite of the variation, the values reported tend to be fairly consistent. It is interesting to note that the VOT values for Spanish [b] and English [p] are practically the same (this also applies to the other pairs [t,d] and [k,g]). This can result in Spanish /p/ being perceived as /b/ and vice versa.

Starting with Keating (1984), languages that contrast lead VOT to short-lag are considered voicing languages and languages that contrast short-lag to long-lag are labeled aspirating languages. A number of factors contributed to the boom that VOT research has enjoyed over the last 50 years. Research on categorical perception using VOT by Eimas, Siqueland, Jusczyk, and Vigorito (1971) showed that one-month-old babies where already able to discriminate between variants of /b/ and /p/ in terms of VOT. This gave support to the innatist camp and, no doubt, helped with the fascination with VOT. When Kuhl and Miller (1978) demonstrated, using the same stimuli as Eimas et al. (1971), that chinchillas also displayed categorical perception around the same +30 ms boundary as newborns, this was interpreted as evidence that speech evolved to exploit sound differences that our brains could already reliably perceive. This meant that VOT is not only innate, but also inherent to the mammalian brain wiring.

VOT had always been a relatively simple attribute to study for any phonetician with a decent lab. However, when tools such as Praat (Boersma & Weenink, 2017) became widely available, the production of research devoted to VOT saw a considerable increase. In a discipline such as linguistics, often focused in the study of abstract units, there is a certain attraction for objective measures that are easily extracted from spectrograms.

Another aspect that has fueled interest on VOT is the issue of the interplay between phonetics and phonology. VOT is clearly a phonetic attribute, but the mapping of [+voice] to negative VOT or [-voice] to positive VOT is not a straightforward matter. We also have the added complexity that phonologists posit the

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>d</th>
<th>g</th>
<th>p</th>
<th>t</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>139</td>
<td>110</td>
<td>108</td>
<td>4</td>
<td>9</td>
<td>29</td>
</tr>
<tr>
<td>English</td>
<td>1</td>
<td>5</td>
<td>21</td>
<td>58</td>
<td>70</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 17.1 VOT values in milliseconds.
[spread glottis] (sg) feature as independently needed. This feature is actually closer to VOT than [voice], and some authors (e.g. Harris, 1984) assume that [sg] is actually at the base of VOT contrasts in aspirating languages.

SLA's interest in VOT follows from this basic question: How do learners with a given L1 parameter for VOT acquire the contrasts in languages that rely on different parameters? Most studies have focused on how L2 learners from aspirating languages acquire the contrasts of voicing languages, and vice versa. VOT has often been the base for inferences about general foreign accent. In that sense studies of VOT have tried to address questions such as ultimate attainment, age of learning (AoL), category formation, the role of individual and social factors, and categorical perception. VOT has also been used to answer general questions about the settings of the bilingual mind.

The target of acquisition

The question of what needs to be acquired to reach native-like VOT settings deserves different answers at the phonetic and at the phonological levels. Phonologically, the goal is arguably to acquire the correct distinctive features for the L2; phonetically, the goal is to fine-tune the timing of the gestures involved in VOT. Unfortunately, at the phonological level, there is no general consensus on the nature of the distinctive laryngeal features in different languages; at the phonetic level, it is increasingly clear that VOT is not the only phonetic cue responsible for stop contrasts in all languages.

There are two main positions: The Single Feature Hypothesis (SFH) and the Multiple Feature Hypothesis (MFH; see Kager, van der Feest, Fikkert, Kerkhoff, & Zamuner, 2007). The SFH (Keating, 1984; Steriade, 1995; Wetzels & Mascaró, 2001) proposes that the underlying phonological features, in both aspirating and voicing languages, is [±voice]. In this view, aspirating and voicing languages have different mappings of [+voice] and [-voice] to different phonetic categories (i.e. lag/synchronous/lead VOT). The MFH claims that in aspirating languages, the feature specification is monovalent [spread glottis] while in voicing languages the relevant distinctive feature is monovalent [voice]. This predicts that in aspirating languages, [voice] is a passive feature, and the same applies to [sg] in voicing languages. There is evidence from English and German for voiced stops surfacing as voiceless when the context cannot supply a [voice] specification. Additional support for the MFH comes from L1 acquisition, as errors favor disproportionally the passive feature while early attempts from children favor the active one in each type of language.

According to the SFH, for the two-way contrast languages, there is no new distinctive feature to be acquired. The difficulties for the L2 learner will be in the phonetic mapping. On the other hand, the MFH assumes that moving from an aspirating to a voicing language, the learner needs to realize at some point that the [sg] feature that she has been using for the L1 is inadequate in the L2. Once the right feature is in place, the phonetic implementation should be unproblematic. The SFH presents no inherent impediment to continued progress extending all the
way to the advanced levels and consequently needs to rely on external factors to account for ceiling effects. The MFH contemplates the possibility of learners never making the switch and continuing to use the wrong feature regardless of the amount of input.

In the SFH, the weight of explanation is shifted toward the acquisition of phonetic mappings. Although seldom explicitly mentioned, this has been the theoretical assumption behind most of the previous L2 VOT literature. On the other hand, the MFH opens the door for the acquisition of phonological features.

The question of what is acquired also has a different answer when we switch our focus onto perception. In perception, L2 learners need to modify the perceptual space and its categories. However, different languages may have different boundaries for what can essentially be considered the same category. These boundaries are relevant because we tend to have high sensitivity for distinguishing two token sounds that belong to different phonemes, but poor within-category discrimination. In classic experiments that synthesize a series of stops varying only in VOT in equal steps and then ask listeners to identify the token sound as voiced or voiceless, it is reported that English speakers normally switch labels around the 30 ms mark while for Spanish speakers the boundary seems to be around 15 ms. This is again something that the L2 learner needs to acquire to perceive the intended phoneme in the target language.

The how and when of VOT acquisition

For perceptual categories, there are again two main positions: One can be considered emergentist, and the other innatist. The debate between these two camps started with Chomsky’s rebuff of Skinner’s views and the proposal that language is a specialized innate module. From a generative perspective, the acquisition of a linguistic contrast is the result of the input activating innate options. This view is better articulated by the Principles and Parameters Theory, which assumes that the innate language module provides a set of parametric options and learners need to discover which of the options matches the inputs of the language context. The most recent development of this position in phonology is Optimality Theory (OT), which posits that the innate module provides a set of universal constraints and the role of the learner is to discover their relative weight. Since OT is a theory of phonological competence, it is not clear that it should be applied to the acquisition of a phonetic feature such as VOT. Nonetheless, Stochastic Optimality Theory combined with the Gradual Learning Algorithm (Boersma, 1997) is a proposal that can handle the formation of phonological and phonetic categories. In this approach, acquiring VOT entails demoting constraints of the type ‘VOT = 100 ms’ through an error-driven algorithm. This demotion would occur when the context language doesn’t have tokens with 100 ms lead VOT. Since the phonetic tokens frequent in the language will not generate mismatch errors when produced by the learner they will not be demoted and will end up with a high
weight. The task of the learner is to rearrange the universal constraints starting from the L1 settings.

From the emergentist point of view, categories and boundaries emerge spontaneously from interaction with the language. Perhaps the most influential model is Patricia Kuhl’s Native Language Magnet (NLM). Kuhl et al. (2008) assume that each perceived token is stored in a perceptual space. With enough hits in that space, clusters with denser traces emerge. In time, these clusters develop exemplars that correspond to the areas of peak density. In perception, these exemplars warp the perceptual space in the sense that they attract toward the center tokens that are experienced toward the margins of the cluster. This magnet effect is thought to be behind the perceptual filters already mentioned by Trubetzkoy (1939/1969) and is able to account for categorical perception. In L2, perceptual filters are thought to be an important factor to explain fossilization and ceiling effects. Another model that competes to account for the acquisition of L2 phonetic categories is the Perceptual Assimilation Model for L2 (PAM-L2) (Best & Tyler, 2007). The main goal of this model is to predict when new categories will be created and when existing categories will interfere with the formation of L2 categories. Some possibilities of the mapping between two languages are multiple assimilation, category merging, or category splitting. Accounting for the mapping of categories between L1 and L2 is also one of the main goals of the Speech Learning Model (SLM; Flege, 1992). This model predicts that an L2 sound that is sufficiently dissimilar to any sound in the L1 will be easier to acquire than one that is too similar to an L1 sound. In these cases, there will be equivalence classification and a single phonetic category merging the perceptually linked L1 and L2 sounds will be formed.

For production, the innatist camp assumes that phonological features are innate and that what is learned is their distributional properties and phonetic mappings. Regarding phonetic categories, Articulatory Phonetics (Browman & Goldstein, 1992) works with gestural scores that control the exact timing of gestures. As we saw above, VOT is a matter of gestural timing, which makes this theory very well suited to account for the acquisition of VOT.

An interesting alternative that links perception and production is the Motor Theory of Speech Perception (Liberman & Mattingly, 1985). According to this theory, what we perceive when decoding speech are the same gestures, and gestural scores, we use in production. From that perspective, the right features of VOT will only be perceived adequately once the speaker develops the correct gestural scores in their production. This hypothesis has obvious implications for the teaching of pronunciation that have, so far, received very little attention.

Finally, the Ontogeny Phylogeny Model (Major, 2001) constitutes an attempt to define stages in the process of L2 acquisition. According to this model, early stages are characterized by a heavy L1 influence. As L2 proficiency increases, UG (universal grammar) starts to play a bigger role only to decrease in later stages as the L2 settings take over both L1 and universal settings.
What VOT research tells us about the advanced learner

Some early experimental studies focused on the limitations of advanced L2 learners in producing and perceiving VOT stops (e.g. Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973). Flege (1980) examined Saudi Arabic speakers learning English in the United States and was one of the first studies to call our attention to VOT ceiling effects. This study focused on the role of length of residence (LoR) and concluded that extended LoR is not a reliable predictor of native-like VOT. Progress in VOT stalls after reaching the advanced level and it seems that additional input does not have the same effect. The low number of participants (6) takes away from the robustness of the claim, but subsequent studies have reported similar results.

Advanced L1 English learners of French performed better than intermediate counterparts in Flege and Hillenbrand’s (1987). However, this applies only to French sounds that were similar to the corresponding L1 sound. In spite of being better, the advanced learners only approximated the target sounds. For sounds that do not have a correspondent in the L1, intermediate and advanced learners were equally close to the target. This suggests that different sounds present different degrees of difficulty in their acquisition, and that for some sounds acquisition is not difficult.

In terms of perception, Flege and Eefting (1987) examined identifications of a VOT continuum by Dutch learners of English. All learners shifted the VOT boundary when going from English mode to Dutch mode, but the shift was greater for the advanced learners. Interestingly, though, VOT values were also shorter in Dutch mode for the advanced learners. The authors’ interpretation was that advanced learners had created two independent phonetic categories in the L1 and L2. Similar results are reported in Flege and Schmidt (1995): in an identification task, Spanish/English bilinguals more frequently label as /p/ tokens with shorter VOT than monolinguals do. Finally, Flege and Eefting (1988) also concluded that the most advanced learners (early learners) had acquired an additional phonetic category that monolingual and late learners did not have. This corroborates previous findings showing that some advanced learners can create new phonetic categories for the L2.

Flege (1991) focused on age-related constraints on the ability to produce native-like VOT. Working with Spanish L1 learners of English, he found that early learners can produce VOT stops with timings close to the monolingual targets, but not late learners. The SLM evolved as an attempt to provide a theoretical model to account for the significant limitations that advanced L2 learners face to produce and perceive target sounds in a native-like fashion. Since this model claims that phonetic categories are not stable-state constructs but rather entities that evolve over the life span, it opens the door to the possibility of highly advanced learners.

Fowler, Sramko, Ostry, Rowland, and Hallé (2008) report that both L1 and L2 VOT measurements are different from that of monolinguals. VOT was found to be shorter than that of English monolinguals, and have a longer lag than that of French monolinguals. Major (1992), examined an English speaker who had been living for 12 years in Brazil at the time of the study and found that her VOT values in English had moved a bit closer to the values of the L2.
In a study examining heritage speakers of Italian, Russian, or Ukrainian living in an English context, Nagy and Kochetov (2013) found that for Russian and Ukrainian speech VOT values “drift away” toward the long lag of English, but not as far as Italian.

In the context of study abroad, Nagle, Morales-Front, Moorman, and Sanz (2016) were able to track improvements for English speakers in the short span of an intensive summer program in Barcelona. Not only did timings move in the L2 direction, but also the standard deviation of the VOT values decreased, indicating consolidation of the L2 category.

Regarding adult learners in the classroom setting, Zampini (1998) examined L1 English advanced learners of L2 Spanish in a phonetics course designed to improve their pronunciation. She reported VOT boundaries that were almost target-like. In terms of production, many learners were almost on target by the start of the study and practically all were reported to reach nearly target-like productions by the third session. Due to the specific design, any development that had happened before data collection was not captured in this study. Nathan (1987) had the reverse L1/L2 pair and was able to report a progressive increase in VOT lag. Reeder (1998) had four proficiency levels (including upper level and instructors), and reported that in terms of perceptual boundaries, all groups used native-like boundaries. However, in terms of production, VOT values decreased with amount of instruction, with the sharper decline taking place at the beginning of the advanced level. Kissling (2012) also studied English speakers learning Spanish and had a cross-sectional design that allowed her to corroborate the progressive decline in VOT lag. González López (2012) examined 16 English speakers at the intermediate level reading and switching from English to Spanish or vice versa. Her findings show that VOT values were significantly different for each language. As expected, VOT values in Spanish were not monolingual; interestingly, the same applies to English. Finally, Nagle (2015) also examined L2 Spanish but this time in a longitudinal study. His design allows the author to track not only the progressive decline in VOT times for the group, but also the individual trajectories—which as expected turned out not to be uniform. This study reports not only gradual improvement for VOT, but a reduction in dispersion that is interpreted as stability and solidification of the categories.

There are articles, such as Bohn and Flege (1993), that argue that VOT is not always important for the perception of voicing. Then, there also others, such as Curtin, Goad, and Pater (1998), that demonstrate that even for speakers of aspirating languages (English in this case), learning stop contrasts in Thai that rely on [voice] is easier than learning equivalent pairs of the same language that use [sg]. This may be open to many interpretations, but none that is consistent with the pre-eminent role assigned to VOT.

Although most previous studies find gradual linear progress that moves VOT timings and boundaries from the L1 to the L2, a couple of studies (Fourakis & Iverson, 1985; Gass, 1984) present cases that challenge this general picture. On the one hand, Gass presented cases of non-native speakers exaggerating the difference between their L1 and English—basically overshooting lag values of the English
target. On the other, Fourakis and Iverson observed that with Arabic-accented English the VOT values were also exaggerated but this time in favor of the L1—making it sound more Arabic than Arabic itself.

Outside of these cases of reverse directionality, the major exception to the idea of an uninterrupted progress comes from the reported ceiling effects. It seems as if L2rs encounter a ‘block on the road’ of sorts, and as a result, there is a gap between expected and actual VOT values.

**Explaining the gap**

For some researchers, the ‘block on the road’ has a biological origin; for others it is a matter of access to UG; and for yet others, it is the result of several interacting factors that combined make progress increasingly harder.

The biological explanation is tied to the concept of a critical period. In the last century behaviorists used this term in the study of bird development to refer to the window of opportunity that the juveniles have for learning their species’ songs. Lenneberg (1967) was the first to apply this concept to SLA by claiming that “foreign accents cannot be overcome easily after puberty” (p. 176). From the early days, this has been a highly controversial claim. We don’t have here the space to review the debates so we will just substitute the statement with an alternative one that seems more in line with current SLA thinking: “the human ability to learn never stops; but throughout our lives learning gets harder.” As learned routines, concepts, or categories become more ingrained in our minds, it becomes harder to change them. These two statements imply opposite approaches to accounting for learning limitations. In one case, the learning gate closes at puberty and some exceptions need to be implemented to handle any new learning; from the other perspective, since the gate never closes, some restrictions are needed to account for learning limitations that come with age.

Some studies propose that incomplete acquisition happens when the quality of the input is insufficient (e.g. MacKay, Flege, Piske, & Schirru, 2001). This seems unlikely to be the main factor in the case of VOT. Perhaps some cases could be explained by appealing to inadequate input but it cannot be extended to the majority of L2rs. Another common source of explanation for failures in acquisition is attributed to the amount and quality of L2 usage (e.g. Flege, Frieda, & Nozawa, 1997; Piske, MacKay, & Flege, 2001). Again, this could be plausible if we were dealing with some measurable number of cases of failed acquisition, but again, failure appears to be the norm and not the exception.

Perhaps the most common explanation for the lack of success in learning phonetic categories is the assumption that learners are not able to perceive the new, or slightly different, phonetic categories of the L2. This is normally connected to the concept of a perceptual narrowing. The basic idea behind perceptual narrowing is that there is significant synaptic pruning that takes place throughout our lives and starts very early in infancy. This pruning is governed by Hebbian learning and is an adaptive response of the brain that keeps and
strengthen connections that get activated and discards others that do not. The conditions of the environment are then the ones that condition neural development, which in turns conditions learning. This adaptive-neural pruning explains the perceptual tuning to the sounds of the environment that has been so often reported in studies of infant perception (e.g. Kuhl, 2010). The same perceptual tuning that makes us so efficient at perceiving the sounds of our L1 in their multiple variants is what gets in the way of perceiving new contrasts in the L2. Perceptual tuning consists in learning to ignore irrelevant phonetic cues. We ‘normalize’ the inherent variability of the auditory signal. This is in part what allows us to perceive constant phonemes in spite of the great inter- and intra-speaker variability that characterizes actual phonetic realizations. In cases were the relevant features have been tuned out by perceptual narrowing, the L2 system may result in being difficult to learn if it relies on the cues that that the speakers of the L1 have learned to ignore.

Khul’s Native Language Magnet, Best’s Perceptual Assimilation Model, and Flege’s Speech Learning Model all share a basic connection to perceptual narrowing. The difficulties that Flege’s model predicts in the case of similar but not identical sounds are the result of perceptual tuning. The magnet effect that predicts that token sounds that are part of the same category will not be perceived as different is the result of learning to ignore some differences. Finally, Best’s predictions about category merging also derive from the acquired ability to ignore information.

Importantly, acknowledging perceptual difficulties for acquisition does not account in itself for the ceiling effects observed for VOT. In fact, a good number of experimental studies with advanced late learners have shown that adults can learn to perceive difficult contrasts and acquire the automatized motor coordination for difficult sounds (e.g. Bongaerts, Mennen, & Slik, 2000).

Consequently, we still don’t have a satisfactory explanation for the kind of incomplete acquisition reported in previous studies. The tenets and hypotheses of the SLM can be put to work toward the goal of articulating a plausible answer compatible with the assumption that there is no hard barrier or gate that blocks learning at some early age. The SLM predicts that in cases were the sounds of the L1 and L2 are similar, equivalence classification can result in category assimilation. This means that no new category is formed to accommodate the slightly different L2 category. Instead the L1 category becomes the base from which to approximate the L2 targets. This is accomplished via realization rules (Flege, 1991, p. 407). Realization rules account for the fact that speakers can accommodate their own pronunciation to match the needs of the social context, or change their pronunciation in accordance with speech rate, noise, distance, emphasis, and so on. The idea here is that this ‘stretching’ of the L1 category via accommodation has limits. As a consequence, without creating the actual L2 category, the learner will only be able to approximate the L2 target.

This is a sensible explanation but, besides opening additional questions such as why realization rules have limits, it misses at least one important aspect of what is reported in the literature. It is not only the L2 category that is not equivalent to the
monolingual target; the L1 category also suffers some shifting in the direction of the L2. An additional problem with this sort of explanation is that the issue of phonetic category formation is most likely not a binary proposition. That is, we cannot simply assume that some learners form the right new L2 category while others are stuck with the L1 category and try to ‘stretch it’ to approach the L2. In fact, if the formation of categories is something gradual then there must be all sorts of intermediate stages.

We believe that perceptual narrowing and category formation are on the right track when it comes to explaining many of the age-related issues in the acquisition of L2 phonology and phonetics; we just do not think that those concepts should be extended to account for VOT ceiling effects. If perceptual filters were the main reason for fossilization, we would expect to see the blocking effect mostly in the early stages of acquisition. Then, as the learner overcomes the perceptual difficulties, with the help of explicit instruction or implicitly with enough quality input, learners should continue to show slow, uninterrupted progress. The fact that there is no evidence of generalized VOT problems due to perceptual difficulties in the early stages suggests that the reported ceiling effects may have a different origin. Similarly, if the inability of some learners to form the right categories for the L2 were the main culprit, we would again expect to see stagnation early on from the very moment L2 categories start to form. We would not predict an outcome where the majority of learners stop learning so close to the native target.

The explanation that we would like to advance here is that the apparent ceiling effects among advanced learners do not result from perceptual errors. In fact, the literature shows that from a very early stage L2ers modify their perceptual boundaries in the right direction and can quickly produce VO timings that are very close to the native target. Gass (1984) found evidence that they can do better than native speakers. When people try to imitate other accents, the adoption of extreme VO timings in the direction of the L2 is again a salient feature. The modification of the perceptual boundaries and the production timings is in fact something we can do on the spot to accommodate our listeners or the social context. The main problem with ceiling effects does not seem to be that perceptual tuning precludes L2rs from noticing the difference and hence adopting the right perceptual boundary or creating the new appropriate category. Instead, what happens is that L2rs settle on the typical bilingual VO timings and those timings are normally a compromise between the L1 and L2 (see among others Avelló & Lara, 2014; Caramazza et al., 1973; Flege & Eefting, 1987; Gass, 1984; Keating, Mikoś, & Ganong, 1981; Kollias, 1993; Raphael et al., 1995; Recasens i Vives, 1986; Yeni-Komshian, Caramazza, & Preston, 1977). These compromise timings are arguably the result of general efficiency constraints on language usage. Under this account, there is no incomplete acquisition and no learning failures or shortcomings; no real ceiling or block, and no fossilization. It is not the case that additional input has no effect on learning; it is rather that the learning goal has already been reached and we were all along assuming the wrong targets (see Cook, 1997, or Grosjean, 1998).
Future research

In spite of the large number of previous studies on the acquisition of VOT, there are still lacunae that should be taken into account in future research. Most of the previous research has focused on the acquisition of voiceless stops in utterance or word-initial position. Voiced and voiceless obstruents are also contrastive in non-initial position and studying in depth those contexts will complement what we already know about initial position.

It can also be argued that it is time to broaden current inquiry to include phonetic features other than VOT that contribute to laryngeal contrasts. Voicing, aspiration, articulatory effort, duration of closure/explosion, intensity of the stop burst, duration of the following vowel, and \( f_0 \) transitions are some of the phonetic dimensions that have been kept in the background while the shining light has been following VOT. The current fixation in measuring VOT values misses an important point: no single cue can determine how the voicing of obstruents is perceived and produced. It is the interaction of multiple cues and factors that determines how these contrasts are implemented and perceived.

There is also a clear need for a better understanding of contextual factors such as the role of stress, point of articulation of adjacent consonants, sonorancy, duration, or the quality of the following vowel. The same goes for stylistic factors such as speed or emphasis; social factors such as context of learning; and psycholinguistic factors such as motivation or working memory.

Although we have learned a great deal about VOT by studying categorical learning and measuring VOT values, future research should not overlook the possibility that the acquisition of laryngeal features (e.g. \([\text{sg}]\) and \([\text{voice}]\)) at the phonological level may play an important role.

Finally, future research should do a better job at trying to capture the whole picture. It is shortsighted to study voiceless stops while ignoring what happens with the parallel voiced stops. Meillet’s famous quote referring to language as “un système où tout se tient” (a system where everything stands) should give us an idea of how woefully inappropriate it is to focus on just one corner of the whole system. If \(/t/\) is defined as what is not \(/d/\), not \(/p/\), not \(/k/\), not \(/s/\), and so on, then the properties and values of \(/t/\) will depend on the values of \(/d/\) and so on, and to acquire the values of one we need to have the appropriate values of the other contrasting phonemes.

In the classroom setting, it would be interesting to see more longitudinal studies, more studies using tasks that do not require an immediate response, and more studies that rely less on controlled lab conditions and instead work with data from spontaneous speech in different learning settings. In this context, empirical research should start to provide the basis for the guidelines that should inform teaching. In the most general terms, for the perception side we need to counter the effects of perceptual narrowing and for the production side we need to work toward attaining automatic motor coordination in the control of gestural scores.
For motor control and coordination, the basic approach has to rely on oral practice. The question is, what kind of practice is most efficient? Currently instructors rely on intuitions, personal experience, and beliefs to decide what activities will help students. However, these decisions should be based on empirical research and not just common sense or personal biases. In that sense, we would like to see more studies attempting to answer questions such as: Is reading aloud in class beneficial for eliminating aspiration? Are whole-class repetitions any good? What about recitation of poems, sections of a story, or a play? Are tongue-twisters worth the trouble? Does explicit knowledge about the articulation of sounds or their acoustic properties help to overcome problems with foreign accent? Should we encourage mimicry or the adoption of different personas?

For perception, the main goal is to get learners to focus on cues that would normally be filtered out by the influence of the L1 perceptual system. In the case of English, for instance, aspiration and vowel duration are important cues, but for Spanish the contrast between /p/ and /b/ relies heavily on true voicing, and focusing on vowel duration does not help.

Future research should be able to better assess the effect on acquisition of computer-assisted visualizations of sound properties. How much effort should be devoted to analyzing your own recorded voice? How much to the study of the accented speech of other learners? How important is it to learn phonetic transcription to be able to capture important details in those analyses? If the teaching of pronunciation is to be raised to the level of importance of other components such as the lexicon, morphology, or syntax, we should start by documenting empirically the effects that different interventions have on the acquisition of the different sounds and contrasts.

In this chapter, we have argued that reaching a highly advanced level of proficiency with the laryngeal features of obstruents is possible. Given the difficulties that a limited amount of time and a limited amount of input pose to L2rs, we are calling for research that not only has a broader scope but also starts the arduous task of determining empirically what the best and most efficient practices are to help learners reach advanced proficiency.

**Conclusion**

We have seen in this chapter that VOT has been a hot topic of research in the past five decades. VOT has been used to argue both in favor and against the Nativist Hypothesis, the Critical Period Hypothesis, and incomplete learning. Given how easy it is to measure and what a central role it has taken over the years at the expense of other laryngeal features, VOT has been seen as a valid proxy for pronunciation and foreign accent in general. We have argued that this focus on VOT is questionable and that future studies should broaden the scope of the research agenda. We have devoted a section of the chapter to reject the notion that VOT is an example of a feature that is difficult to master and that results in incomplete acquisition, ceiling effects, or fossilization.
REFERENCES


Part IV    Advanced Grammar
Advanced-Level Mood Distinction

AARNES GUDMESTAD
Virginia Polytechnic Institute and State University

Introduction

The mood system is a component of language that has been of great interest to linguists (e.g. Palmer, 2001; Rothstein & Thieroff, 2010), including those in second language acquisition (SLA). Although its semantic function has been debated, Thieroff offers the following description of mood:

Mood, or, more precisely, morphological mood … is a morphological category of the verb, just as are the verbal categories person, number, aspect, tense, and voice. Mood categories express modalities such as orders, wishes, (non-)factivity, (non-)reality and the like. (2010, p. 2)

To date, much of the SLA research on mood distinction has focused on the subjunctive mood in second language (L2) French and Spanish (e.g. Ayoun, 2013; Collentine, 1995; McManus & Mitchell, 2015; Montrul, 2011). This body of research has included analyses of adult learners at different proficiency levels and has assessed learners’ ability to both use and interpret verbal moods. It spans various theoretical and analytical approaches and consists of data collected from different types of tasks. This chapter includes a concise overview of mood distinction in French and Spanish, a discussion of advanced-level mood distinction based on existing empirical findings, and suggestions for future research.

Mood distinction in French and Spanish

In Spanish and French the subjunctive–indicative alternation occurs primarily in subordinate clauses, which suggests that L2 learners need to be able to produce
and interpret complex syntax before they can acquire mood contrasts (e.g. Collentine, 1995). Since the indicative mood is the dominant mood in both languages, the main acquisition challenge seems to be that learners must learn to identify appropriate contexts for the subjunctive. The linguistic contexts in which the subjunctive occurs—or the functions it serves in each language—have received the attention of linguists for years. While scholars have offered and debated various accounts of mood, including semantic and pragmatic ones (e.g. Confais, 1990; Haverkate, 2002), reference grammars (e.g. Butt & Benjamin, 2011; cf. Poplack, Lealess & Dion, 2013) often explain the use of the subjunctive according to an established group of semantic or lexical triggers (e.g. quiero que in Spanish, je veux que in French, both meaning ‘I want that’). Not only do learners need to acquire the contexts in which mood contrasts occur, but they must also develop the ability to use and interpret morphological verb endings that correspond to both verbal moods. With the simple present, for example, the morphological endings differ between the subjunctive and indicative for all verb classes and all person/number forms in Spanish. One example in the first-person plural is cantar ‘to sing’: cantamos [IND] and cantemos [SUB]. For some verbs in Spanish the stem also differs between the two moods. For instance, the third-person-singular forms of ser ‘to be’ are: es [IND] and sea [SUB]. However, in French, most present-tense verbs do not carry morphology that overtly denotes mood. The distinction, which is evident more often in written than oral French, is limited to morphological endings of certain verb classes and person/number forms. Danser ‘to dance’ is a regular verb that overtly marks mood in the first-person plural: nous dansons [IND] and nous dansions [SUB]. The distinction is also explicit in the stems of suppletive verbs. One example, shown here in the first-person singular, is faire ‘to do/make’: je fais [IND] and je fasse [SUB].

In addition to the learnability issues of using and interpreting subjunctive verb morphology in appropriate contexts, the acquisition of mood distinction is perhaps made more challenging by two other characteristics of this linguistic phenomenon. First, in comparison to the indicative, the subjunctive occurs infrequently in both languages. In their analysis of the Corpus del Español (Davies, 2002–), Biber, Davies, Jones, and Tracy-Ventura (2006) reported that the subjunctive constituted 7.2% of the verb forms used by native speakers (NSs) of Spanish. Similarly, Ayres-Bennett, Carruthers, and Temple (2001) observed that NSs of French produced no more than 5–10 contexts for the subjunctive in 30 minutes of speech. Given the ambiguity between most present subjunctive and indicative forms in French, the number of overt subjunctive forms that NSs use would most likely be even fewer than the number of subjunctive contexts they produce. Another acquisition challenge for learners is the variable nature of mood distinction. Mood distinction has been undergoing change for centuries in both languages (Harris, 1974), with French considered to be further along in the process (Obaid, 1967). This change in progress is generally characterized by a gradual decrease in the use of the subjunctive, coupled with an increase in indicative use. It explains at least in part why researchers have documented variable use of verbal moods among NSs of French (e.g. Gudmestad & Edmonds, 2015; Poplack, 1992, 2001) and Spanish.
Variable mood distinction means that NSs do not use the subjunctive and indicative moods categorically in many linguistic contexts, even though prescriptive grammars may state that use of a given mood is obligatory (cf. Poplack et al., 2013). Thus, the input that learners receive regarding mood distinction likely consists of infrequent subjunctive forms and variability in the use of verbal moods.

Advanced-level learners and mood distinction

SLA research on both languages has found that the subjunctive–indicative alternation is acquired late in the developmental process (e.g. Bartning & Schlyter, 2004; Geeslin & Gudmestad, 2008). Researchers have typically made this observation after analyzing data on mood distinction from learners who are considered to be at an advanced proficiency level but who do not exhibit target-like mood contrasts (e.g. Gudmestad, 2012; Howard, 2008; McManus, Tracy-Ventura, Mitchell, Richard, & Romero de Mills, 2014). A thorough examination of research on mood distinction in L2 French and Spanish, though, reveals that researchers use different benchmarks to classify learners as advanced and that the advanced-level groups are often comprised of learners whose experience with the target language differs in some ways. For instance, groups of advanced-level learners sometimes consist of both undergraduate- and graduate-level participants (e.g. Montrul, 2011) or of learners who have spent varying amounts of time in the target-language environment (e.g. Sánchez-Naranjo, 2009). Other means of categorizing learners by proficiency have included course level and internal proficiency tests (e.g. Ayoun, 2013; Gudmestad, 2006), and in some cases details on the criteria employed were not provided (e.g. Borgonovo, Bruhn de Garavito, & Prévost, 2005). While this makes it difficult to make generalized, empirical observations about what constitutes advanced-level mood distinction, the fact that an assortment of criteria has been used to characterize advanced proficiency is not unique to research on L2 mood distinction in Spanish and French. Hulstijn, Alderson, and Schoonen (2010) argue:

> we must equally recognise that especially SLA, but language assessment as well, have to date operated with notions of development and levels of development which have all too frequently been hopelessly imprecise. (p. 16)

Although classifications of advanced proficiency in research on mood distinction in L2 French and Spanish are diverse and at times vague, four different ways of identifying advanced-level learners prevail. They are based on the characteristics of the participants who have taken part in previous research on the subjunctive-indicative contrast and are not mutually exclusive.¹ This body of research suggests that learners may be considered advanced if:

1. they are undergraduate-level learners who have spent an academic year abroad in a target-language environment;
Advanced Grammar

2. they are graduate-level learners;
3. their proficiency has been rated as advanced according to an established proficiency metric; and/or
4. they are categorized as near-native speakers based on a series of characteristics.

The first three definitions were chosen because multiple previous investigations classified learners as advanced based on these criteria. Although the fourth definition is limited to one study, it provides an opportunity to examine non-native speakers (NNSs) whose primary experience with the language was naturalistic. In contrast, many of the learners who correspond to the other definitions of advanced proficiency were individuals who have had extensive experience with the target language in an instructional setting. For each of these possible characterizations of advanced proficiency, this chapter offers a description of the participants and the data-collection instruments and an examination of the ways in which these learners use, select, and interpret verbal moods in target-like and non-target-like ways. In light of the plethora of practices that SLA researchers use to evaluate L2 proficiency and the challenges that these varied assessments bring to the field’s ability to build comparable and generalizable findings (e.g. Thomas, 1994), the classifications of advanced proficiency are largely kept separate in this chapter.

Academic year abroad

A few investigations have examined mood distinction among undergraduates who have spent an academic year (about nine months) in the target-language environment. This section starts with a brief summary of the participants and the data analyzed in each of these studies. Next, the frequency with which these learners produced subjunctive verb forms and the linguistic contexts in which the subjunctive occurred are examined.

Beginning with the research on L2 French, Howard (2008) investigated subjunctive use among six English-speaking Irish undergraduates (see also Howard, 2012). They had spent a year abroad studying at a university in France. During this time they took courses with NSs of French, rather than language courses, and lived in dorms with NSs. Prior to their stay abroad, they had studied French for eight to nine years. The participants completed an hour-long oral interview. In another study, McManus and Mitchell (2015) reported on 29 learners who had also spent an academic year in France. They were all third-year undergraduate students from the United Kingdom and had an average of 10.4 years (range: 8–15 years) of instruction in the target language. Most were first language (L1) English speakers. At the end of their stay abroad, all participants completed an oral interview, an argumentative essay, and a grammaticality judgment task (GJT) that focused on mood distinction. This group of learners was further divided into two groups (low- or high-scoring) based on their score on an elicited imitation task that the researchers used as an oral-proficiency measure.
Isabelli and Nishida (2005) and Bonilla (2015) are two studies that have investigated mood distinction in L2 Spanish after an academic year abroad. Isabelli and Nishida (2005) examined oral interviews from 29 American L2 learners of Spanish who spent the third year of their undergraduate education studying in Barcelona, Spain. Before going abroad they had taken two years of beginning- and intermediate-level coursework in Spanish at a university in the United States. They were all L1 speakers of English. Bonilla (2015) examined seven L2 learners of Spanish from the Spanish Learner Language Oral Corpora (http://www.splloc.soton.ac.uk/) who had spent a year abroad and had received approximately 895 hours of instruction. The oral data she analyzed consisted of an interview and picture-description task. Thus, in general these learners of French and Spanish were NSs of English and most of the data on mood distinction comes from production tasks. In terms of analysis, learners’ use of the subjunctive in these investigations was primarily compared to prescriptive norms.

A common finding among these studies is that after an academic year in a target-language context, learners use the subjunctive infrequently. Overall the six Irish learners in Howard (2008) produced 29 linguistic contexts where the subjunctive is required in French and just three verb forms that were unambiguously marked as subjunctive in these contexts. While each participant produced between three and six subjunctive contexts, the three subjunctive verbs in the dataset were produced by two of the six learners. No instances of the use of the subjunctive in indicative contexts were observed.

McManus and Mitchell’s (2015) participants produced the subjunctive infrequently in French as well. In terms of the rate of occurrence of subjunctive forms (calculated per 1,000 words) in the interview and the essay, these learners produced more subjunctive forms in writing than in speech. The researchers did not observe dramatic differences in subjunctive frequency between the low- and high-scoring groups. For example, the total number of subjunctive occurrences ranged from six cases by the low-scoring group (15 participants) in the essay to 24 cases by the same participant group in the interview. The rate of subjunctive occurrences for the high-scoring group (14 participants) fell in between the rates of the low-scoring group. However, McManus and Mitchell did find differences between the two participant groups when they examined rates of acceptable (subjunctive triggers with subjunctive verb forms) and unacceptable (subjunctive triggers with indicative verb forms or indicative triggers with subjunctive verb forms) use for mood distinction. They found that the low-scoring group produced higher rates of unacceptable than acceptable mood use in writing and in speech. The opposite trend was observed for the high-scoring group, who exhibited higher rates of acceptable mood use.

Turning to L2 Spanish, the highest rates of subjunctive use were observed with the learners in Isabelli and Nishida (2005). The researchers analyzed nominal, adjectival, and adverbial subordinate clauses that prescriptively require the subjunctive or allow both moods. Focusing on the data collected at the end of their stay abroad (at nine months), the findings showed that overall these learners produced the subjunctive about half of the time in relevant contexts (34/69 contexts).
Moreover, the researchers also observed variability among the participants. Not all learners produced subjunctive triggers (i.e. linguistic contexts for the subjunctive), nor did each participant use subjunctive verb forms. Instead, 21 (72.4%) of the learners produced subjunctive triggers and just over half of this L2 group (16/29 participants) used the subjunctive in these contexts. These subjunctive users each produced between one and seven forms during the interview and their accuracy rates for the subjunctive range from 0 to 100%. Although less detailed information regarding mood distinction is available on the learners in Bonilla (2015), the results of this study suggest that this learner group exhibited a less productive use of the subjunctive than those in Isabelli and Nishida’s study and that there were differences among the participants. Bonilla found that, while the learners she analyzed produced sentences with subordinate clauses, they rarely produced subjunctive triggers and subjunctive verb forms. Even though detailed information about mood use was not provided for each participant, Bonilla reported that “Several learners are starting to produce subjunctive marking (e.g., Kyle and Esther) … One learner, Michelle … marked the subjunctive in 2/2 contexts” (p. 67).

A comparison of the results of these studies reveals that the frequency with which learners used the subjunctive varied by linguistic context. The learners of French in Howard (2008) produced 10 different triggers. Three were what he called “syntactic devices” (e.g. jusqu’à ce que ‘until’) and seven were “lexical verbs” (e.g. falloir que ‘to be necessary that’) (p. 192). The triggers with which this L2 group used the subjunctive were restricted to falloir que (subjunctive use in 2/3 contexts) and vouloir que (‘to want that,’ subjunctive use the one time it was produced). Although McManus and Mitchell (2015) did not investigate differences in subjunctive use across linguistic contexts on the interview or in the essay, they did find that on the grammaticality judgment test (GJT) their participants performed significantly better on affirmative contexts (e.g. vouloir que) than on adverbial conjunctions (avant que ‘before’) and significantly better on adverbial conjunctions than on negative triggers (ne pas penser que ‘to not think that’). They also compared the learners’ performance on the GJT to that of 10 NSs of French who completed the same task. In general the NSs scored significantly higher than the low-scoring and high-scoring L2 groups on the three aforementioned trigger types. The one exception to this finding was that the high-scoring L2 group performed similarly to the NSs with the affirmative triggers. Finally, although Bonilla (2015) did not report on the linguistic contexts in which the subjunctive occurred, Isabelli and Nishida (2005) analyzed three clause types. They found that, after a school year in Spain, learners of Spanish used the subjunctive most often in nominal clauses (67% or 8/12 cases), followed by adverbial clauses (49% or 26/53 cases), and least often in adjectival clauses (0% or 0/4 cases). The learners produced eight different nominal triggers (e.g. espero que ‘I hope that’) and six different adverbial triggers in the interviews, with hasta que ‘until’ being the most frequent.

Taken as a whole, it appears that L2 learners of French and Spanish who have spent a year of their undergraduate education in an immersion context, and who are considered to be at an advanced level of proficiency by the researchers, have not yet acquired the subjunctive-indicative contrast. It may be worth noting,
however, that the learners in these studies were primarily compared to an idealized prescriptive norm, rather than a native-speaker norm. This could mean that, while the learners are not in line with prescriptive grammar, they may exhibit some patterns similar to those of NSs, as was observed with the high-scoring group and the NSs in McManus and Mitchell (2015) who performed similarly on mood-distinction items in affirmative contexts on a GJT. Another important observation that emerges from these studies is the presence of individual variability. Subjunctive use was not uniform among learners in the same participant group, which suggests that some learners were further along on the developmental path than others.

Graduate-level NNSs

The largest body of research on (possible) advanced proficiency comes from analyses of graduate-level NNSs—participants who are currently enrolled or have completed a graduate program. This section consists of Universal Grammar (Iverson, Kempchinsky & Rothman, 2008), error-analysis (Cheng & Mojica-Diaz, 2006), and variationist (Geeslin & Gudmestad, 2008, 2010; Gudmestad, 2012, 2013, 2014, 2015; Kanwit & Geeslin, 2014) research. Given the methodological differences among them, they are discussed in turn.

In Iverson et al. (2008), they investigated a group of NNSs of Spanish (N = 18) who were professors and graduate students of Spanish. They had studied Spanish between 9 and 22 years and had spent between one and five years living in a Spanish-speaking country. Thirteen NSs also participated in the study, serving as a control group. The data on the subjunctive-indicative contrast came from a “scalar grammaticality judgment task testing for knowledge of the permissibility of indicative and subjunctive complement clauses with volitional and negated epistemic verbs” (p. 151). The participants rated how natural the sentences seemed on a scale of one to five. The instrument consisted of six sentence types. There were two sentence types in which the independent clause contained a volitional trigger. One of these sentence types had a subjunctive verb in the subordinate clause and the other had an indicative verb. The remaining four sentence types contained a negated epistemic trigger in the independent clause. Half of them were followed by a subjunctive verb in the dependent clause and the other half were followed by an indicative verb. These sentences also ended with a coordinated clause that showed that the speaker either committed to the truth of the subordinate clause or did not, which provided important discourse information that was necessary to judge these sentences. The results revealed that overall the NNSs and the NSs judged the sentence types similarly. However, in terms of the distinctions that the participants made between sentence types, the NSSs differed significantly from the NSs in “both their contrast of negated epistemics with indicative clauses in which the speaker commits to the truth of the embedded clause and those in which the speaker does not” (p. 155). The researchers concluded that this NNS group had near-native-like knowledge of the mood-distinction contexts under investigation.
While the aforementioned study in this section focused on assessing knowledge of mood distinction, Cheng and Mojica-Diaz (2006) were interested in examining the role that pedagogical intervention could play in accuracy rates of the subjunctive among advanced NNSs. The NNSs (N = 6) were Spanish teachers in secondary schools in the United States who were enrolled in a graduate program. They were all NSs of English. When the data were collected, they were participating in a nine-week intensive program in Mexico. They lived with Mexican families and were taking courses on Spanish grammar and L2 pedagogy. The data came from simulated oral proficiency interviews. Each NS completed one interview. The NNSs completed three interviews: at the beginning of the program (pre-test), at the end of the four-week course (post-test 1), and at the end of the summer program (post-test 2). In addition to analyzing L2 subjunctive use in the interviews, the researchers also reported that all of the participants were rated as being at the advanced level of proficiency according to the American Council on the Teaching of Foreign Languages guidelines (ACTFL). Three were classified as advanced-low, two as advanced-mid, and one as advanced-high. The researchers analyzed two types of subjunctive contexts: “complement clauses of a matrix expressing modality and the if-clauses referring to hypothetical situations” (p. 20). The results revealed that the NNSs’ accuracy rates in obligatory subjunctive contexts changed over time. At the pre-test, they used the subjunctive in 66.7% of subjunctive contexts. They showed improvement at post-test 1, producing the subjunctive in 87.2% of obligatory contexts. However, their accuracy rate fell to 68.3% at post-test 2. In terms of variable contexts, although the researchers observed a decrease in the use of the subjunctive over time (83.3% at the pre-test, 75.0% at post-test 1, and 67.9% at post-test 2), they noted that the NNSs produced very few contexts overall (two contexts at the pre-test and four contexts at each post-test). Accuracy rates of the subjunctive in hypothetical contexts also decreased over time (from 63.2% at the pre-test, to 27.2% at post-test 1, and 24.3% at post-test 2). These findings suggest that graduate-level NNSs who were rated as advanced-level speakers according to ACTFL did not exhibit sustained improvement in their use of the subjunctive after a nine-week immersion experience, which included formal instruction on mood distinction.

A majority of the recent studies that have examined graduate-level speakers and mood distinction have been informed by variationism (see Chapter 28, Variable Structures and Sociolinguistic Variation). What they all have in common is the inclusion of a group of NSs of Spanish who completed the same tasks as the NNSs. In light of the variable nature of mood distinction in Spanish, these investigations compared L2 data to NS data as a means of assessing target-like behavior rather than conducting error analyses based on prescriptive norms.

Similar to Iverson et al. (2008), Kanwit and Geeslin (2014) investigated the interpretation of mood distinction in Spanish (see Chapter 28). The graduate-level participants, who were NSs of English, were teaching Spanish at the university level while pursuing their graduate degree in Spanish. They had lived abroad for an average of 13.6 months (range: 6 weeks–3 years) and had studied Spanish formally for 11.5 years on average (range: 6–14 years). A group of NSs of Spanish (N = 16)
provided baseline data with which to assess the target-like nature. The data on mood distinction came from a written interpretation task. This instrument consisted of sentences that contained an independent clause with a present-indicative verb and a subordinate, adverbial clause with either a present-subjunctive or a present-indicative verb. Following each sentence, the participants were presented with two interpretations of the sentence: whether the event had yet to occur or was habitual. They could choose one of these interpretations or select an option that both interpretations were possible. The researchers designed the instrument to examine a range of linguistic and extra-linguistic independent variables. The findings revealed many differences between the NNSs and NSs, and Kanwit and Geeslin concluded that, “at the highest level of proficiency in the current study, our participants performed largely as would be predicted by prescriptive norms” rather than like the NS comparison group (pp. 527–528).

Continuing with variationist research, Geeslin and Gudmestad (2008, 2010) investigated the use and selection of verbal moods in Spanish. In Geeslin and Gudmestad (2010) one group was comprised of 16 NNSs who were NSs of English. They had studied Spanish formally for an average of 8.07 years (range: 3–17 years) and lived in a Spanish-speaking country for a length of time that ranged from three months to 11 years. The second group, NSs of Spanish (N=16), came from various Spanish-speaking places. All participants were graduate students, had experience teaching Spanish at the undergraduate level, and reported using Spanish socially. They completed a semi-guided interview with two NSs of Spanish. The researchers analyzed all finite-verb contexts that the participants produced in the interview and found that, while both groups used non-subjunctive forms (indicative and non-finite) at a higher rate than the subjunctive, the NSs produced more subjunctive forms (6.7% or 586/8,716 contexts) than the NNSs (2.0% or 124/6,065 contexts). In Geeslin and Gudmestad (2008) a subset of these participants (10 individuals in each group) also completed a written contextualized task (WCT) focused on the subjunctive–indicative contrast in the present tense. This instrument consisted of a short story separated into segments. Following each section the participants were presented with a sentence that served as dialogue in the story. Two versions of the sentence were provided with each context: one with the subjunctive and one with the indicative in the mood-choice context. Participants were asked to select which of the sentences they preferred or to choose that they like both. Similar to Kanwit and Geeslin (2014), the task was designed to analyze a series of linguistic factors. In terms of frequency of verbal moods, the NSs exhibited rates of the subjunctive on the WCT and the interview (WCT: 61.6% or 215/349 contexts; interview: 18.4% or 311/1,688 contexts) that were significantly higher than those of the NNSs (WCT: 49.1% or 172/350 contexts; interview: 6.6% or 75/1,141 contexts). However, even though these advanced NSs showed lower rates of the subjunctive mood than the NSs, they selected and used this verbal mood in most of the same linguistic contexts as the NSs.

In a series of studies, Gudmestad (2012, 2013, 2014, 2015) analyzed oral-production data from 20 NNSs and 20 NSs of Spanish. All were graduate students and most had experience teaching undergraduate-level Spanish. The NNSs were NSs of
English and 19 of them had lived in a Spanish-speaking country ($M=19.3$ months, range: 1.5 months–5 years and 1.5 months). The NSs came from different Spanish-speaking places and been in an English-speaking country for an average of 4.8 years (range: 1.8–14.5 years). Gudmestad (2012) examined the frequency and predictors of mood use and found that both participant groups used subjunctive forms (NSs 52.7% or 1,425/2,704 contexts; NNSs: 53.8% or 1,229/2,285 contexts) more than the non-subjunctive (i.e. indicative and non-finite forms produced in finite-verb contexts) forms (NSs: 47.3% or 1,279/2,704 contexts; NNSs: 46.2% or 1,056/2,285 contexts) in the mood-distinction contexts under investigation. The difference in frequency between the two groups was not significant. Moreover, all of the independent variables under investigation were significant for both participant groups. However, the direction of the effects for some of these variables differed between the two groups. These findings demonstrated that this group of advanced L2 speakers used verbal moods in largely native-like ways.

Continuing with this dataset, Gudmestad (2014) investigated the relationship between mood distinction and individual lexical triggers. One of the independent linguistic variables analyzed the frequency of verbal moods among NNSs according to the frequency with which the NS group used the subjunctive with different individual lexical triggers (see Gudmestad, 2014/2012). This variable consisted of two categories: high and lower subjunctive use. Limiting the analysis to a subset of the lexical triggers analyzed in Gudmestad (2012), she found that the rate with which the NNSs used the subjunctive with individual lexical triggers varied according to the frequency with which the NSs used the subjunctive with the same triggers. Specifically, the results demonstrated that the NNSs produced the subjunctive at a higher rate than the NSs with lexical triggers that exhibited lower NS subjunctive use. In contrast, their frequency of subjunctive use was lower than that of the NSs with lexical triggers that showed high subjunctive use among the NSs. Thus, although Gudmestad (2012) found evidence of near-native-like use for this L2 group when all of the data were analyzed together, Gudmestad (2014) indicated that an examination of individual lexical triggers could identify additional components of mood distinction that this advanced-level L2 group had yet to acquire.

Finally, while most studies on L2 mood distinction have focused on the subjunctive mood in opposition to the indicative, Gudmestad (2015) argued that, given the variable nature of mood distinction in Spanish, it was important to understand the full range of forms that NSs and NNSs use in mood-distinction contexts. In other words, the fact that NSs tend not to exhibit categorical subjunctive use in contexts where the subjunctive is possible means they are using other verb forms in these contexts. Gudmestad sought to identify these forms and their frequency of occurrence and to compare these findings to L2 use. She found that this group of NSs used no fewer than 23 verb forms (both subjunctive and non-subjunctive forms) in the mood-distinction contexts under investigation. These forms ranged from those that were very infrequent (e.g. pluperfect indicative, present progressive subjunctive) to those that the NSs used often (e.g. present indicative, imperfect subjunctive). The NNSs produced most of these forms as well.
Advanced-Level Mood Distinction

(20/23 forms) and they used them at a similar rate to the NSs. Furthermore, detailed analyses of the conditional and the imperfect revealed that the NNSs exhibited many of the same patterns of use as the NSs.

In a similar vein, because much of the L2 research has centered on the present subjunctive, Gudmestad (2013) aimed to expand this body of research by investigating the complete inventory of tense-aspect forms of the subjunctive that these two speaker groups used in mood-distinction contexts. This analysis revealed that this NS group produced eight subjunctive forms in this dataset with varying frequency. The present, imperfect, and pluperfect subjunctive forms constituted almost 98% of all subjunctive forms (1,395/1,425 forms), while the present perfect, present periphrastic, past progressive, past periphrastic, and present progressive subjunctive forms occurred infrequently (30/1,425 subjunctive forms). Although the NNSs produced no past progressive or past periphrastic subjunctive forms, their frequency of use of the other subjunctive forms was similar to the NS rates of use. Moreover, an examination of the three most frequent subjunctive forms (present, imperfect, and pluperfect subjunctive) indicated that the NNSs seemed to be more target-like with the present and imperfect subjunctive than with the pluperfect subjunctive. Thus, the analyses that move beyond a focus on the present subjunctive suggest that graduate-level NNSs used a range of verb forms in mood-distinction contexts in many of the same ways that NSs did.

In summary, although Cheng and Mojica-Diaz (2006) compared NNSs to a prescriptive norm, much of the research on graduate-level NNSs analyzed L2 data in relation to a NS comparison group. Collectively, studies with an NS baseline suggest that graduate-level NNSs tend to exhibit patterns of interpreting and using verbal moods that are similar to those of NSs. Although each of these investigations identified ways in which these advanced-level learners were not yet target-like with regard to the subjunctive–indicative contrast, the learners appeared to be able to make mood contrasts in a range of linguistic contexts.

Established proficiency metric

Another possible means of classifying learners as having achieved an advanced level of L2 proficiency is through formalized assessment metrics. A small number of studies have described learners’ proficiency based on the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001) or ACTFL guidelines (ACTFL, 2001). Beginning with the CEFR, two studies included learners who were considered to be advanced (Level C in the framework). Ahern, Amenos-Pons, and Guijarro-Fuentes (2014) analyzed learners’ interpretation of mood distinction in conditional statements containing if-clauses in Spanish. For this task, the participants were presented with 20 if-conditional and 10 distractor sentences. The subordinate clause contained the imperfect indicative or the imperfect subjunctive, and either the conditional or the imperfect indicative were used in the matrix clause. The participants were
asked to select one interpretation for each item: irrealis, habitual, or quotative-echoic. The L2 participants who corresponded to Level C were either NSs of French (N = 32) or English (N = 14). Their performance on the task was compared to a group of 35 NSs of Peninsular Spanish and to a group of intermediate-level (CEFR B2) learners, 16 of whom were NSs of French and 26 of whom were NSs of English. What is interesting to note about this study is that the researchers began their analysis by examining CEFR proficiency levels separately, but, because, they found that “overall proficiency level systematically undermines task performance” (p. 188), the researchers analyzed both proficiency levels together. This methodological decision prevents us from making target-like assessments of the Level C group alone. When analyzed together, these learners exhibited a mix of target-like and non-target-like interpretations. For example, they performed similar to the NSs on sentences with a habitual reading but differed from the NSs on items with a quotative-echoic interpretation.

Gudmestad and Edmonds (2015) also categorized L2 proficiency according to the CEFR, and in their investigation of mood distinction in L2 French, they analyzed learners who corresponded to different proficiency levels as separate groups. The Level C learners (N = 19) had been living in a French-speaking country between one month and 28 years. At the time of data collection all participants were enrolled in a degree program or intensive French courses at a French university. Their length of time spent studying French ranged from six weeks to 34 years. They were NSs of various languages. Target-like use was assessed by comparing the learners’ mood distinction to that of NSs who completed the same tasks. The NSs of Hexagonal French (that of France, N = 31) were students at the same university as the L2 speakers. The data on the subjunctive–indicative contrast came from two written tasks that were designed to investigate a series of independent variables: a contextualized-clause-elicitation instrument and a contextualized-verb-elicitation instrument. In terms of frequency of use of verbal moods, the results indicated that the NSs used the subjunctive more often (53.3%) than the Level C learners (36.9%). Next, Gudmestad and Edmonds distinguished between lexical triggers that exhibited categorical use of a verbal mood and those that were variable. The NSs used both moods with 19 of the lexical triggers included in the tasks. Their indicative or subjunctive use was categorical (i.e. 100% of the time) with six and five triggers, respectively. The NNSs were variable in their use of verbal moods with a higher number of triggers (28 triggers) than the NSs. They exhibited categorical indicative use with the remaining two triggers. The analyses of the independent variables revealed that this group of NNSs used the subjunctive in similar contexts to the NSs. Nevertheless, they had not yet acquired a target-like frequency of use of the subjunctive, which includes categorical use with some lexical triggers.

In addition to CEFR, two investigations have classified L2 learners of Spanish as having advanced proficiency according to ACTFL guidelines. While the focus of Asención-Delaney and Collentine (2011) was not mood distinction, they observed that learners (L1 English) enrolled in third-year Spanish courses at a university in the United States, whose writing abilities were rated as
advanced-low, used subjunctive verb forms in written narratives. Lastly, Cheng
and Mojica-Diaz (2006), which was previously discussed in the section on graduate-
level NNSs, found that advanced-level NNSs used the subjunctive less often at
the end of a nine-week immersion program than when the program started.
They used the subjunctive more often in obligatory and variable subjunctive
contexts than in hypothetical contexts. Taken together, these studies suggest
that learners who have been classified as advanced according to the CEFR or
ACTFL have the ability to use and interpret mood but they have not yet reach
target-like attainment.

Near-native speakers

The three previous characterizations of advanced proficiency focus on L2
learners in an instructional setting. While much of SLA research stems from
classroom learners, many NNSs around the world are predominately or exclu-
sively naturalistic learners. Medina-Rivera (2004) is one study that offers details
about mood distinction among NNSs whose current experience with the target
language appears to be naturalistic. The participants were “two non-Hispanic
clergy and two religious sisters (all of them dominant in English) who used
Spanish in their ministry within the Catholic church in Hispanic communities”
in the United States (p. 135). Each participant had received some instruction in
college but Medina-Rivera reported that most of their experience with Spanish
took place over years of interacting with NSs either in South America or in the
United States. While one participant had been speaking Spanish for four years
and working with Spanish-speakers for two years, the other three individuals
had more than 30 years of experience with the language and working with NSs.
The researcher classified the participants as near-native speakers because they
were individuals who speak Spanish as a second language, are capable of hav-
ing a complex conversation in that language, are able to understand any speaker,
and have the ability to function as professionals using Spanish in their field of
work (p. 133).

The data on mood distinction came from an interview. Medina-Rivera conducted
an error analysis to assess the participants’ use of the subjunctive but did not pro-
vide details on the linguistic contexts examined. His assessments of accuracy were
based on his intuitions as an NS, as well on characteristics of regional variation. The
results indicated that, while the participants used the indicative correctly over 99
percent of the time, their accuracy rates with the subjunctive differed. Individually,
they used the subjunctive correctly 36.4, 66.7, 85.7, and 100 percent of the time.
These accuracy rates did not appear to be strongly connected to their years of
exposure to the language. One participant (María), who had more than 30 years of
experience with Spanish, exhibited the lowest rate of subjunctive use. In contrast,
David, who had been speaking Spanish for four years, used the subjunctive cor-
rectly two-thirds of the time. In sum, Medina-Rivera found evidence of individual
variation in the use of the subjunctive among these near-native speakers.
Conclusion

This chapter has offered four possible characterizations of advanced proficiency, which stemmed from the participant characteristics of previous research on mood distinction in L2 French and Spanish. For each of these possible definitions of advanced proficiency, learners’ ability to make mood distinctions was examined. In general, each study found evidence of non-native-like mood distinction, though the degree to which each L2 group diverged from the target varied across studies. Some L2 groups appeared to be far along in their acquisition of verbal moods (e.g. the graduate-level participants in Gudmestad, 2012), whereas others used the subjunctive with such low frequency that it appeared that they needed quite a bit more experience with the target language before (near-)native-like abilities would be possible (e.g. the undergraduates who had spent a year abroad in Howard, 2008). At a quick glance these diverging observations may suggest that one characterization of advanced proficiency represents learners who are more advanced than others. Although this may very well be true, it is important to keep in mind, as noted earlier in this chapter, that the body of research on mood distinction in L2 Spanish and French is diverse. Not only are there notable differences among the participant populations examined, but researchers have also employed a range of tasks to investigate the use, selection, and interpretation of verbal moods. The means with which native-like mood distinction has been assessed differs as well. While some studies measured learner accuracy based on prescriptive norms, others assessed target-like mood distinction by comparing learners to NSs. These differences among studies within a single characterization of advanced proficiency and across all of the research examined in this chapter make it difficult to determine objectively whether the four descriptions of advanced proficiency converge on the same findings and therefore make it challenging to establish a single empirical profile of advanced-level mood distinction (cf. Thomas, 1994).

Given the diversity among existing investigations on mood distinction in L2 French and Spanish, a clear avenue for additional research is replication (specifically, approximate replication, cf. Abbuhl, 2012). In order to determine empirically whether there are commonalities in the ability to make mood distinction among the learners in these four classifications of advanced proficiency, future investigations need to replicate the tasks and data analyses employed in the study of a given learner population and investigate a different characterization of advanced proficiency. These types of replications will enable researchers to arrive at a clearer understanding of advanced-level mood distinction than is currently available in the literature. Additionally, there are other areas that are in need of investigation. Since much of the existing research has come from instructed learners who are NSs of English and has focused on L2 French and Spanish, future research should make an effort to examine naturalistic learners, L1s other than English, and additional target languages. This chapter was also limited to four descriptions of advanced-level learners, but there are other ways of assessing proficiency. Expanding the participant populations will further knowledge of advanced-level mood distinction as well.
NOTES

1. It is recognized that whether these learner groups are indeed at an advanced level of proficiency is an empirical question.
2. Many of these studies contain analyses of other learner groups who are not discussed in this chapter because they do not correspond to the definitions of advanced offered in this chapter.
4. The goal of this study was to track changes in learners’ use of the subjunctive over 21 months and six data-collection periods. Given this section’s focus on undergraduate-level learners who have spent an academic-year abroad, the discussion centers on the data that were collected at the end of their stay abroad.
5. See McManus, Tracy-Ventura, Mitchell, Richard, and Romero de Mills (2014) for further details on this proficiency measure.
6. Geeslin and Gudmestad (2008) also analyzed six independent extra-linguistic variables. Since small cells likely impacted the results of these factors, they are not discussed here.
7. Geeslin and Gudmestad (2010) analyzed all finite verbs in their analysis of mood use. In Geeslin and Gudmestad (2008) they examined dependent clauses with non-past reference only. See the articles for justifications of these dependent variables.
8. Each participant completed three mood-distinction tasks: a monologic role-play, a contextualized-clause-elicitation task, and a contextualized-verb-elicitation task. See any of these studies for details on the tasks.
9. See Gudmestad (2014) for findings concerning another independent linguistic variable—relative frequency of the lexical item, which provides additional details about the development of mood distinction among learners at different proficiency levels.
10. ‘Quotative-echoic’ is defined as “unique situations that were not (yet) accomplished at some specific point in the past” (Ahern et al., 2014, p.178).

REFERENCES


Gudmestad, A. (2014). On the role of lexical items in the second-language


Introduction

Within the realm of multi-layered concepts such as tense and aspect, learners can make use of frequently used morphological markers in association with specific verbs (e.g. perfective marker with telic events) or with specific narrative structures (e.g. imperfective markers with descriptions). The acquisition of such prototypical conceptualizations of tense and aspect has been the main focus of analysis of a number of empirical studies starting in the late 1980s (e.g. Andersen, 1986; Bardovi-Harlig, 1992). The use of concrete hypotheses and the collection of a significant body of empirical data for over 30 years have helped the field make significant progress on the analysis of tense-aspect development (see Salaberry & Comajoan, 2013, for a recent review). On the other hand, the majority of previous studies have primarily focused on the analysis of data from the initial stages of development. Spearheaded by Andersen’s hypothesis, by and large, it was assumed that once learners were able to overcome the influence of lexical aspectual values inherent in the lexicon, the acquisition of tense/aspect would proceed unimpeded given enough exposure to the L2. Alternative hypotheses proposed from within different theoretical models (i.e. Minimalism) did not focus on advanced levels of development either because, as a matter of theoretical premise, non-prototypical pairings of lexical and grammatical aspect (e.g. coercion as described in de Swart, 1998) were regarded to be outside of the grammatical system (i.e. they were considered part of the realm of pragmatics). As a consequence, despite the number of studies focused on the acquisition of tense and aspect, there is yet a lot to learn about the advanced stages of acquisition of temporality in L2 acquisition.
In the present chapter, I will review the various definitions of tense and aspect, concentrating primarily on the concept of aspect, given its inherent complexity for the development of advanced knowledge, to conceptualize and mark temporality in the L2. The main objective of this analysis is to show how the proper description of the concept of aspect (e.g. the limitations of imprecise and incomplete definitions) is crucial to both assess and understand the development of advanced knowledge in the use of tense-aspect phenomena. For the analysis, I will make reference primarily to Spanish data, because there is a substantial body of empirical data that can be used to evaluate the various theoretical proposals. I will then review two recent hypotheses that have focused on the study of the advanced knowledge of tense-aspect morphology. Even though each hypothesis is framed within a very distinct theoretical framework, they both describe the development of the advanced understanding of temporality in the L2 with reference to the use of complex conceptualizations. To conclude, I will reframe these two hypotheses within the perspective of the definition of aspect and temporality in general in light of the description to be presented in the first section of the chapter.

Aspect: From representation to acquisition

The representation of aspect

The standard definition of aspect highlights the effect of a broad range of contextual factors on the linguistic construal of a situation (see Comrie, 1976; Klein, 1994, inter alia for extended discussions). The consequence of considering such a broad contextual level of support to define aspect leads to differing interpretations about speakers’ selections of aspectual markers. Not surprisingly, thus, most definitions of aspect make reference to perspective or construal of aspectual knowledge. Klein (1994, p. 16, italics added), for one, points out that aspect “concerns the different perspectives which a speaker can take and express with regard to the temporal course of some event, action, process, etc.” Similarly, Michaelis (2004, p.5, italics added) describes “aspectual categorization as a product of the manner in which people, as producers and processors of texts, construe scenes, rather than as a reflection of the properties which situations have ‘in the world’.” As a consequence, there is a component of subjective interpretation and expression that baffles both researchers and learners looking for a delimited and precise way to determine “right and wrong” selections of aspect reflected in verbal morphology.

The variable selection of aspectual predicates is magnified by the effect of various levels of linguistic representation to define aspect. In concrete terms, restricting the construct of aspect to the level of lexical information (Aktionsart concept) has resulted in an incomplete account. Ever since the publication of the groundbreaking paper of Verkuyl (1972), there has been a general consensus on the value of the linguistic context beyond the verbal predicate including the composite effect of both internal and external arguments on the representation of aspect (e.g. Depraetere, 1995; Filip, 1999; Klein, 1994; Tenny, 1994). In some cases, definitions of
aspect include the compositional value of adjuncts (e.g. adverbial phrases). Consequently, aspectual information straddles a number of layers of linguistic interpretation, including lexical, syntactic, discursive, and pragmatic levels of analysis (see Salaberry, 2008, for an overview of relevant studies). In fact, the complexity of aspectual knowledge is understated once we consider the range of phenomena affected by aspect such as the aspectual uses of Spanish copula (e.g. Luján, 1981), the so-called Spanish impersonal se (e.g. Suñer, 1990; Zagona, 1994), the use of the perfective–imperfective contrast to segment the discourse structure of a narrative (e.g. Silva-Corvalán, 1983, 1984), and the effect of repeated instances of an event as reflected in habituality or iterativity (e.g. Salaberry, 2013a; Slabakova & Montrul, 2007), among many other topics.

A (broad) grammatical representation of aspect

The combined effect of the speaker’s construal of a situation and an ever-expanding range of contextual information (that the speaker may rely on to profile a situation) has led some researchers to compartmentalize such complexity. One immediate way to do this is to hierarchically organize the various levels of linguistic information available to determine aspectual representation. To this extent, the layers of meaning closest to the verbal predicate are the most stable, thus the most likely to limit the variability of aspectual composition. One of the most well-known multilevel models has been the one proposed by Smith (1997).

Smith proposes a basic two-level system: viewpoint or grammatical aspect (broadest level of contextual support) and situation type or inherent lexical aspect (restricted to internal and external arguments of the verbal predicate). The concept of situation aspect has a long history through the division of verbal predicates into lexical aspectual classes. Vendler’s (1967) classification of four main verb types has become the benchmark classification in studies in L2 acquisition: states, activities, accomplishments, and achievements (some researchers argue for three classes, or, in other cases, five: see Salaberry, 2008, for a review). Viewpoint aspect, on the other hand, is encoded not only on the verbal predicate, but in components beyond the head of the verb phrase as well, such as adverbials as in (1), and contextual information including world knowledge and pragmatics as in (2).

(1) Suddenly, I was asleep.
(2) Last year I fed my cat.

In (1), the adverbial prompts an inchoative interpretation of the verbal predicate focusing our attention on the inception of the state of being asleep. In (2), the adverbial provides a timeframe that forces a habitual interpretation, rather than a punctual one. As a consequence, it is not necessarily the case that we can avoid the inherent complexity of aspectual contrasts by means of restricting our analysis to the effect of lexical aspectual classes. At a minimum, we face the challenge of deciding where to draw the line between one type of aspect (i.e. situation versus viewpoint) as we attempt to partition the effects of the contextual information that aspectual contrasts always require.
In some cases, adverbial information can bring up complex interpretations that require a fine-grained analysis of the compositional value of adjuncts in combination with internal and external arguments and the predicate proper. Note, for instance, that in Spanish the adverbial phrases in the sentences below (from Güell, 1998, p. 102) set up a conflict with the prototypical aspectual marker for states.

(3) a. *Lo supo (PRET)/sabía (IMP) durante mucho tiempo.
(S/he) knew it for a long time.

b. Lo *supo (PRET)/sabía (IMP) desde hacía mucho tiempo.
(S/he) knew it from a long time ago.

In (3a) there are two pieces of information that would normally steer L2 speakers to use the imperfective form. First, states are prototypically associated with the imperfect. Second, the contextual information provided by the adjunct (the aspectual role of the adverbial phrase *durante mucho tiempo*) contradicts the meaning of inchoativity (i.e. the beginning of the state) which is the one that normally triggers the use of the perfective. In fact, despite the effect of the durational adverbial phrase, the preterit is the preferred marker in (3a) among native speakers. Along the same lines, in (3b) the presence of the adverbial *desde hacía mucho tiempo* provides an initial point of that state and would—in principle—trigger an inchoative interpretation (requiring the use of the perfective). But, this is not the case here as the use of the imperfective maintains the focus on the actual state irrespective of the explicit highlighting of the inception point.

Analyses from other Romance languages are also relevant and useful for the present discussion of Spanish aspect given the almost identical representation of aspectual construals in Romance languages. Brisard (2010, p. 489) proposed that a contextualized definition of aspect is necessary to understand the concept of aspect and what non-native speakers need to learn: “interpreting the concrete (temporal or modal) values of this tense [imperfective] depends on pragmatic inferences on the basis of contextually provided information and is, as such, not to be attributed to the semantics of the [imperfective] proper” (italics added). In this regard, one of the earliest studies to (indirectly) assess the relative effect of contextual effects on the selection and use of aspect was the analysis of Coppieters (1987). Among a number of phenomena, Coppieters analyzed the selection of past tense endings in French (i.e. *passé composé-imparfait*) and concluded that the biggest contrast in the judgments of grammaticality among native and near-native speakers was between imperfective and perfective meanings (highest deviation of 39.5% between groups). Coppieters noted that whereas native speakers had a strong sense of which choices to make, near-native speakers were more ambivalent about their selections. He argued that the locus of such discrepancy was probably due to the highly contextualized nature of aspectual markers: “it may be difficult (particularly for one whose native language does not formally mark the category or distinction in question) to separate contextual from grammatical information” (p. 567, italics added). For a review of the relevance of Coppieters’ study, see Salaberry (2016).
Imperfective as the true aspectual marker

The complexity of aspectual contrasts in Spanish appears to be closely associated with the imperfective given that the latter is the marker most relevant to carry aspectual information. On this point, Doiz (1995, 2002) argues that whereas the Spanish preterit is used to convey a perspective associated with speech time (tense), the imperfect is used to signal a view of the given situation from a past viewpoint (strictly aspectual) given an implicit reference point situated in the past.

(4) Mis hermanos y yo crecimos en una familia muy grande.
   My siblings and I grew up (PRET) as part of a large family.
(5) Mis hermanos y yo hacíamos mucho ruido.
   My siblings and I made (IMP) a lot of noise.

The sentence in (4) uses speech time as a point of reference to mark past tense. In contrast, sentence (5) relies on the use of a relative moment in the past as the reference point. Doiz distinguishes the present viewpoint (the speech moment or the here-and-now) from the past viewpoint (an alternative here-and-now). The latter is the one that serves to visualize the situation (and mark it) as imperfective. In essence, the imperfective may be regarded as the true aspectual marker, although it does achieve its function when contrasted with a true tense-marker, the perfective. The proposal made by Doiz is further justified given that previous models of aspect seem to point in a similar direction, as is the case of Klein’s (1994) proposal for the use of the concept of Topic Time as an additional point of reference to be differentiated from Time of the Situation.

Recent L2 empirical data (e.g. Labeau, 2005, and McManus, 2013, for L2 French; Salaberry, 2011, for L2 Spanish) confirm previous claims (e.g. Salaberry, 1999; Wiberg, 1996) that there is an increase in the use of prototypical pairings of lexical and grammatical aspect along with increased L2 proficiency. Furthermore, the same publications show that the perfective form acts as a default marker of past tense, whereas the imperfective form takes longer to acquire given its more complex aspectual connotations. In the case of L1 English speakers learning L2 Spanish, the perfective is a productive marker to the extent that it functions as a grammatical category applicable to old and new lexical items. In contrast, the imperfective is added to the system in an associative fashion.

Iterativity and habituality: fine-grained distinction

As stated above, the inherent complexity of the imperfective form as the basic carrier of aspectual information obtains by reference to the perfective (even if the latter acts as a carrier of tense). By the same token, the aspectual meaning of the perfective obtains in contrastive use with the imperfective. One particular case in which such complexity can become a daunting learning target for L2 learners is the case of iterated eventualities that can be marked with either one of the aspectual markers of verbal morphology in Spanish, as shown in sentences (6a) and (6b).³
In principle, the situation that is described in (6a) and (6b) can be exactly the same one (let us assume, for instance, that Lucas played soccer for exactly 10 years in both cases). The linguistic representation is, however, different. The representation of the eventuality has been aspectually qualified. We conclude, thus, that the effect of the adverbial phrase is associated with the type of aspectual representation. The identification of the specific aspectual effect of various adverbial phrases remains, however, a challenge (see Menéndez-Benito, 2002).4

One theoretical proposal has been more promising with regards to the identification of the specific conceptualization of iterated eventualities differentially classified as instances of habitual versus iterativity. Langacker (1999, p. 251–253) makes the case for the existence of two distinct types of aspectual concepts associated to the repetition of events in the past: iterative sentences and habitual sentences.5 The conceptual distinction, Langacker claims, hinges on the existence of two types of knowledge that he labels the actual plane and the structural plane. Langacker proposes that iterativity (‘repetitive’ is his choice of words) shows the component events of individual instances of the entire eventuality anchored to particular points in time (“conceived as actually occurring” on the actual plane), whereas in a habitual sentence, the component events are not anchored to any particular points in time (“with no status in actuality”) (p. 251).

Doiz (1995, p. 107) contrasts the distinctive meanings of habituality and iterativity as they are represented through imperfective and perfective morphology, respectively, in Spanish:

(7) a. El año pasado iba a nadar todos los días.
   Last year I went (IMP) swimming every day.
   b. El año pasado fui a nadar todos los días.
   Last year I went (PRET) swimming every day.

In line with Langacker’s explanation, Doiz states that the implicature of sentence (7a) is that the repetition of the event of swimming does not continue into the present, thus signaling the concept of habituality. In contrast, there is no such implicature for the interpretation of sentence (7b), in which case the repetition conveys the aspectual concept of iterativity. Doiz notes further that the concept of habituality allows for the failure of the event to take place at one particular time during last year (i.e. events not anchored on the actual plane allow for this interpretation). In sentence (7b), on the other hand, Doiz points out that the use of the perfective form (associated with iterativity) conveys the notion that the speaker went to swim every day last year (i.e. constrained by events anchored to specific points in time on the actual plane).
As we can see, in both habituality and iterativity, a series of events are iterated or repeated. How should learners apprehend the distinction? Salaberry (2013a) and Salaberry and Martins (2013) analyzed the effect of iterativity in contrast with the concept of habituality among monolingual Spanish speakers and L1 English speakers with near-native competence in the L2. The study was carried out with the use of grammaticality judgments based on contextualized sentences (short passages). The findings revealed that Spanish near-native speakers, despite their extensive experience with the L2, did not distinguish fine-grained representations of the aspectual concepts of iterativity versus habituality. The same speakers demonstrated, nevertheless, native-like judgments of more prototypical uses of aspect. Both findings combined confirm the positive effect of language experience in general, but the failure to acquire conceptual distinctions that are rarely (if ever) taught explicitly. The role of explicit descriptions of nuanced semantic descriptions may lead to novel conceptualizations of aspect, thus pedagogical intervention may have an important role to play (e.g. Lantolf, 2011). I return to the role of pedagogical intervention in subsequent sections.

Regardless of the theoretical perspective used to assess the acquisition of the phenomena reviewed above, it is clear that the acquisition of some components of the definition of aspect continue to be a challenge for advanced L2 learners. Studies within the perspective of Minimalism (e.g. Montrul & Slabakova, 2002; Rothman & Iverson, 2008; Slabakova & Montrul, 2007) have shown similar empirical results even if the theoretical description differs from the one provided above within the framework of cognitive linguistics. However, some recent analyses hint at the possibility that advanced L2 learners are not able to acquire the concept of aspect in the L2. For instance, Diaubalick and Guijarro-Fuentes (2016) collected data to assess the relevance of two hypotheses (i.e. the Interpretability Hypothesis and the Feature Reassembly Hypothesis) on the interpretation of “coerced” meanings (de Swart, 1998). They conclude that neither hypothesis can be confirmed with the data from their study, noting, furthermore, that “even advanced speakers do not reach native level” with respect to the “coercion condition in the grammatical judgment task” (p. 192).

Temporal vs. non-temporal meanings of aspect

The representation of aspect is even more complex when we consider the fact that the traditionally labeled past tense morphology is actually used to convey non-temporal information as well (e.g. Fleischman, 1989, 1990; Waugh, 1990). Fleischman (1990, p. 5–6), for instance, pointed out that the grammatical category past might convey multiple oppositional properties at different levels of analysis: referential (based on truth conditionals especially related to temporality), textual (organization of discourse: foreground versus background), expressive (conative, affective, and social functions), and metalinguistic (styles, registers, or types of language). In line with the previous argument about the inherent aspeсtual value of the imperfective with regard to the perfective, Brisard (2010) contends that the imperfective “presents a situation as part of a mentally construed reality which
does not coincide with the speaker’s and which is not to be considered as actual for that reason” (p 488). Brisard thus argues that strictly temporal accounts of the imperfective as past tense cannot account for all of its possible uses, including the non-temporal (modal) ones. For reasons of space, I will not discuss the uses of aspectual contrasts to convey non-temporal meanings.

The acquisition and development of aspect across the bilingual continuum

The task of providing a comprehensive definition of aspect is further complicated once we take into account two important theoretical dimensions that must be factored into any description of aspectual knowledge: not only do we need to describe the representation of aspect among monolingual native speakers, but we also need to consider (i) the description of aspect among bilinguals as well as (ii) the description of the development of aspectual meanings among L2 learners of varying levels of proficiency as they approach a bilingual stasis. Given the wide range of contextualization of aspectual knowledge (subjective viewpoint), it is relevant to consider a continuum of bilingual speakers with monolingual native speakers on one end of the spectrum and non-native speakers on the opposite end. This continuum of various levels of representation of aspect is most useful to account for the wide range of outcomes across many different empirical studies. Most famously, Silva-Corvalán (1991, 1994) demonstrated that aspect is among one of the last concepts to be acquired by English-Spanish bilingual speakers, and it is also one of the first grammatical concepts to be negatively affected by language attrition (once bilinguals start to shift away from the use of Spanish). Given that the bilingual speakers studied by Silva-Corvalán spanned many different levels of Spanish competence (i.e. she collected data from up to three generations of bilingual speakers), it is necessary to consider a wide spectrum of aspectual knowledge, including L2 learners.

The use of a continuum of language proficiency is also useful to match discrepancies in outcomes in empirical studies with specific theoretical claims. For instance, Salaberry (2011) empirically demonstrated that the use of the imperfective continues to increase proportionally relative to the increase of the perfective form across a wide range of proficiencies in L2 Spanish among L1 English speakers. Using these findings, Salaberry argued that as non-native speakers gain more experience with the target language, they may develop an increasingly accurate system of proceduralized knowledge (e.g. DeKeyser, 2009; Paradis, 2009) that allows them to track target-like selections of aspectual markings based on probabilistic frequencies associated mostly with lexical aspectual values (cf. Andersen, 1991, 1994), and to some extent, with discourse grounding (cf. Bardovi-Harlig, 1995; Salaberry, 2011). In general, theoretical claims about the nature of the acquisition/learning system will be most informative in the context of a wide range of proficiencies across the spectrum of language competence, from monolingual to bilingual. In this respect, among Spanish L2 learners, aspectual contrasts are among the most difficult topics to be mastered in the traditional language
classroom (along with other semantic-based phenomena such as ser-estar, the subjunctive, various functional uses of se, etc.).

**Instructional settings and pedagogical rules**

*Imprecise and incomplete descriptions of aspect*

There is an abundant literature on the topic of pedagogical rules to explain the aspectual contrast in Spanish (e.g. Frantzen, 1995; Lunn, 1985; Westfall & Foerster, 1996). In most cases, the effort to provide learners with descriptions that can capture the entire spectrum of meanings of aspect has resulted in simplistic descriptions. These incomplete (and thus misleading explanations) have been referred to as rules of thumb (e.g. Whitley, 1986). In general, most rules of thumb focus on examples that carry only prototypical marking of verbal morphology according to lexical aspectual classes (i.e. imperfect is associated with stative and preterit with telic events), thus limiting the applicability of the concept that learners need to develop. Despite numerous expansions, modifications, and new classifications, alas, rules of thumb do not appear to be more than mere ‘crutches’ to help learners move along in the process as they gather more experience in the L2 (and can start to make sense of aspectual contrasts on their own).

*Advanced knowledge of aspect: Beyond rules of thumb*

To address the challenge of using inaccurate rules of thumb, two theoretical proposals have argued for a reconceptualization of the way traditional deductive instructional procedures guide learners to develop a representation of aspectual contrasts in L2 Spanish: Concept-Based Instruction (CBI) (e.g. Lantolf, 2011; Negueruela, 2003; Negueruela & Lantolf, 2006; Yáñez-Prieto, 2008) and the Competing Systems Hypothesis (CSH) (e.g. Long & Rothman, 2012; Rothman, 2008). Rothman, for instance, proposes the pedagogical value of (good) instruction to develop important complex concepts (or, at a minimum, not to “compete” with the positive effects of access to language data in context). Even more strongly, the proposal from the CBI hypothesis relies on the development of explicit linguistic knowledge (organized in a way to promote understanding, control, and organization) as the solution: “models must raise learners’ awareness of what linguistic resources are available to them to carry out concrete linguistic actions with specific purposes across all contexts” (Negueruela & Lantolf, 2006, p. 84–85). Both hypotheses (CBI and CSH) identify the same problem (i.e. oversimplification of grammatical concepts through the use of rules of thumb), and they favor, in general, the use of more developed conceptual representations of aspect. On the other hand, each hypothesis defines the role played by explicit instruction within different theoretical frameworks. Lantolf (2011), for instance, privileges the perspective on grammar derived from cognitive linguistics, whereas Rothman (2008) aligns himself with the Universal Grammar (UG) model associated with the theory of Minimalism.
Developing an advanced conceptualization of aspectual knowledge: CBI and CSH

Negueruela and Lantolf’s (2006) position against the beneficial effect of traditional pedagogical rules is predicated on the deficiency of an instructional program narrowly focused on a mechanistic approach to learning: “Simplified and reductive rules of thumb have the potential to do more harm than good because, for one thing, they depict language as a sedimented entity that appears to have a life of its own independent of people” (p. 83). The conceptualization of language of CBI is inherently defined by social phenomena and is demarcated by the co-construction of linguistic meanings among interlocutors (more precisely, developed through intermental mediation). Interestingly, Long and Rothman (2012, p. 67) make a similar case for the futility of most instructional intervention through the teaching of oversimplified (and misleading) rules of grammar: “Oversimplification in classroom instruction can lead to the formation of a static system of learned rules.” The latter suggestion is framed within the perspective of a definition of rules less affected by social interaction and more so by underlying language abilities residing within each individual speaker as the outcome of a genetically guided developmental process. Furthermore, both CBI and CSH argue for the pedagogical relevance of explicit and rigorous linguistic explanation. Rothman, for instance, claims that “Most language teachers are not trained in formal linguistics despite compelling reasons to expect they could benefit from such training” (Rothman, 2008, p.77).

In sum, both CBI and CSH have important points in common as they both reject the adduced benefits of traditional pedagogical rules. Both hypotheses also acknowledge the development of knowledge (probably akin to a transition from declarative to procedural knowledge) as learners gain more experience with the language (through classroom instruction). The hypotheses do differ on the causal effect of the instructional setting (classroom versus natural, non-instructional setting of communication) on the accurate and complete conceptualization of aspectual knowledge.

The problems with the advanced conceptualizations of aspect of CBI and CSH

Despite their bold theoretical claims about learning, the definition of aspect used by both hypotheses (CBI and CSH) is in conflict with the broad definition of aspect outlined in previous sections. For instance, previous classroom-based studies that have been advanced as evidence for CBI have used shallow descriptions/definitions of aspect that do not do justice to the full range of meaning potentials of aspect. Whereas studies such as Negueruela (2003) and Yáñez-Prieto (2008) have addressed important components of the CBI hypothesis (i.e. visualization and verbalization of conceptual knowledge), they have not incorporated a comprehensive definition of the concept of aspect. Negueruela,
for instance, specifically restricts his analysis to the definition of aspect from Bull (1960, 1965), and in particular, Bull’s contrast between cyclic versus non-cyclic verbs. As a consequence, the definition of aspect chosen by Negueruela does not match the level of complexity of aspect as a matter of perspective/construal and/or the existence of various layers of representation of aspectual meanings. For her part, Yáñez-Prieto uses the same operational definition of aspect from Bull (1960, 1965), while making some passing reference to the concept of non-prototypical pairings of lexical and grammatical aspect described in Salaberry (2008). While the notion of non-prototypical meanings helps to focus our attention on a slightly more complex definition than the one provided by Salaberry, it is still not as comprehensive as the definition described in the first sections of this chapter.

The bulk of the analysis of both these studies (i.e., Negueruela and Yáñez-Prieto) is focused squarely on the ways students verbalize their understanding of aspect as they react to various teaching procedures that communicate that “the selection of aspect depends on the perspective and focus that the speaker or writer wants to adopt” (Yáñez-Prieto, 2008, p. 426). In fact, while Lantolf and Poehner (2014, p. 121) note that the pedagogical scaffolding based on flowcharts used by Negueruela (2003) was significantly improved with the use of images in Yáñez-Prieto (2008), it is noticeable that the definition of aspect used by both authors remained essentially the same. As much as there is an improvement in the process of visualization of the concept, the latter was not matched by an expanded definition of aspect. In essence, if the real target of L2 instruction is to develop learners’ knowledge of the linguistic resources available in the L2 to be able to convey the intended information, we have a problem insofar as the descriptions of the target grammatical item (in this case aspect) have not been properly configured for explicit presentation, analysis, and implementation.

For its part, the CSH adopts the definition of aspect used in all studies carried out within the Minimalist framework (e.g., de Swart, 1998; Giorgi & Pianesi, 1997; Slabakova & Montrul, 2007). The theoretical compromise of such a theoretical approach is that many contextual effects that are part of viewpoint aspect (the higher aspectual phrase or AspP) are regarded to be outside of the realm of grammar proper (i.e., they are part of pragmatics or world knowledge). That is, the most variable interpretations of aspect are left outside of the scope of the definition of aspect (see Salaberry, 2013b, for an extended discussion of related problems). At the same time, the least variable component of lexical aspect (mostly associated with the lower aspectual phrase) becomes the main target of analysis for most UG-based studies. The methodological decision (i.e., related to research design) to leave some components of the concept of aspect outside of the scope of analysis is not under scrutiny. As a matter of research design, by definition, researchers can delimit their research space. On the other hand, we cannot blur the lines between methodologically efficient decisions and theoretical representations. For instance, Slabakova and Montrul (2007) methodologically “constrain” the theoretical representation of aspect by considering the context above the lower verb phrase as
information representing pragmatic information, thus outside of the realm of aspectual representation. The concern with regards to the delimitation of the concept of aspect is relevant because, as we have seen in previous sections, it is not necessarily the case that information provided by adjuncts should be relegated to pragmatics knowledge (or, at least, they should not necessarily be completely isolated from grammatical knowledge).

Furthermore, unlike the claim advanced by CBI, the proposal of CSH does not focus on specific pedagogical procedures to acquire aspectual knowledge in the L2. Instead, the CSH shifts the focus of attention to the potential effect of the natural setting of communication, under the assumption that direct access to language data will be enough for learners to activate their knowledge and modify the representation of aspect. While access to rich and varied input is a necessary component of understanding and learning aspectual contrasts, it is not entirely clear that access to the non-instructional setting is a solution to the negative effect of pedagogical rules of thumb. For one thing, loosely defined conditions of natural settings may include access to some type of pedagogical scaffolding in the form of some explicitly stated generalizations about language structure. Long and Rothman concede as much: “We do not intend to suggest that naturalistic learners are never offered ‘rules’ by native speakers with whom they interact or do not attempt to form their own version of descriptive rules” (2012, p. 71). Furthermore, previous research findings on natural learners with no access to any type of pedagogical intervention are not very promising. Many longitudinal studies of naturalistic learners of various ages and with varying levels of exposure to non-instructional uses of the L2 reveal little marking of past tense morphology (e.g. Sato, 1990; Schumann, 1987; Trévise, 1987). In general, most studies of natural learners show that the development of verbal endings is a slow and gradual process which in some cases takes years, and in others merely leads to fossilization (e.g. Klein & Perdue, 1992). Dietrich, Klein, and Noyau (1995) concluded that natural language learners seem to be especially affected by the particular contextual features of natural discourse: the use of verbal morphology is not necessary to establish communication in the L2 among natural learners.

The proposals outlined by the CBI hypothesis and the CSH provide an auspicious opportunity to focus our attention on the conceptualization of aspect among advanced learners of the L2. Accordingly, both hypotheses rightfully point out the inadequacies of providing learners with incomplete and misleading generalizations (rules of thumb) to guide the development of knowledge of aspect among adult L2 learners. Even if not completely on target, both hypotheses represent the first step toward developing an agenda that can help L2 learners become better users of language resources. In the final analysis, however, the CBI hypothesis underestimates the potential of the inductive processes managed by learners as the agents of change, whereas the CSH makes the opposite mistake by ignoring the fact that adult (especially literate) learners normally make a connection between form and function even when not placed in a (formal) instructional setting.
Guided (enhanced) induction

*Guided induction: language data and metalinguistic awareness*

There are two basic principles that present a straightforward solution to the previous deficiencies of CBI and CSH, leading to an alternative proposal: (i) the conceptualization of grammatical concepts such as aspect, being as complex as it is, can only be achieved (developed) through numerous encounters with samples of (mostly natural) language data, and (ii) (deductive) pedagogical interventions (mostly represented as metalinguistic awareness events) can be useful to guide the development of conceptualizations of the L2. The first principle follows as the natural consequence of considering a broadly contextualized definition of tense-aspect meanings, whereas the second one focuses on the type of processing needed to modify the form-meaning mapping of aspectual knowledge in the L2. The integration of both of these principles into a single theoretical proposal creates an inductive-deductive continuum that is critical to address the multilayered representation of tense-aspect meanings.

First, as described above, the deficiency of the CBI proposal is that the development of metalinguistic descriptions of highly complex grammatical concepts (such as aspect) was predicated on an incomplete and inaccurate definition of aspect. To use the more comprehensive definitions of aspect described in previous sections, learners need access to an extensive language database. That is, to properly conceptualize the complex notion of aspect requires the use of a comprehensive, data-based approach to represent the concept of aspect along with a reduced focus on the attempt to develop complete models of grammatical concepts (i.e. more inductive than deductive). Contrary to the main claim inherent in the CBI proposal, the actual force of this process resides primarily in the inductive process managed by the learner with the guidance of the instructed process. Interestingly, the proposed (qualified) improvements in the conceptualizations of aspect among learners in the studies conducted by Negueruela, Yáñez-Prieto, and other proponents of CBI happened despite the fact that the proposed explanations from instructors are incomplete or misleading (see critique of Bull’s proposal above). In sum, pedagogical interventions are most useful when used in synchrony (over time) with inductive processes of analysis of large samples of language data.

Second, the development of (explicit) metalinguistic awareness is necessary to develop advanced knowledge about aspect. The proposal of CSH eschews instructional procedures, arguing for a reliance on extended access to sociolinguistically contextualized language use (orthogonal to classroom-based language use, and thus primarily inductive). Contrary to the claim of CSH, however, the effect of metalinguistic awareness, especially in natural language use contexts, cannot be discounted. Among some early studies in SLA, Klein (1986, p. 16) posited a communication-learning paradox: for a learner to be able to communicate and interact with other speakers, “he must learn the language,
and in order to learn it, he must communicate.” This paradox notwithstanding, Klein points out that learning occurs when a mismatch is perceived between the learner’s own output and that of others (p. 141). For that mismatch to be perceived, learners must engage in some type of metalinguistic awareness. In line with Klein’s proposal, Jessner (2008) defined metalinguistic awareness as “the ability to focus on linguistic form and to switch focus between form and meaning.”

Along those lines, Swan (2005) argues that it is precisely in areas above the sentence level (for which we may erroneously adduce the positive effects of unguided induction) that we need a more explicit approach to guide the learner’s metalinguistic focus. Swan stated that, for instance, “… if students do not already know the linguistic conventions for opening and closing conversations, interrupting and challenging, etc, how are they supposed to learn them without input from the ‘dominating’ teacher? One cannot teach by eliciting what is not there” (p. 350). The focus of analysis chosen by Swan is relevant for our discussion given the prevalent belief that study-abroad settings are ideal for the development of sociolinguistic-appropriate language use. In fact, the basic concept of metalinguistic awareness (spanning the entire range from traditional grammar points to sociocultural norms) reminds us that we cannot—idealisti­cally—define the context of non-classroom language use and interaction as devoid of any opportunity for the metalinguistic analysis of language. In fact, the available empirical evidence of study-abroad settings casts doubts on some of the proposed benefits of unguided natural settings of communication (Collentine, 2004; Freed, Segalowitz, & Dewey, 2004) while highlighting the positive effects of intensive at-home programs that are based on extensive access to input data along with pedagogical scaffolding of various types (the latter more akin to enhanced classroom-based environments).8

Conclusion

For L2 learners to demonstrate advanced knowledge of aspect, they must know and be able to manage not just the shallow meanings represented in prototypical pairings of aspectual markings (i.e. rules of thumb), but also the integrated representations of aspect that underlie the non-prototypical choices that, despite not being common in the data, are managed and used precisely by competent speakers of the language. Such advanced knowledge of aspect cannot be achieved by making reference to verbal forms in isolation which can only convey general referential meaning. As Silva-Corvalán (1986, p. 244) claimed early on, “… the general referential meaning of a verbal form may in part overlap with the meaning of another and … form-specific meanings must be identified in contexts of use.” L2 learners with access to advanced knowledge of aspect are able to integrate the superficial layer of meaning conveyed by lexical aspect, and internal and external arguments, with additional layers of information that contribute to the complete value of the aspectual contour of eventualities. Even though proficient learners of a
language with complex aspectual representations may function reasonably well in
the use of the L2 while making reference to the components that are most salient
and most used to compute the aspectual value of verbal predicates, only advanced
learners are able to integrate the deeper layers of aspectual meanings when they
are required to produce native-like interpretation and production of aspectual
morphology.

Future studies of tense-aspect among advanced L2 learners should probe
further into the highly contextualized layers of meaning that are part of the con­
ceptualization of temporality that have been described in this chapter. As noted
above, one of the first studies to assess the advanced knowledge of temporality
in the L2 (among other grammatical categories), Coppieters (1987) analyzed
native and near-native speakers’ judgments of uses of French passé composé
imparfait. Even though Coppieters’s study was focused on decontextualized
sentences, the judgments of native speakers reflected access to nuanced contex­
tual meanings apparently recovered from such decontextualized utterances. For
such a recovery to happen, new empirical studies need to expand the context of
use of aspectual meanings while precisely identifying the sources of information
used by participants in future studies. A few initial studies have already
eschewed the sentence-level approach to studying aspectual interpretations to
incorporate instead a comprehensive view of temporality that provides access
to the analysis of the complete construct of aspectual representation (e.g.
Salaberry, 2013a). Future studies should continue to circumscribe the dependent
variable they will use to ascertain the level of knowledge about aspect to be
investigated. From a methodological perspective, new studies of aspectual
knowledge should consider the effect of such a broad level of contextualization
of aspect for the description of independent variables such as lexical aspect,
discourse grounding, and other lexical and discursive factors. In a recent review
of the methodological options available to future researchers of advanced levels
of knowledge of tense-aspect, Salaberry and Comajoan (2013) provide a sum­
mary of multiple perspectives on how to approach this task. Finally, some ped­
agogical proposals, such as the CBI, have also moved our field toward the
analysis of the acquisition of aspect beyond the narrow perspective of the sen­
tence level (typically associated with rules of thumb) to a broader realm of contex­
tualization, and consequently a more accurate conceptualization of the
grammatical representation of aspect.

The first 30 years of research on the acquisition of tense and aspect in the L2
have laid the foundation for the next stage of collection of empirical data to develop
a comprehensive picture of the acquisition of the complex, multilayered
grammatical concept of tense-aspect marking in the L2. Future studies should
expand on the knowledge acquired so far and incorporate new methodological
approaches that are in line with the theoretical knowledge developed to date. The
advantages of a broad view on the description of aspectual knowledge stand to
benefit the field of second language acquisition given the multifaceted description
of aspect and the need to include such a concept as part of the definition of the
advanced L2 learner.
NOTES

1 The examples are taken from Smith (1997).
2 In some cases, even in the absence of adverbial information, the same effect obtains when relevant information is implicit: see Bybee (1995).
3 The sentences are taken from Salabery and Martins (2013).
4 Menéndez-Benito proposed that Spanish preterit and imperfect cannot combine with generic adverbs (e.g. normalmente ‘normally’) and durational phrases (e.g. durante dos años ‘for two years’), respectively. Nevertheless, there are many counter-examples that challenge that assumption, as acknowledged by Menéndez-Benito.
5 Langacker considers generic sentences together with habitual sentences.
6 An exception to this developmental/acquisition trend represented by studies of natural learners happens when learners focus more explicitly and consistently on grammatical form, in which case they show some clear signs of development of aspectual morphology (e.g. Giacalone-Ramat, 2002).
7 Jessner (2008) added that “[i]ndividuals who are metalinguistically aware are able to categorize words into parts of speech; switch focus between form, function, and meaning; and explain why a word has a particular function.”
8 For the context of classroom instruction, Toth, Wagner, and Moranski (2012, p. 19) demonstrated the notion of “richness of engagement” whenever learners can “formulate L2 rules and use contributions from other learners to further their thinking.” In fact, even when students seem not to be actively engaged in the analysis of language, as measured by behavioral indices (actual talk), they may still be focused on the active analysis of language as shown in the extended study of Amy Ohta (2001).

REFERENCES

Bull, W. (1960). Time, tense, and the verb: A study in theoretical and applied linguistics, with particular attention to Spanish (First
Advanced Conceptualizations of Tense and Aspect in L2 Acquisition


20 Inflectional Morphology

ROUMYANA SLABAKOVA
University of Southampton, UK
UiT The Arctic University of Norway

Introduction

Producing and understanding messages in a second or additional language is a complex process with many facets, which have all been the focus of theoretical and experimental investigations. For example, learning motivation, learning strategies, and vocabulary acquisition, among many other topics, are important elements of the acquisition process. There is one part, however, that is at the core of the linguistic experience and without which none of the other facets of language acquisition can be brought together. This indispensable part is acquisition of the second language grammar. Contemporary linguistic theory makes a useful distinction between lexical and grammatical linguistic information. In the context of advanced learners, this distinction takes on added significance because new lexical items can be acquired through the lifetime, while the acquisition of the L2 grammar is expected to be largely in place for advanced L2 learners. In this chapter, I shall explore what it means for a target language grammar to be in place. I shall use inflectional morphology as a yardstick for this knowledge. In particular, I will distinguish between underlying knowledge of linguistic features, production of the inflectional morphology, and sensitivity to this morphology as reflected in language processing studies. I will argue that, even at advanced proficiency levels, underlying linguistic knowledge may be under-represented by production and processing. Thus, I will argue that only by looking at all these different types of linguistic performance can advanced linguistic competence be adequately described.
The function of inflectional morphology and its place in the language faculty

**Inflectional morphology expresses grammatical and semantic features**

Generative linguistics, and especially its most recent account of language, the Minimalist Program (Chomsky, 1995), attributes a special place to functional morphology in grammar and acquisition. Functional morphemes, both bound (inflectional) and free morphemes (e.g. articles, auxiliaries), carry grammatical meanings that radically change the overall interpretation of a sentence and discourse. No sentence meaning is complete without taking the functional and especially the inflectional morphology contributions into account. The following example illustrates the division of labor between lexical and functional categories. Suppose we want to combine the lexical items *works, Dexter, hospital, in, a* into a message with this meaning: “Dexter is gainfully employed by some hospital or other.” No language would just string the words together without taking grammatical and without taking word order into account. In English, for example, the meanings of count, singular, non-familiar noun are imparted by the indefinite determiner *a*; the aspectual meaning habitual action is reflected in the present simple tense; the temporal meaning of this action including the moment of speech is captured by the present tense form; the verb also agrees with a third-person-singular subject.¹

(1) Dexter works in a hospital.

The final makeup of the sentence in (1) carries a complete meaning (its truth-value) composed of lexical and grammatical meanings that are considerably removed from the raw assembly of notional verb and two nouns. These additional meanings that anchor the event in time and space are contributed by the functional morphemes, including the inflection marker –s.

In addition to being indispensable in carrying grammatical meanings, the functional morphemes are the main locus of linguistic variation among languages of the world. As per the Minimalist Program assumptions (Chomsky, 1995), the core syntactic operations (e.g. Merge, Agree, etc.) are the same in all natural languages, hence they are universal. Parametric differences are no longer explained by postulating that the syntax of two languages differs. Of course, syntactic structure needs to be correlated with semantic structure and with some expression in sound or sign or writing for the form–meaning mapping to be effected. If syntax is universal, then variation among languages can only be located in the functional lexicon and at the two interfaces with meaning and form (or expression). Apart from learning the equivalents of the native open-class lexical items (verbs, nouns, adjectives, etc.), acquiring functional morphology with all its features secures the skeleton of the second language grammar (see Slabakova, 2016, chapters 1 and 2, for more discussion of language architecture).
Furthermore, in languages such as English and many others, functional morphemes can have multiple grammatical meanings. Linguists express this state of affairs by postulating that a number of grammatical features reside (are checked) in the same functional category. As we saw in the example sentence above, the agreement morpheme –s signals present tense, habitual aspect, and a third-person subject. These are the grammatical meanings reflected by the morpheme in the sense that the morpheme is their expression. However, two other grammatical meanings are also captured by features in the same functional category, Tense: the subject is obligatory, as illustrated in (2), except in certain diary registers, and there is a lack of verb movement over functional morphemes such as negation, as shown in (3). The empty set symbol stands for a null subject.

(2) *Ø works in a hospital.
(3) *Dexter works not in a hospital.
(4) Dexter doesn’t work in a hospital.
(5) *Works Dexter in a hospital?
(6) Does Dexter work in a hospital?

In addition, knowing how Tense works in English includes knowledge that auxiliary verbs such as do are needed to carry the agreement in negative and interrogative sentences—see examples in (3)–(6). Thus, multifunctional Tense in English carries at least five morphosyntactic features (or grammatical meanings): present tense, habitual aspect, subject agreement, obligatory subject expression, and lack of main verb movement. We can certainly imagine these meanings being separated in acquisition, as the features can be learned one by one, or in various bundles. It seems inconceivable, though, that advanced learners of English can ignore these grammatical facts, even if they don’t have explicit and conscious knowledge of them.

The example above, based on basic linguistic facts about English inflectional morphology, serves to illustrate that knowledge of grammar is not, and cannot be, monolithic. A learner doesn’t acquire the agreement morpheme –s and its various meanings at one go. The dissociation between expression and grammatical meanings is worth emphasizing. Considering all this linguistic information as constituting functional category knowledge is not only linguistically correct but also profitable in explaining acquisition patterns. In sum, we can view the acquisition of a functional category as comprising at least three different types of knowledge:

1. morphological reflexes: target-like usage of inflectional morphology (if any);
2. syntactic reflexes: knowledge of features and feature strength encoded in this functional category, which would be reflected in movement, case-marking, etc.; and
3. semantic reflexes: knowledge of the semantic properties of the functional category: what meanings are computed when the particular functional category is engaged.
The Bottleneck Hypothesis (Slabakova, 2008, 2016) assumes precisely this understanding of what it is to know a specific piece of inflectional morphology. The hypothesis identifies the inflectional morphology as the hardest part of the grammar to acquire. The reason for this heightened difficulty is that the inflectional morphology encodes all the grammatical and semantic features distinguishing between languages. On the other hand, core syntax and semantics are universal computations once the features, or feature bundles, of the inflectional morphology are acquired. It is important to keep in mind this language architecture of the grammar: inflectional morphology is the surface realization of at least three types of features (morphophonetic form, formal syntactic features, and semantic reflexes). In the next two sections, we explore the consequences of this complex relationship among these three types of information.

The syntax-before-morphology view

White (2003, pp. 182–184), following the idea of morphological separation (Beard, 1987; Halle & Marantz, 1993), as well as Lardiere (1998a, 1998b), provides convincing evidence for the so-called syntax-before-morphology view. She argues that L2 acquisition data from children and adult learners point to the following dissociation. Morphological expressions of functional categories (e.g. the –s in our example above) are being produced with accuracy between around 5 and 45% in obligatory contexts, while the syntactic reflexes of the same functional category, Tense, are observed with close to a 100% accuracy. These syntactic reflexes are lack of verb movement, nominative case marking on the subject, and knowledge that the subject is obligatory. It seems that this syntactic knowledge is in place much earlier than the obligatory suppliance of the morphological expressions is completely achieved.

Lardiere (2006) builds on this idea to present a comprehensive idea of dissociation of knowledge at advanced proficiency levels. Lardiere provides the consummate case study of Patty, a native speaker of Mandarin and Hokkien Chinese, whose acquisition of English as a fourth or fifth foreign language began when she was an adult. While Patty’s performance has most often been given as an example of fossilization (Montrul, 2014), there are facets of her performance that firmly establish her as an advanced learner. In particular, at the two times of testing she had already spent about 10 and 18 years, respectively, in the United States, she was married to an American native speaker, she had supplemented her education with a master’s degree, and she had been working as a professional for a US company. While Lardiere discusses Patty as a case study of a learner at end-state, it can also be argued that her massive exposure to American English input and seamless functioning in society make her an advanced learner.

Let us now look more closely at her famously uneven linguistic production. Patty only supplied 4.41% of third-person agreement –s morphology and 34.66% of past tense morphology in obligatory contexts. At the same time, her production of overt subjects in the nominative case was above 98% accurate. This dissociation cannot be attributed to transfer from Chinese: both her native languages do not mark tense and agreement on the verb, but also do not case-mark subjects overtly.
and allow subjects to be dropped freely when they are Topics. While fossilization in Patty was evident in her morphophonology (especially for affixal inflection), aspects of her syntax related to the Tense functional category were completely native-like. In addition, she had native-like knowledge of pronominal case marking, case marking on subjects as a function of finiteness, placement of verbs and adverbs, relative clause formation, wh-movement, preposition stranding, subject–auxiliary inversion, and do-support in questions. It is worth reiterating that with respect to this long list of properties, Patty’s native languages work differently and straightforward L1 transfer is not possible.

Why do we see this dissociation between morphological and syntactic competence? Lardiere (2006) and White (2003) attribute it to morphological production lagging after target-like comprehension, both of the syntax and of the inflectional morphology. What are some theoretical explanations of this discrepancy? One theoretical proposal is the Missing Surface Inflection Hypothesis (Haznedar & Schwartz, 1997; Prévost & White, 2000), which argues that the correct functional morphology may not be produced due to a breakdown in the mapping between the morphosyntactic features in the syntax, on the one hand, and the morphophonology of the correct expression, on the other. Such a breakdown is attributed to imperfect lexical access, possibly under production pressure. Another proposal is the Prosodic Transfer Hypothesis (Goad & White, 2006), which suggests that there may be phonological reasons for morphophonology omission. More specifically, learners need to organize the new morphemes prosodically so that they can pronounce them, and that development may be delayed as compared to morphosyntactic development. Whatever the reasons for the morphology–syntax dissociation in performance, Patty’s case study clearly demonstrates that syntactic knowledge may be advanced even if the functional morphology lags and is highly variable in production. Finally, it is worth remembering what Hawkins (2001, p. 46) points out: even if an L2 speaker provides a piece of functional morphology with 100% accuracy, this does not mean that they attribute to it the same interpretation as native speakers do. On this issue, see the next section.

The semantics reflexes of functional morphology knowledge

Another interesting dissociation to consider is the one between semantics and morphology. The logic goes like this: if learners are aware of semantic reflexes (meanings) that are related to the functional morphology and its features, does it really matter that they are not supplying it in production? If knowledge of a functional category is thus dissociated, is one type of knowledge more important than another, just because it is visible? In this section, we will review findings from three representative studies of semantics, which point to the same conclusion: semantic knowledge of functional categories is indeed attested to be native-like in advanced L2 learners.

The first study is Dekydtspotter and Sprouse (2001), an investigation of tense interpretations in L2 French. Consider the example below from French wh-movement.
(7) a. Qui de célèbre fumait au bistro dans les années 60?
who of famous smoked in the bar in the 60s?
b. Qui fumait de célèbre au bistro dans les années 60?
who smoked of famous in the bar in the 60s?
‘Which famous person smoked in bars in the 60s?’

When the *wh*-word *qui* ‘who’ moves to the top of the clause to check an interrogative feature, its so-called restriction *de célèbre* ‘famous’ can travel with it, or be left behind. These are known as discontinuous interrogatives in French grammatical description. It is interesting that this optional movement has interpretive consequences. The question in (7a) may be answered by naming a present or a past celebrity. On the other hand, it is impossible to answer the discontinuous interrogative constituent as in (7b) by naming someone who is famous now but was not famous in the past. This complex and subtle semantic knowledge is captured in grammatical features regulating the *wh*-movement and the Tense functional category. Advanced learners and even intermediate learners were reliably sensitive to the interpretive distinction. The researchers argue that this type of knowledge exemplifies a Poverty of the Stimulus (PoS), in the sense that positive evidence for this subtle meaning is very unlikely to present itself to learners (see more examples and an extended discussion of PoS in Chapter 23, Advanced-Level Semantics).

In contrast, the next study we review does not present a PoS learning situation. The generic and specific meanings under investigation are semantic features pertaining to the nominal functional categories of English and Spanish and expressed by number morphemes and determiners. The question being asked here is: Even if L2ers produce nominals in the target language, do they really know the fullness of interpretations they carry? The examples are from Ionin, Montrul, and Crivos (2013, p. 485).

(8) Tigers eat meat. [✓ generic reading, *specific reading]
(9) The tigers eat meat. [*generic reading, ✓ specific reading]
(10) *Tigres comen carne. 
tigers eat meat
(11) Los tigres comen carne. [✓ generic reading, ✓ specific reading]
the-pl tigers eat meat
‘The tigers eat meat.’

A syntax-semantics mismatch is in evidence between English and Spanish generic and specific meanings, such that genericity is expressed by bare plurals in English and definite plurals in Spanish. Using a bi-directional truth-value judgment task, Ionin, Montrul, and Crivos (2013) documented initial difficulty but ultimate success for this property, in both learning directions. One conclusion that we may draw from this study, and others like it, is that in order to evaluate
complete knowledge of certain functional morphology, it is of utmost importance to investigate interpretive knowledge of semantic reflexes.

The third study that we review in this section has to do with lack of morphological marking. Slabakova (2015) asked the following research question: If there is no overt morphological expression of Tense, as in the acquisition of tense marking in Mandarin Chinese as a second language, do learners know which sentences are to be interpreted as encoding a past action? More specifically, can learners use adverbials, aspectual morphology, and discourse context to fix the temporal interpretation of a sentence? It turns out that they do, and even at intermediate levels of proficiency. Advanced learners in this study were completely native-like in their interpretive choices.

The studies reviewed in this section (among many other studies, see Slabakova, 2008, for review) supply evidence that indeed syntactic and semantic knowledge can be dissociated from their morphological reflexes. Target-like syntactic knowledge and semantic interpretation have been attested before fully accurate realization of the inflectional morphology that expresses them. To recapitulate, such findings suggest that production of functional morphology can significantly under-represent, and in some cases also over-represent, knowledge of the associated syntax and semantics. In a sense, we cannot discuss advanced learners' knowledge of inflectional morphology without adding syntax and semantics to the big picture.

**Feature interpretability**

Another theoretical distinction that we should introduce in this section is feature interpretability. It is important because it is used by theoretical accounts to predict learner behavior in L2 acquisition. In a nutshell, grammatical features can be interpretable, that is, useful for the meaning computation of the sentence, or uninterpretable, hence formal features only relevant for the core syntactic operations (Chomsky, 1995). Feature interpretability becomes another factor that can engender morphological variability in competence and performance. The Interpretability Hypothesis (Hawkins & Hattori, 2006; Tsimpi & Dimitrakopoulou, 2007) argues that uninterpretable features cannot be acquired in the L2 if they are not selected and exemplified by the native language. Studies supporting such an account in the area of inflectional morphology are Franceschina (2001, 2005) and Tsimpi and Mastropavlou (2008). The latter study investigated acquisition of pronominal object clitics in Greek by Russian and Turkish early bilinguals. The participants had been using Greek for at least eight years at the time of testing and were considered to be advanced learners. The behavioral predictions depend on a specific analysis of clitic features. In the authors' analysis, first- and second-person clitics, as well as genitive clitics, exhibit interpretable person features which allows them to point to various referents in the discourse. Unlike first- and second-person and genitive clitics, however, third-person object clitics are proposed to carry a bundle of uninterpretable features. The prediction is that the bilinguals will be much more accurate in producing the former but not the latter. Tsimpi and Mastropavlou
found, indeed, that their experimental participants produced first- and second-
person clitics significantly more often than third-person clitics, and that genitive
clitics were produced more reliably than object ones. They interpreted these find-
ings to mean that third-person object clitics are unacquirable even at very advanced
levels of proficiency. We will come back to this hypothesis in the section discussing
processing, where its predictions will be checked with other uninterpretable fea-
tures: the gender agreement between nouns and adjectives. In the next two sec-
tions, we turn to verbal and nominal morphology separately, to track the sources
of morphological variability.

Verbal inflectional morphology

In this section, I will survey some recent studies on person, number, and gender
agreement morphology as marked on the verb. The common denominators of
these studies are the participation of advanced learners as well as the treatment of
the morphology as only an external exponent of deeper grammatical knowledge.
It has to be announced from the start of the section, however, that most research on
verbal paradigms focuses on child second language learners and on instructed
learners, only some of whom are advanced.

The Missing Surface Inflection Hypothesis (MSIH, Haznedar & Schwartz, 1997;
Prévost & White, 2000) has already been invoked above as an explanation of the
dissociation between production of the inflectional morphology and underlying
knowledge of grammatical features. Bruhn de Garavito (2003) and Herschensohn
(2001) provide additional evidence for the main claims of this hypothesis. One pre-
diction is that variation and errors would be more frequent in production than in
comprehension, a prediction challenged by McCarthy (2008). Since all of the
studies mentioned in this paragraph test beginning to intermediate learners, we
cannot make any important conclusions for the competence and performance of
advanced learners. This is hardly surprising because the MSIH is an account of
variation in non-stable grammars. If it is the case that production and even com-
prehension of morphology lags after syntax, then we expect errors to be largely
gone by advanced proficiency levels.

What the MSIH fails to predict, however, is which form in a morphological par-
adigm will be adopted as a default by the learner. A level of individual variance is
implied, whereby some learners may choose one form as a default while other
learners may choose another. To address this gap, McCarthy (2012) proposes the
Morphological Underspecification Hypothesis (MUH). This hypothesis is based
on a universal feature hierarchy (Harley & Ritter, 2002) and suggests that in lan-
guage acquisition the default, or morphologically least specified forms, will be the
ones to be acquired early and with fewer errors. They may also be substituted
when a more complex form is not reachable in the mental lexicon. Following this
hierarchy, non-finite forms should be acquired before finite forms, third-person
before non-third-person, first-person before second-person, and singular before
plural. For instance, the third-person value is not associated with any person and
number features. Consequently, third-person forms will have fewer features than
the more highly specified first- or second-person forms, and may be inserted in
first- and second-person contexts, if the mental lexicon does not oblige by providing
the needed form.

The data in McCarthy (2012) came from learners at two stages of development
in the Spanish Learner Language Oral Corpora (SPLLOC, 2010), more specifically
the guided interview portion of the corpus. McCarthy checked the development
predictions for three grammatical features in spontaneous oral production: finiteness,
person, and number. She found that indeed third-person acted as a default in
first-person contexts, and non-finite forms were used as defaults in finite contexts.
(There were very few second-person substitution errors beyond the low-proficiency
level, but also few occasions for the learners to use the second-person in the
interview.) Even keeping these limitations in mind, the underspecified member of
the paradigm, third-person, appeared to be acquired first and to be sometimes
substituted when other members of the paradigm were attempted. Non-finite
forms were also demonstrated to be defaults. For the feature of number, while
some development was observed, there was no clear developmental trajectory. In
sum, McCarthy (2012) showed systematic development of inflectional morphology
that was predicted by a universal, arguably innate, feature hierarchy. It is inter-
esting to note that this research direction is not interested in checking the syntax-
before-morphology prediction, but considers the inflectional morphology in its
own right with its own developmental trajectory.

In a more recent study, Blom, Chondrogianni, Marinis, and Vasic (2016) investi-
gated the MUH with Turkish-native children acquiring Greek and Dutch as second
languages. Since we have to be mindful of learner proficiency in this chapter,
we should state that the children were tested at the average age of 7.1 (range; 4.8–
8.8), after being exposed to the second language for over 2.5 years. While the par-
ticipants were rather accurate in using agreement inflection, their performance
was not error-free. An important finding of this study was that the MUH was
largely supported in the acquisition of Greek but not in the acquisition of Dutch.
The authors argue that language-specific properties conspire with the universal
feature hierarchy. The MUH appears to be more successful at explaining the acquisi-
tion patterns in Greek because it is a language with rich agreement morphology,
while Dutch is a language with poor agreement morphology (just like English)
and it takes children less time to become highly accurate with the inflectional
forms. One caveat that the reader should keep in mind is that the findings of this
study, while in line with monolingual L1 acquisition, may not be easily extended
to adult L2 learners. As Schwartz’s (2009) Domain by Age model argues, child L2
acquisition is like adult L2 acquisition in the domain of syntax, and like L1 acqui-
sition in the domain of inflectional morphology.

A fruitful line of investigation that has been pursued since the very beginning
of generative second language acquisition research is investigation of null subject
interpretation (Belletti, Benatti, & Sorace, 2007; Hilles, 1986; White, 1985; see also
discussion in Chapter 23). In order to understand who the missing subjects refer to
in Spanish and Italian, L2 speakers have to make use of the verbal person and
number endings. There is a widely held consensus in the literature that learners, especially those at advanced proficiency, have no trouble at all identifying the missing subject. This finding has been confirmed with various experimental methods. Learners have more trouble when subjects are overt, but that is because they carry additional discourse information about topic and focus. We have to keep in mind, however, that for the most part context also helps the learners to identify the missing subject, not just the verbal inflection.

A look at a rare property, ergative agreement, will conclude this section. Rodríguez-Ordóñez (2015) considers adult learners of Basque who are Spanish-native and Spanish-dominant. Basque verbs agree with the subject and the object in person, number, and case. Basque is an ergative language so the agreement pattern is structurally different from the Spanish one. In addition, auxiliary selection is also involved. As a result, the Basque paradigm is much more complex: just for the present indicative, Spanish verbs have six different forms while Basque has 102. We shall focus here on Rodríguez-Ordóñez’s advanced participant group \( (n=8) \), who started learning Basque after the age of 12. The test instruments were an oral interview, an elicited production task, and an acceptability judgment task (AJT). The advanced participants were over 99% accurate on the oral production tasks, including subject–verb agreement and auxiliary selection. In the AJT, they were at ceiling on accepting correct endings for all types of verbs; they also rejected the unacceptable endings reliably. The author argues that the L2 advanced speakers demonstrate superior knowledge of auxiliary selection and verbal case agreement. Results from nominal inflection in L2 Basque will be discussed in the next section.

In explaining these results, we must be mindful of the elevated exposure to Basque these learners enjoy in the Basque country. Although they are Spanish-dominant bilinguals, the L2 input they get is clearly sufficient for them to achieve inflectional mastery of the complex paradigm.

In summary, in this section we considered additional evidence for the MSIH. We discussed the MUH, which found support from instructed young learners and from child L2 learners. It was suggested that only in languages with rich morphology can a developmental trajectory be clearly detected, as the learning of morphologically poor paradigms may occur too fast. Copious research on null subject interpretation shows that learners make use of the verbal inflection to identify missing subjects. Finally, in a comprehensive study of Basque ergativity marking, advanced learners were demonstrated to have completed paradigm acquisition successfully.

**Nominal inflectional morphology**

The acquisition of gender and number has been the subject of a great deal of research using a variety of methodologies and theoretical frameworks. Within the generative tradition, both the mental representation and the processing of gender and number have been extensively investigated (Bruhn de Garavito & White, 2002; Franceschina 2001, 2005; Sagarra & Herschensohn, 2011; White, Valenzuela,
Kozlowska-Macgregor, & Leung, 2004). Combining research findings from different studies suggests that while gender marking knowledge is acquired fast to relatively high accuracy, gender errors persist to very advanced levels of proficiency. To wit, Bruhn de Garavito & White (2002) found that learners were 81.5% target-like in performance with gender agreement after a year of studying Spanish. On the other hand, Franceschina’s (2001) near-native speaker of Spanish continued to make gender errors after living in a Spanish-speaking environment for 24 years.

However, when more research is taken into account, it is clear that advanced learners are highly accurate with number and gender agreement. White (2003, p. 136, table 4.6) summarizes research supporting this claim. To take a concrete study, White et al. (2004) used the phenomenon of N-drop to test unconscious knowledge of gender and number in Spanish L2. As the example in (12) illustrates, it is possible to drop the noun in Spanish while the adjective and/or determiner gender and number morphology licenses the missing noun.

\[(12) \quad \text{¿Me compro este negro?} \]
\[\quad \text{clitic buy this-masc.sg black-masc.sg} \]
\[\quad \text{‘Shall I buy this black one?’} \]

The answer to the question in (12) has to be a masculine noun. In a picture choice task, subjects had to identify which object (out of three) was the appropriate answer, agreeing in number and gender with the features provided in the test question. Learners from English and French backgrounds were tested. Advanced and even intermediate groups showed considerable accuracy on both number and gender features, and in fact were no different from native speakers. In addition, the fact that French has grammatical gender, unlike English but like the target language, does not appear to have played any role. Learners were equally accurate regardless of their native language.

In more recent work, Bruhn de Garavito and Otalora (2016) test whether learners are aware of the fact that under noun ellipsis, the elided noun and its antecedent can differ in number, but they cannot differ in gender, as illustrated in (13).

\[(13) \quad \text{Visité a mis maestros y tú visitaste a los tuyos / *a las tuyas / al tuyo} \]
\[\quad \text{visited-I my teachers-masc.pl and you visited yours-masc.pl/*yours.fem.pl/} \]
\[\quad \text{yours-masc.sg} \]
\[\quad \text{‘I visited my teachers and you visited yours.’} \]

The researchers show that their L2 learners are able to access the gender and number features of the elided noun and then are able to compare these features to those of the antecedents. In addition, just like the native controls, they are sensitive to the prohibition of gender mismatch. This is just the last of a long list of studies documenting unimpaired mental representation of gender as well as number features, especially in advanced learners.
In thinking about the gender feature, it is important to distinguish knowledge of the agreement process (e.g. nouns agree in gender with determiners and adjectives) and accurate knowledge of the gender of each individual noun in a language. The latter is known as “gender assignment” and is lexical, not grammatical knowledge. The studies we reviewed in this section are looking at gender agreement, and they check gender assignment as a condition on participation in the experiment, or make gender assignment clear as part of the experimental design. This interesting dichotomy will be explored in the next section.

Dissociations between representation and processing

The majority of studies that have documented indisputable evidence that functional morphology can be acquired to native-like levels by second language learners have employed a variety of off-line tasks. Among such tasks are truth-value judgment tasks for interpretation and grammaticality judgment tasks for form acceptability, as well as various elicited production tasks. It is in principle possible that some of the high-level performance attested with such tasks is due to non-automatic or learned responses based on metalinguistic teaching. In order to tease the contribution of rote learning from natural language processing, researchers turn to on-line processing measures. However, we must be clear from the outset of this section that such measures do not replace but are complementary to acceptability, interpretation, and elicitation measures. All studies discussed in this section find that L2ers’ sensitivity effects are not that pronounced, or that language processing is slower and less efficient, in comparison to that of monolinguals, who are not operating with two competing grammars in their brain.

In principle, it is possible for a learner to have the underlying morphological knowledge pertaining to a functional category in place, but still to exhibit processing effects such as being accurate on more local operations than on long-distance ones. We shall see such examples below. The relationship between processing and underlying knowledge is a complex one. On identifying non-native-like processing in L2 users, one can conclude that it is indeed solely a processing issue, or one can conclude that learners’ mental representations are impaired. Therefore, it is important for processing studies to conduct grammatical competence tests as well.

An example of the first type of conclusion is provided by Hopp (2010). Hopp tested Dutch, English, and Russian native advanced learners of German on a complex construction called “scrambling” depending on case, and subject–verb agreement knowledge. While all three groups’ behavior was similarly accurate on an off-line grammaticality judgment task as well as on a self-paced reading task, differences emerged when the researcher applied a speeded grammaticality judgment task. Only the Russian native speakers were reliably able to detect the case marking violations under processing pressure, while the Dutch and English groups could not. Hopp interpreted this finding as evidence for lingering L1 influence even at the level of ultimate attainment. Since Russian
has case marking but Dutch and English do not, it seems that this language feature has contributed to those learners being sensitive to case-marking violations. The grammatical process of accessing and computing overt case endings may have been transferred from the native language. Hopp concluded that the near-natives’ lack of automaticity is a reflection of slower and less efficient processing, which is not qualitatively but just quantitatively different from that of native speakers.

A very similar finding is interpreted in the opposite way by Jiang, Novokshanova, Masuda, and Wang (2011). In this experimental study, the researchers compared Russian and Japanese native speakers’ sensitivity to plural agreement marking in their L2 English. It turned out that only the Russian group performed in a native-like manner, while the Japanese group did not. The explanation, just as in Hopp’s study, was based on Russian (but not Japanese) having overt plural morphology. According to Jiang et al. (2011), the overt marker coming from the native language aids the L2 processing of the same feature. These results are in keeping with Jiang’s (2004, 2007) findings, supporting his Morphological Congruency Hypothesis. This hypothesis argues that native-like competence with respect to a morphologically expressed feature can be reached only if the native language expresses the same feature with an overt morpheme. Thus, we see that on-line and off-line performance dissociation is interpreted as a competence fault by Jiang and colleagues, but as pointing to a processing issue by Hopp.

An interesting study that directly addresses the two possible sources of processing difficulties is Keating (2009). This is an eye-tracking study that exploits the proximity of the agreeing elements, in this case gender agreement between the noun and adjective, as exemplified in (14)–(16).

(14) Agreement within the NP:

Una cas-a pequeñ-a cuesta mucho en San Francisco.

‘A small house costs a lot in San Francisco.’

(15) Agreement in the matrix clause VP:

La cas-a es bastante pequeñ-a y necesita muchas reparaciones.

‘The house is quite small and needs a lot of repairs.’

(16) Long-distance agreement in the subordinate clause VP:

Una cas-a cuesta menos si es pequeñ-a y necesita reparaciones.

‘A house costs less if it is small and needs repairs.’

In the eye-tracking data, both longer fixation times on the critical elements (in this case the adjective pequeñ-a) and regressions from the adjective back to the noun for (literally) checking the agreement indicate sensitivity to the grammatical property as well as processing difficulties. The learners’ behavior showed an effect of proficiency as well as an effect of the distance between the noun head and the adjective. The native speakers displayed a response pattern including fixation and subsequent regression for all three noun–adjective distances. Only the highly advanced learners revealed the same native-like pattern of behavior, and only for the most local configuration as in (14). However, if we consider the
stimuli from the point of view of grammatical representations and linguistic competence, the same formal features are involved in the noun–adjective agreement no matter what the distance between head and modifier is. Therefore, Keating argued that his eye-tracking data suggested successful acquisition but more labored processing. His learners had acquired the feature-checking mechanism, but had a harder time applying it at a distance, in cases where they had to keep more words in working memory. These findings address the claims of the Morphological Congruency Hypothesis, since they show that underlying linguistic competence may not always translate into seamless morphological processing, particularly where that processing involves more resource-consuming operations.

In addition to L1 influence and proximity of the agreeing elements, other factors have been recently discussed in the psycholinguistic literature as influencing the processing of L2 functional morphology. The first additional factor to examine here is attention to form or meaning, as suggested by Roberts and Siyanova-Chanturia (2013). It turns out that learners in an eye-tracking experiment reported by Foucart and Frenck-Mestre (2012) did spend longer fixated on an error equivalent to (15) in L2 French, while Keating’s L2 Spanish learners did not. In Keating’s study, however, learners performed a meaning-related task, while in Foucart and Frenck-Mestre’s experiment, participants’ attention to grammar may have been heightened because they were asked to offer grammaticality judgments.

Second, in a study addressing the contribution of phonology to morphological processing, Carrasco-Ortiz and Frenck-Mestre (2014) manipulated subject–verb agreement in such a way that it was either correct, incorrect but orally realized, or incorrect and silent, not pronounced. This is possible in French where a fair amount of agreement morphology is silent. It is important to point out that the participants read sentences silently while their brain activity was captured through event-related brain potentials (ERP). In other words, the experimental factor was present in the language but not in the presentation of the stimuli. The results revealed more robust ERP responses, that is, more sensitivity to errors, when the inflection errors were orally realized in French as opposed to when they were unpronounced. These findings suggest that phonological cues actually enhance learners’ sensitivity to morphology, even though both types of cues can be used when they are present.

Finally, another important factor that we have been mentioning is lexical access to the functional morphemes and words. Earlier, I made the distinction between gender agreement computation and gender assignment, i.e., associating a noun with its gender class as part of its mental lexicon entry. The lexical access factor is tested by employing shared or divergent gender assignment stimuli. A number of recent studies suggest that even if languages have overt grammatical gender marking, the way they mark gender on the morphological and lexical level can vary. For example, Klassen (2016) reports that Spanish-German bilinguals show different responses in lexical decision and naming to neuter gender in German, vis-à-vis masculine and feminine, which have counterparts in the Spanish
lexicon. Bilingual L2 gender processing appears to be affected by shared lexical representations. Weber and Paris (2004) also report that bilinguals activate the gender of L1 words in their L2 processing.

A recent study taking lexical representation of gender into account (Grüter, Lew-Williams, & Fernald, 2012) explored the production, comprehension, and processing of very advanced L2 learners of Spanish. Processing sensitivity was measured using the eye-tracking-while-listening paradigm: Can learners look at the picture of the noun with the correct gender at hearing the introducing gender marker? Different-gender and same-gender trials were compared: only in the former was faster shifting to the target image expected. Both familiar and novel nouns with canonical gender marking were used. All participants were at ceiling in off-line comprehension; errors in production were mainly due to wrong gender assignment. Results of the eye-tracking experiment showed that while native speakers used gender marking predictively, the advanced learners in this study were not able to do so, on familiar real nouns trials. However, when they heard novel nouns, learners exhibited a contrast between the different-gender and same-gender conditions. The researchers concluded that lexical rather than syntactic properties of gender are the source of difficulty even for very advanced learners. At the same time, the findings from novel noun processing suggest that these learners can in principle use gender cues predictively. The slower retrieval time of gender information was attributed to “less effective use of gender cues in online processing” (p. 210).

In a continuation of this line of research, Hopp and Lemmert (2016) investigated similar predictive gender processing among Russian learners of German. Russian and German being languages with lexical gender marking, enduring influence of the native language was expected. The advanced participants demonstrated native-like gender prediction throughout. At lower levels of proficiency, lexical congruency (that is, whether the noun was assigned to the same gender class in Russian and German) had an effect on prediction. Syntactic congruency, that is, whether gender is realized on determiners as in the L2 or as a suffix on the noun as in the L1, also had an effect. The researchers demonstrated that processing was affected not only by L1–L2 syntactic congruency in gender marking but also by lexical congruency of individual nouns. They suggested a developmental trajectory from a more lexically based to a native-like syntactic processing of gender assignment.

Processing studies have recently been used to check predictions of impaired competence accounts comparing between features, too. For example, Alemán Bañón, Fiorentino, and Gabriele (2014) examined the processing of grammatical gender and number in English L1 learners of L2 Spanish, arguing that ERPs could be used to test the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007). Alemán Bañón et al. (2014) proposed that if the hypothesis was correct, then one should expect L2 learners to show qualitatively different processing for the purportedly unacquirable L2 features. Since English has grammatical number but not grammatical gender, the predictions of the Interpretability Hypothesis would be
that the L2 learners might show evidence of native-like processing for number violations, but not for grammatical gender violations. In the event, the study showed that by advanced stages of L2 acquisition, English learners of Spanish do have qualitatively similar processing for both gender and number violations. In an extension study, Alemán Bañón, Miller, and Rothman (2017) combined on-line and off-line measures to look at gender and number knowledge while checking for markedness and default effects (as in MacCarthy, 2008). Since advanced learners were able to process gender and number in a native-like way, the researchers argued that markedness explanations of L2 morphological variation were not supported by their findings.

To sum up this section, L2 learners exhibit variable processing behaviors regarding sensitivity to morphological marking. Current research is addressing these explanations with a variety of processing measures. When off-line and on-line measures of morphological processing differ in the same group of learners, one can entertain both an impaired competence and an impaired performance accounts. Attested with some properties such as plural marking in Jiang et al. (2011), this type of behavior was interpreted as due to L1 influence and demonstrating impaired linguistic competence. When functional morphology processing is native-like, however, there is no doubt that native-like competence is also in place. The Hopp and Lemmert (2016), Keating (2009), and Alemán Bañón et al. (2017) studies reviewed here, especially the latter, provide such evidence.

Conclusions and directions for further research

Adult L2 learners often exhibit variability in their use of inflectional morphology, even at very high levels of proficiency and across the verbal and nominal domains. In this chapter, we defined morphological variability as a learner’s inconsistent use of obligatory functional morphology. The literature on morphological variability is quite extensive and growing every day. Some generalizations that have emerged from the literature are that functional morphology errors are systematic, but that nevertheless some properties exhibit considerably more variability than others. Following White (2003) and Lardiere (1998a, 1998b), I suggested that mere suppliance of the inflectional morphemes in obligatory contexts probably under-represents learner knowledge of functional categories. This conclusion was supported by findings showing that the same learners who underproduce inflectional morphemes are actually quite accurate with the syntactic effects and semantic interpretation of those morphemes.

To account for the asymmetry in variability, it has been proposed that L2ers resort to morphological defaults, or forms that are underspecified and hence easier to insert under fully specified syntactic nodes. This markedness effect is attributed to processing pressures by some (the Missing Surface Inflection Hypothesis) but interpreted as impaired competence by others (the Morphological Under specification Hypothesis). Furthermore, variable production of morphology exhibiting
uninterpretable features is also attributed to impaired representations (the Interpretability Hypothesis).

These accounts are being investigated these days by comparing off-line production and comprehension with on-line processing measures. Several factors have been identified as affecting advanced competence and performance. Those include:

1. the syntactic and morphological differences between the native and second language;
2. lexical access difficulty and speed in L2ers;
3. unmarked (default) versus marked forms.

However important, these factors do not lead to advanced learners and native speakers behaving in a qualitatively different manner. Functional morphology may be a challenge for L2 learners, but it is not an insurmountable one.

Future research should advance current explanations by investigating additional properties beyond gender and number, by involving fresh pairs of languages, and by comparing lexical and morphosyntactic competence with real-time processing. Time pressure experiments can also reveal whether native and near-native processing are qualitatively the same or different.

NOTES

1 I use small caps to set out grammatical meanings with predominantly semantic content, as opposed to purely grammatical meanings such as subject–verb agreement.
2 In this view, the form of inflectional and derivational affixes is separated from their function, as we argued above. Beard distinguishes two types of rules: one type regulates grammatical processes (e.g. plural), while the second type of rules are affixation rules that spell out grammatical functions. For example, English plurals are formed in a number of ways, as is shown in (i):

(1) girl–girls, bus–busses, criterion–criteria, tooth–teeth

Under the Separation Hypothesis there is a single rule of plural formation, which adds the feature [plural]. The resulting abstract morpheme is input to different rules of the second type, and these rules spell out the actual phonological form of the plurals in (i). A similar idea is also captured by the Distributed Morphology framework (Halle & Marantz, 1993). This separation of form and function in morphology has proved fruitful in explaining acquisition.
3 Patty was born in Indonesia and lived with her family among the expatriate Chinese community. She moved to China as an adolescent and then to Hong Kong to complete high school. Patty acquired most of her English as an adult immigrant to the United States.
REFERENCES


Rodríguez-Ordóñez, I. (2015). The acquisition of nominal and verbal inflectional morphology: Evidence from


21 Advanced Lexical Development

STUART A. WEBB
University of Western Ontario, Canada

Introduction

The vast majority of studies of lexical development have tended to focus on beginner- and intermediate-level learners. When research has looked at more advanced learners, attention has mainly centered on the extent to which L2 learners have been able to attain a native-like knowledge of words (e.g. Forsberg Lundell & Lindqvist, 2012; Mizrahi, 2016; Siyanova & Schmitt, 2008; Zareva, 2005). This line of research has been useful because it highlights the different ways that we might define advanced lexical development. However, it also reveals the need for a better understanding of what is involved in vocabulary learning and how we might help learners progress in their lexical development. In this chapter we will look at different ways to define advanced lexical development, how learners might work toward achieving an advanced level of lexical proficiency, how advanced learners might continue to improve their vocabulary knowledge, and what research is needed on the topic of advanced lexical development.

How might we define advanced lexical development?

Vocabulary size

Researchers have long agreed upon the importance of lexical knowledge to language learning. Words are building blocks of communication; it is difficult to convey our intended meanings effectively without using words, and it is difficult to understand a great deal without understanding the words that we encounter. Vocabulary is thus at the heart of language learning.
There are no established definitions of beginner, intermediate, and advanced levels of lexical development. This may be due in part to the fact that vocabulary knowledge is a multidimensional construct and so while one approach to measuring lexical proficiency may differentiate between stages of lexical development, another approach may indicate quite different findings. Perhaps the most transparent way of defining lexical proficiency is by the number of words that are known. This has been perhaps the most common way of establishing a person’s lexical knowledge. Those who do not know many words might be classified as being at a beginning stage of lexical development, while those who know a great number of words might be classified as being at an advanced stage.

Research has shown that educated native speakers of English know about 15,000–20,000 word families (a word family is made up of a headword such as smile, its inflections smiles, smiled, smiling, and its derivations smilingly, unsmiling), so we can look at the extent to which L2 learners’ knowledge compares (Biemiller & Slonim, 2001; Goulden, Nation, & Read, 1990). Research indicates that advanced L2 learners may be able to score as well as native speakers on such tests (Arnaud & Savignon, 1997). However, it also indicates that the majority of learners in English as a foreign language (EFL) contexts learn a relatively small proportion of these words (e.g. Danelund, 2013; Webb & Chang, 2012).

Although a comparison between non-native and native speaker vocabulary size may show the extent of progress in vocabulary learning, the threshold for different levels of lexical proficiency would be arbitrary. However, lexical profiling studies may provide a more objective target for advanced lexical proficiency. Research on the lexical profiles of different text types builds on studies of vocabulary size by indicating the number of words that are necessary to achieve different language learning goals. Lexical profiling studies reveal the proportion of words at different word frequency levels in a text or corpus. This line of research is based on studies of lexical coverage (the percentage of known words in a text) that have indicated that a text may be understood if a certain proportion of words is known. Studies investigating the lexical coverage of written text tend to suggest that readers may achieve adequate comprehension if 98% or more of the words are known (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). Research examining the lexical coverage of spoken text tend to suggest that listeners may understand speech when they know 95% or more of the words that they hear (van Zeeland & Schmitt, 2012). These findings are supported by research investigating the extent to which vocabulary size is correlated with comprehension. Research has consistently shown that the greater the vocabulary size, the higher the comprehension test score (Laufer & Ravenhorst-Kalovski, 2010; Qian, 1999; Stehr, 2009). Laufer (1992) found positive correlations ranging from .50 to .75, Qian (2002) reported a correlation of .74, Qian (1999) found a correlation of .78, and Laufer and Ravenhorst-Kalovski (2010) reported a correlation of .80 between reading comprehension and vocabulary size. In a study looking at the relationship between listening comprehension and vocabulary size, Stehr (2009) found a large positive correlation of .70. Taken together, the research indicates the large impact that vocabulary knowledge may often have on reading and listening comprehension. The explanation for this
is that as lexical knowledge increases, a greater proportion of the words within spoken and written text are known, and as a consequence there is a greater precision in understanding. In other words, when vocabulary size increases, comprehension is also likely to increase (Webb & Paribakht, 2015).

Over the past 30 years, corpus-driven studies have indicated the lexical demands of different types of discourse. Lexical profiling research indicates that the most frequent 3,000 word families may provide 95% lexical coverage of speech (Adolphs & Schmitt, 2003; Nation, 2006), television programs (Rodgers & Webb, 2011; Webb & Rodgers, 2009a), and movies (Nation, 2006; Webb & Rodgers, 2009b) and that the most frequent 4,000 word families may provide 95% of academic spoken discourse (Dang & Webb, 2014). Corpus-driven studies investigating the vocabulary size necessary to reach 98% coverage of written text indicate that there can be a great deal of variation among the vocabulary in written texts, with knowledge of the most frequent 8,000–9,000 word families, 10,000 word families, and 14,000 word families necessary to reach 98% coverage of novels and newspapers (Nation, 2006), texts written for young learners (Webb & Macalister, 2013), and the reading comprehension passages found in English proficiency tests (Webb & Paribakht, 2015), respectively. Webb and Nation (2017) suggest that these vocabulary sizes make appropriate learning goals for students at different levels. Table 21.1 shows how different vocabulary sizes might be linked to proficiency level.

Table 21.1  Defining lexical development by vocabulary size.

<table>
<thead>
<tr>
<th>Level</th>
<th>Vocabulary size</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>800 lemmas</td>
<td>Understand 75% of the words in spoken and written text</td>
</tr>
<tr>
<td>Intermediate</td>
<td>3,000 word families</td>
<td>Understand most forms of spoken discourse</td>
</tr>
<tr>
<td>Advanced</td>
<td>8,000–9,000 word families</td>
<td>Understand written text without external support</td>
</tr>
</tbody>
</table>

Perhaps the most useful test to measure vocabulary size is the Vocabulary Size Test (Coxhead, Nation, & Sim, 2015; Nation & Beglar, 2007). It indicates the extent to which test takers can recognize the meanings of words. Thus, it provides a measure of receptive knowledge of words. Receptive lexical knowledge refers to the ability to understand words when they are encountered in writing or speech, whereas productive lexical knowledge refers to the ability to produce L2 words. It is important to note that receptive lexical knowledge tends to be easier to gain and is usually acquired prior to productive lexical knowledge (e.g. Waring, 1997; Webb, 2008). At present there are no established tests designed to measure productive vocabulary size (although Laufer and Nation’s (1999) Productive Levels Test has been used for this purpose). Tests of productive lexical knowledge are of great value for assessing advanced learners because the use of words rather than the recognition of words may be of particular importance at this stage of learning.
Lexical sophistication

A second way that we might reveal advanced vocabulary development is by measuring lexical sophistication. Lexical sophistication is an assessment of the extent to which lower frequency words are used in speech or writing. The rationale behind its use as a measure of lexical knowledge is that beginning learners tend to use a large proportion of high frequency words in their speech and writing, but as they become more advanced, they tend to use fewer high frequency words and more lower frequency words. For example, a beginning learner might say ‘The woman talked to the man.’ While a more advanced learner might convey a similar but more precise meaning in the following ‘The actress chatted with the actor.’ The rationale behind using lexical sophistication as a measure of lexical proficiency is that higher frequency words tend to be learned before lower frequency words (e.g. Schmitt, Schmitt, & Clapham, 2001), so a greater proportion of lower frequency words used is likely to indicate further development in lexical knowledge. Research by Laufer and Nation (1995) provides support for this. They examined the distribution of words at different frequency levels in short texts written by learners at different levels of proficiency and found that the lexical frequency profiles of the texts were able to distinguish between the different groups of learners. Meara and Bell’s (2001) P-Lex application has also been found to accurately differentiate between learners at different levels by analyzing the lexical sophistication in a text.

The challenge with using lexical sophistication as an indicator of lexical development is that other factors such as topic and background knowledge might also influence the proportion of higher frequency words used in a text. Similarly, the ability to produce lower frequency words does not ensure that they are used correctly. Software such as Range (Nation & Heatley, 2002) is used to measure lexical sophistication and it does not take into consideration whether there are errors in the text that is analyzed. Moreover, it may be that the speaker or writer is able to produce lower frequency words but simply chooses not to. After all, it can be argued that L1 texts that include a larger proportion of lower frequency words would not always necessarily be judged as superior to those that include a smaller proportion of these words. Because there will be variability in the lexical sophistication of different texts it is probably best to look at sophistication scores over a range of texts to best determine its influence on lexical development. Despite these limitations, there is inherent value in examining the words that are produced by learners as a measure of vocabulary learning progress.

Formulaic language

A third method that can be used to indicate advanced lexical development is the degree to which learners can successfully recognize or use formulaic language. Perhaps the easiest way to distinguish advanced learners from native speakers is through the use of odd combinations of words. For example, Chinese learners might use the combination open the TV or close the light because these combinations
occur in their L1. Similarly, Japanese learners may sometimes say *take lunch* rather than *have lunch* because this is the more typical sequence in their L1. It is these unusual sequences of words that will often reveal someone to be a learner rather than a native speaker. Research has compared the degree to which advanced learners and native speakers have knowledge of multi-word combinations. Arnaud and Savignon (1997) investigated the extent to which advanced and highly advanced English language learners could recognize collocations. They found that none of the non-native participants were able to score in the native speaker range. Forsberg, Lundell, and Linqvist (2012, 2014), however, revealed that advanced and very advanced L2 learners of French had scores consistent with those of native speakers on a test measuring productive knowledge of collocations. Moreover, the participants in their study had been exposed to the L2 for between 5 and 35 years.

Measuring knowledge of strings of words, however, poses some challenges. Although there have been several studies that were designed to measure knowledge of collocation, at present there are no established tests that might be used for this purpose. Perhaps the most useful approach to revealing an advanced knowledge of multi-word items would be to measure knowledge of word combinations at different frequency levels in a test format similar to the Vocabulary Levels Test (Nation, 1983; Schmitt, Schmitt, & Clapham, 2001; Webb, Sasao, & Ballance, 2017). An initial study using this approach suggests that learners’ knowledge of L2 word combinations does diminish as the words found in the items become less frequent (Nguyen & Webb, 2016). This is similar to findings with individual words (Schmitt, Schmitt, & Clapham, 2001; Webb & Chang, 2012) and suggests that looking at word strings according to their frequency levels might provide a useful indicator of lexical development.

The most common method of investigating knowledge of formulaic language is by examining the word combinations produced by L2 learners in their writing or speech. This line of research has shown that non-native speakers tend to produce many unusual multi-word combinations, that they tend to overuse the most frequent multi-word items and underuse less frequent sequences (Durrant & Schmitt, 2009; Granger & Bestgen, 2014), and that it takes a long time to learn formulaic language (Levitzky-Aviad & Laufer, 2013; Li & Schmitt, 2010). Thus, the degree to which students can successfully produce natural strings of words may indicate the extent to which they have mastered an L2.

**Lexical processing**

A fourth indicator of advanced lexical knowledge is the speed at which learners can process words or multi-word combinations. Processing speed provides an indication of lexical fluency; processing words at a fast rate similar to that of native speakers provides an indication of a person’s level. For example, Siyanova and Schmitt (2008) compared advanced L2 learner and native-speaker performance on a judgment task examining the acceptability adjective-noun collocations. Their findings revealed that the native speakers’ reaction times were faster and more accurate than those of the advanced L2 learners. Siyanova and Schmitt also looked
at exposure to English as a factor within their study and found that L2 learners who had greater exposure to English tended to perform at a higher level than those who did not.

It is intuitively logical that lexical processing speed may indicate stages of lexical development because the speed with which we can comprehend the words that we encounter, and the rate at which we can produce words, may reveal how fluent we are with a language. Comparisons between native and non-native speakers may indicate the extent to which L2 learners have attained a native-like lexical proficiency (Mizrahi, 2016).

Taken together, we find that there are a number of different ways that advanced lexical development can be defined. The easiest and perhaps most pedagogically useful approach may be to determine whether learners have reached a vocabulary learning target such as knowing the most frequent 8,000–9,000 word families that should allow them to understand most written texts without support. The problem with using vocabulary size as the sole indicator of lexical development is that it will not indicate the degree that learners will be able to successfully use words. Similarly, if we use another indicator as the sole criteria for advanced lexical development, it is unlikely to reveal the extent to which learners are able to achieve other aspects of lexical proficiency. Thus, the most accurate assessment will include the use of multiple measures and acknowledge where measurement is lacking.

**What are the keys to helping learners reach a level of advanced lexical development?**

Many L2 learners fail to attain a high level of lexical proficiency. In fact, research indicates that a large proportion of EFL students fail to master the most frequent words after many years of study. For example, Webb and Chang (2012) found that after nine years of formal English language instruction, only 47% of EFL students in Taiwan had mastered the most frequent 1,000 words. Danelund (2013) found similar results in Denmark. Her study showed that after nine years of study, only 48% of high school students knew the most frequent 2,000 words. Research investigating the number of words per year tends to show a great deal of variation among learners. Milton (2006) found that on average EFL learners in Greece who received about 100 hours of instruction per year tended to learn approximately 500 lemmas per year (a lemma is made up of a headword such as *play* and its inflections: *plays*, *played*, and *playing*). However, the most productive students were able to learn as many as 1,000 lemmas per year while the least productive learners made little to no gains in lexical development. Very similar findings were reported by Orosz (2009) with primary school students in Hungary. He found that on average students were able to learn 300 to 400 lemmas per year, with some students learning as many as 1,000 lemmas while others made no progress. Webb and Chang (2012) also reported large differences in lexical growth among EFL learners in Taiwan, with the number of words learned per year ranging from 18 to 430 word families among different groups of learners.
Reaching an advanced standard of lexical development solely through classroom-based instruction may not be realistic. If we consider vocabulary size as the indicator of advanced lexical proficiency (receptive knowledge of 8,000–9,000 word families), there is unlikely to be sufficient classroom time to achieve this goal. It is rare that the primary aim of a language learning course is vocabulary learning, instead vocabulary tends to be taught in relation to the material that is encountered by students. Although vocabulary may be learned in every class, the end result may be a relatively small number of words that are deliberately learned in the classroom, and only a proportion of these words may be retained.

Although relatively few learners in the EFL context may reach an advanced level of lexical development, there is reason to be optimistic about increasing the proportion of learners who achieve this level of proficiency. This is because principled vocabulary learning programs are rare. As mentioned above, words are often taught as a byproduct of completing another task such as reading a passage, understanding content, or writing an assignment. It is likely that if greater emphasis were placed on learning vocabulary within language learning programs, the lexical development of students would improve (Webb & Chang, 2012).

The four strands of a vocabulary learning program

Nation (2007) suggests that a language learning program should include the four strands of meaning-focused input, meaning-focused output, language-focused learning, and fluency development, and Webb and Nation (2017) discuss how each strand contributes to vocabulary learning. Meaning-focused input refers to situations where language is encountered and the focus is on understanding the input (rather than learning specific aspects of language). Meaning-focused output refers to situations in which language is used to communicate, such as in conversation, presentations, and writing emails. Language-focused learning involves learning specific language features. In this strand students deliberately learn words in activities and exercises. Fluency development involves trying to use words in a more native-like way. This typically involves processing and using words at a faster rate. The value of each of these strands to lexical development is apparent if we look at how they contribute to vocabulary learning.

Meaning-focused input

In recent years, many researchers have advocated learning with meaning-focused input as a means to increase L2 input and fuel lexical growth. The justification for this approach is that individual words (e.g. Horst, Cobb, & Meara, 1998; Waring & Takaki, 2003) and sequences of words (Pellicer-Sánchez, 2017; Webb, Newton, & Chang, 2013) can be learned incidentally through reading. The greater the amount of reading that is done, the greater potential to learn words because more input provides the opportunity for words to reoccur. However, Cobb (2007) found that even if learners did relatively large amounts of extensive reading, they might still struggle to master the most frequent 3,000 word families. This has led to recommendations that extensive
reading be supplemented with other forms of meaning-focused input such as L2 television programs to increase the amount of input (Webb, 2015; Webb & Nation, 2017).

There are several reasons why different types of meaning-focused input should be used for learning. First and most importantly, not all people like to read extensively in their L1 so expecting extensive reading to fuel L2 vocabulary growth may not be realistic. Second, research suggests that people may encounter more L1 input through sources other than written text. For example, Americans view television more than five times as much as they read (United States Department of Labor, 2006). Many students may thus be motivated to learn from alternative sources of input, and in fact research does support this (Grieve & Clark, 2005; Sueyoshi & Hardison, 2005). Third, research suggests that L2 words are learned incidentally through viewing short videos (e.g. Neuman & Koskinen, 1992; Rice & Woodsmall, 1988), online internet video clips (Montero, Peters, Clarebout, & Desmet, 2014), and television programs (Rodgers, 2013), and that the amount of learning may be similar to that gained through reading (Neuman & Koskinen, 1992; Rodgers, 2013).

It is useful to consider the contributions of meaning-focused input in relation to each of the four components of vocabulary knowledge that might be indicators of an advanced level of lexical proficiency. Knowledge of the form-meaning connection of unknown and partially known words is gained through repeated encounters in meaning-focused input. Of particular importance to learning is the amount of input; a small amount of input leads to relatively few words encountered a sufficient number of times for learning to occur. However, as the amount of input increases, the number of words that are encountered with sufficient repetitions for learning to occur increases exponentially (Webb & Nation, 2017).

Knowledge of formulaic language is also likely to be gained through learning with meaning-focused input in the same way as individual words, through repeated encounters in context (Webb et al., 2013). Readers or listeners are likely to gradually become aware of the syntagmatic relationships of words as they encounter them again and again in context. This might be particularly important in developing productive knowledge of words; it can be difficult to use words when you do not know which words they typically co-occur with. Improvement in lexical sophistication is likely to occur in part through these gains in knowledge of form-meaning connection and formulaic language. As knowledge of form-meaning connection increases, there is greater potential to use more words. Similarly, as knowledge of formulaic language occurs, learners may have greater awareness of how to use more words. Learning with meaning-focused input may also improve processing speed. Beglar, Hunt, & Kite (2012) found that through reading large amounts of simplified text, the number of words read per minute increased.

Meaning-focused output

There are several reasons why meaning-focused output is important in developing vocabulary knowledge. First, research suggests that it is more difficult to gain productive knowledge than receptive knowledge (Griffin & Harley, 1996;
Webb, 2009a, 2009b), and that we tend to have less productive knowledge than receptive knowledge of words (Webb, 2008). This means that while we may understand many words when we encounter them, we may be unable to produce these words when needed. This stands to reason because when we use words we need to pay attention to aspects of vocabulary knowledge such as grammatical functions, collocations, and constraints on use that are of lesser importance for comprehension. Second, research comparing receptive and productive learning indicates that productive knowledge is more effectively gained through productive learning rather than receptive learning (Griffin & Harley, 1996; Webb, 2009a, 2009b). This reveals the need to include output in learning to improve the potential for students to use words. Moreover, Joe (1998) found that how words were used had an effect on learning; the greater the variation in use, the better the chance that a word would be learned. Together, this suggests that for learners to develop the ability to effectively use words, meaning-focused output will likely need to play a major role.

The challenge of learning through meaning-focused output in the classroom is that it tends to be more time consuming than learning in the other three strands, and there is no guarantee that students will use target words; some learners may avoid using partially known words to reduce the chances that they make mistakes. Manufacturing opportunities for students to use the vocabulary that they learn in meaning-focused output outside of the classroom is likely required to further develop the ability to use target vocabulary. There are many types of meaning-focused output available for autonomous learning, such as writing L2 blogs, commenting on online forums, tweeting, making spoken journals, participating in speaking corners, and singing songs. Familiarizing students with these types of activities within the classroom may be necessary to ensure that they are encouraged to learn in this way outside of the classroom. Including meaning-focused output in learning will help students to move beyond knowledge of form-meaning connection and toward a richer knowledge of vocabulary that includes improvement in knowledge of formulaic language and using words in speech and writing at a faster rate.

**Language-focused learning**

Language-focused learning is the strand that is most typically associated with vocabulary learning because it involves the deliberate learning of words. Research suggests that deliberate learning of vocabulary is a faster and more efficient way of learning words than incidental learning of vocabulary through reading (Laufer, 2003). In particular, paired associate learning has been found to be a particularly effective method of facilitating lexical development through the learning of the form-meaning connections of words (e.g. Thorndike, 1908; W. B. Webb, 1962). For example, W. B Webb (1962) found that one person was able to learn 666 word pairs in six hours. Although decontextualized approaches to learning words such as paired associate learning are focused solely on linking L2 form to meaning, there is evidence that when there is overlap in L1-L2 knowledge of other components of vocabulary knowledge, such as collocation and association, that other aspects of
lexical knowledge may also be gained (Webb, 2007). However, it is important to note that deliberate learning of a word may often focus on developing knowledge of a single meaning sense, a single collocation, and a single form of a word. While this knowledge may be gained relatively quickly, further learning in meaning-focused input is likely necessary to expand on this knowledge. Thus, while language-focused learning has great value in vocabulary learning, it is important that it be supplemented with learning in the other three strands to more effectively improve other aspects of lexical knowledge, such as formulaic language and processing speed.

**Fluency development**

Perhaps the strand that receives the least attention in the classroom is fluency development (Nation, 2013). It may also be the strand that has the greatest relevance in achieving an advanced level of lexical development because fluency with words is closely aligned with proficiency. Fluency development should be a focus of learning with each of the four skills. Activities designed to improve lexical fluency tend to have four key characteristics: time pressure, repetition of the activity, the use of few if any unknown words, and being meaning-focused (Nation, 2013). Improving reading fluency involves increasing reading speed to a words-per-minute rate that is similar to that of native speakers. Speed reading courses have consistently been found to have a positive impact on increasing reading rate (Macalister, 2010). Writing fluency involves writing at a faster rate. The 10-minute writing activity in which students are encouraged to keep track of how many words they write over 10 minutes, and work to write more each session, is an example of a task designed to improve writing fluency. The 4/3/2 activity is designed to improve the number of words spoken per minute. Students talk to different partners about the same topic first over four minutes, then over three minutes, and finally over 2 minutes. The pressure to speak at a faster rate over subsequent iterations focuses students on talking at a faster pace. Listening activities tend to use repetition to try to improve understanding of authentic speech. Repeated listening, repeated reading while listening, and repeated viewing (of television) all involve hearing the same content several times. The rationale behind this type of activity is that in the first encounter with the language, listeners may be overwhelmed with trying to understand the overall meaning, as well as smaller details, sequences of words, and individual words. However, through subsequent encounters hearing the content, listeners are able to focus more attention on the aspects of the audio that they did not understand, and this results in better comprehension, as well as increased vocabulary learning (Webb & Chang, 2012). Forms of repeated listening may also help listeners to process the language in chunks rather than as individual words (Brown, Waring, & Donkaewbua, 2008; Webb & Chang, 2012), which may help to improve knowledge of formulaic language.

Lexical development through a balanced learning program that includes learning in each of the four strands provides a principled approach to reaching an advanced level of lexical development. It is also essential that students understand
that with so many words to learn and so much to learn about words, a balanced language learning program needs to occur both inside the classroom and outside of the classroom to optimize progress.

**Where should the focus of learning be for those who have already achieved an advanced level of lexical development?**

This first section in this chapter described how vocabulary knowledge is a multi-dimensional construct that involves developing different types of word knowledge such as vocabulary size, lexical sophistication, formulaic language, and increased processing speed. Further increasing knowledge of these components of vocabulary knowledge should be the primary aim of advanced learners.

The four strands should still guide learning at an advanced level of proficiency. Meaning-focused input provides the opportunity to continue to encounter and develop knowledge of unknown words and partially known words. Encountering these items in context reveals how the words are used, which should make it easier to use vocabulary in meaning-focused output. At an advanced level, students should be able to understand most discourse types, such as novels and newspapers, conversation, radio programs, television, and movies, without the support of dictionaries and teachers. Encountering large amounts of meaning-focused input should play a large role in fueling further lexical development.

Finding opportunities to use words in meaning-focused spoken and written output may be more challenging if learners are not living in an English-medium environment. However, meaning-focused output will likely have great impact on moving students from an advanced level to a near-native level of lexical proficiency. The value of meaning-focused output for advanced learners is that it will help to further develop productive knowledge of form-meaning connection, lexical sophistication, and formulaic language, as well as the rate of using words in speech and writing.

Language-focused learning should continue to play an important role in lexical development. In particular, a focus on increasing knowledge of formulaic language will remain important. Perhaps the most useful tool for this purpose is a concordancer. A concordancer is software that generates a collection of examples for a keyword. Concordancers list all of the occurrences of a word in a text or corpus according to preset parameters. For example, using the concordancer at the Compleat Lexical Tutor, a search for the keyword *regression* in a six-million-word academic corpus that is sorted alphabetically according to the first word on the left shows that *a, the, linear, multiple, of,* and *poisson* are all words that might be used directly before *regression* in academic text. If we sort according to the first word on the right we find that *to, analysis, the, and, equation, into, line,* and *model* are words that might be used immediately after *regression* in academic writing. The concordance (the set of examples for the keyword) thus reveals to users how a word can be successfully used.
A focus on fluency development might also be challenging when learning autonomously. If learners are aware of the four principles that set up ideal conditions for fluency development (time pressure, repetition of the activity, the use of few if any unknown words, and a focus on meaning rather than form), they can work toward improving lexical processing speed.

**Directions for further research**

Advanced lexical development is an under-researched area and so there are many topics that would be useful to investigate. In particular, there are two lines of research that require attention: vocabulary growth and test development. Studies of vocabulary growth are surprisingly rare. The following questions need to be examined in research:

- How many words can English as a second language (ESL) and EFL students be expected to learn per year?
- On average, how many years are required to learn the most frequent 1,000, 2,000, 3,000, 5,000, and 8,000 word families?
- What proportion of English language learners in the ESL and EFL contexts reach intermediate and advanced lexical proficiency targets?
- How many years of English language learning are required to reach an advanced level of lexical proficiency?

It would be useful to conduct longitudinal studies of L2 lexical growth that look at the number of words that are learned per year by students in different contexts, as well as the time that it takes students to achieve different vocabulary learning targets such as learning the most frequent 8,000 word families. It would also be important to examine the proportion of learners who reach the different vocabulary learning targets over different periods of time. This would help to reveal the efficacy of language learning programs and may indicate where emphasis needs to be placed in vocabulary learning. Conducting studies of lexical growth should involve measuring vocabulary knowledge at regular intervals (perhaps once per year) over the duration of learning. With an aim of helping learners reach an advanced level of lexical proficiency, there would be a need to develop several new instruments to gauge lexical development.

Although there are tests that are designed to measure receptive vocabulary knowledge of higher frequency words (Vocabulary Levels Test: Nation, 1983; Schmitt, Schmitt, & Clapham, 2001; Webb, Sasao, & Balance, 2017), vocabulary size (e.g. the Vocabulary Size Test: Coxhead, Nation, & Sim, 2015; Nation & Beglar, 2007; V_YesNo: Meara & Miralpeix, 2017), and lexical sophistication (Lexical Frequency Profile: Laufer & Nation, 1995; V_Size: Meara & Miralpeix, 2017), there are few established tests of productive vocabulary knowledge. The Productive Levels Test (Laufer & Nation, 1999) is perhaps the best-known productive test. However, it is based on rather old word lists, does not include all
frequency levels, and may not include a sufficient number of items per level to
provide a reliable measurement. Meara and Miralpeix’s (2017) V_Size test and
V_Capture and Laufer and Levitzky-Aviad’s (2016) CATSS (Computer Adaptive
Test of Size & Strength) are more recently designed exploratory tests that may
fill this gap. Establishing a valid and reliable measure of productive vocabulary
knowledge is particularly useful for learners at an advanced level of lexical
development because these learners may require regular use of words at this
stage of proficiency.

Although there have been several studies that have explored different for-
mats for measuring knowledge of formulaic language (e.g. Eyckmans, 2009;
Gyllstad, 2009; Jaén, 2007; Keshavarz & Salimi, 2007; Nguyen & Webb, 2016), at
present there are no established tests designed to measure receptive or produc-
tive knowledge of sequences of words. A diagnostic test designed to measure
productive knowledge of formulaic language at different word frequency levels
would be particularly useful because this would indicate (i) to what extent a
learner could correctly use word combinations, (ii) how far a learner’s knowledge
of formulaic language had progressed, and (iii) which frequency level they
needed to focus on when learning multi-word combinations. Nguyen and Webb
(2016) found that receptive knowledge of collocation followed a similar pattern
as that of single-word items; learners tend to know collocations according to the
frequency of the individual words in the collocations. It would also be useful to
explore whether productive knowledge of sequences of words followed this
same learning pattern.

Currently there is one test designed to measure lexical processing speed. Q_Lex
(Meara & Miralpeix, 2017) tests whether learners can identify five-letter words
within a longer string of letters such as \textit{angel} in \textit{todangeloant}. Although Q_Lex is still
in the early stages of development and only measures words relatively high in fre-
quency, it provides a useful starting point for formats aimed at measuring lexical
processing. Establishing tests that measure lexical processing speed in each of the
four skills would have great value for learners at all levels. It could specifically
help to show the degree to which learners are fluent with words, and this might
help to reveal whether further emphasis should be placed on developing fluency
in language learning programs.

Conclusion

In this chapter, it has been suggested that advanced lexical development be
examined from several different perspectives, because there are many words to
learn and many things that we need to learn about each word. For example, we
need to be able to produce the word to convey its intended meaning. Similarly,
we must learn to recognize or recall the meaning when we see or hear the word.
Eventually, there is a need to produce the word together with other words with
which it commonly occurs, and do so at a rate that is similar to that of native
speakers. Because there is much to be learned about words, there are many
Research indicates that the majority of learners in EFL contexts fail to reach an advanced level of lexical proficiency. Increasing vocabulary learning likely begins with a principled plan within institutions that aims at both learning more words and learning more about words. Nation’s (2007) four strands provide a useful framework that might be incorporated within classrooms to develop both breadth and depth of vocabulary knowledge. However, because classroom time is limited, it is important that both teachers and students recognize that the bulk of learning needs to occur autonomously outside of the classroom. Creating a long-term program within institutions that involves principled vocabulary learning inside and outside of the classroom is likely required to help more students reach and move beyond an advanced level of lexical development.

A balanced program that incorporates the four strands may be the best way for advanced learners to continue vocabulary learning. Although learners should continue to learn words as they progress, improving lexical accuracy and fluency may provide the greatest value for more advanced learners. A focus on improving knowledge of formulaic language should increase accuracy, and using words at a more native-like rate is the key to improving lexical fluency.

As we have seen, there is a need for more research on lexical development. In particular, further development of instruments designed to measure vocabulary knowledge in different ways is essential to investigations of advanced lexical development. The creation of instruments that provide a valid and reliable measure of receptive and productive knowledge of formulaic language, productive knowledge of form-meaning connection, and lexical processing would allow researchers to provide a more accurate assessment of vocabulary growth.

<table>
<thead>
<tr>
<th>Aspect of lexical knowledge</th>
<th>Abilities of advanced learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary size</td>
<td>Comprehend English by knowing 98% or more of the words in spoken and written discourse</td>
</tr>
<tr>
<td>Lexical sophistication</td>
<td>Use both high and low frequency words to communicate in a precise and effective manner</td>
</tr>
<tr>
<td>Formulaic language</td>
<td>Use words in the sequences in which they frequently co-occur in English</td>
</tr>
<tr>
<td>Lexical processing</td>
<td>Recognize and produce words at native-like speeds</td>
</tr>
</tbody>
</table>
REFERENCES


Introduction: Information structure (IS) at the syntax–discourse interface

A well attested phenomenon in advanced and near-native L2 acquisition is learners’ difficulty in the acquisition of (morpho-)syntactic alternations when constrained by principles of information structure (IS), e.g. topic and focus, at the interface of syntax and discourse. Interface relations, opaque form-meaning mappings, and non-obligatory discourse-motivated preferences are typically areas of difficulty in second language acquisition (SLA), which may lead to divergence even at the end-states of acquisition (DeKeyser, 2005; Sorace, 2011). IS management is problematic even for advanced L2 learners who have only a limited awareness of the appropriate use of lexical and (morpho-)syntactic devices of information highlighting in formal and informal registers (see Callies, 2009, and Lozano, 2014, for overviews). IS is often considered the “final threshold” or “hurdle” (von Stutterheim, 2003; Verheijen, Los, & de Haan, 2013) and is among the “late-acquired features” (Ortega & Byrnes, 2008).

The investigation of interlanguage has progressed from the sentence level to the discourse level over the past few years. A series of constructions are constrained by IS principles like topic and focus, e.g., non-canonical word order (postverbal subjects, dislocation, clefting) and the distribution of pronominal
forms in discourse. We will pay attention to these constructions within two broad research traditions in SLA:

- Formal approaches, which have focused on the syntax–discourse interface. Most of the evidence comes from experimental studies.
- Functional approaches, which have investigated IS at large as well as specific focus constructions and alternations. Most of the empirical evidence comes from learner corpus studies.

Researchers working in one of the two paradigms have often overlooked studies conducted in the other. We will bring these two approaches together since the overarching question they ultimately try to answer is one of the key questions identified in SLA research: “Can L2 learners achieve native-like competence?” (VanPatten & Benati, 2010, p. 9). The learning task is not only whether learners can acquire formal grammatical properties of specific constructions, but also whether they can acquire their associated discourse function(s) in view of the constraints represented by IS concepts such as information status and topic/focus.

We will review the major findings on the acquisition of word order and the key principles of IS by advanced learners by considering two major methodological approaches (experimental and corpus methods), and a range of L2s (Spanish, English, German, etc.). We survey relevant studies focusing on advanced and near-native adult L2 learners, leaving aside other acquisition contexts such as child L2 acquisition and heritage speakers.

The chapter is structured as follows: In the second section, we present studies testing the Interface Hypothesis with a focus on (morpho-)syntactic alternations at the syntax–discourse interface (pronominal subject distribution, subject–verb inversion, and left-periphery structures). In the third section, we discuss functional approaches to IS that have examined transfer of principles of IS, cleft constructions, alternations, and aspects of grammatical variation. In the fourth section, we discuss how the combination of corpus and experimental data represent a way forward in the study of IS in advanced and near-native L2 acquisition.

The syntax–discourse interface and the Interface Hypothesis in advanced SLA

The Interface Hypothesis (IH; Sorace, 2011, 2012; Sorace & Serratrice, 2009) has been the theoretical framework for a series of studies testing IS: overt/null pronominal–subject alternations, subject–verb inversion, and left-periphery structures.

The Interface Hypothesis

Sorace’s (2011, 2012) IH1 assumes a modular approach to the language faculty. The computational system interfaces with the mental lexicon and also with other modules such as the sensory-motor systems (phonetics/prosody) and the conceptual-intentional systems (discourse/pragmatics/semantics), see Figure 22.1.
Interface relations have been widely reported to be problematic in advanced L2 acquisition. Sorace and Serratrice’s (2009) IH predicts that it is only the external interfaces (e.g. syntax–discourse), and not the internal interface (lexicon–syntax), that are persistently problematic and lead to residual optionality and divergence from native grammars in near-native learners. In the latest IH version (Sorace, 2011, 2012), learners’ deficits may not be observed in off-line tasks testing knowledge, but will be observable in on-line tasks due to the difficulty of simultaneously integrating purely linguistic with discourse information during on-line processing/parsing, which requires more processing resources. Deficits could be the result of a trade-off between inhibitory control and integration/updating (Sorace, 2016). However, research has shown that many properties at the syntax–discourse interface are eventually acquirable/processable, which then leaves an open question: Why are some interface properties more vulnerable than others? (see White, 2009, 2011, for discussion). Note that recent SLA theories (Slabakova’s, 2014, Bottleneck Hypothesis) agree that learners can acquire syntax, but argue instead that their main challenge is inflectional morphology and its associated formal features. In what follows, we will review the evidence provided by selected studies that tested the IH.

**Pronominal subject distribution and IS**

The distribution of overt and null pronominal subjects in discourse has been extensively tested under the IH. In null-subject languages (e.g. Spanish, Italian, Greek, Arabic) the overt/null pronominal subject distribution is grammatically allowed but pragmatically constrained at the syntax–discourse interface. Overt and null pronouns (Spanish él/Ø ‘he/Ø’) can syntactically alternate in subject position as in (1a). However, when there is a continuity in topic with either one or two potential antecedents, as in (1b) and (1c), respectively, a null pronoun felicitously
marks topic continuity (Juan, ..., Ø), whereas an overt pronoun would be pragmatically infelicitous, as shown by ‘#’, because it signals a shift in topic: él typically refers to the object (Pedro). The learnability question is whether advanced/near-native learners of L2 Spanish can acquire not only the syntactic alternation, but also its subtle IS nuances (topic-continuity vs. topic-shift).

(1) a. Él/Ø abría la puerta.
   ‘He/Ø opened the door’.

b. Juan i tropezó mientras Ø i/é abría la puerta.
   ‘Juan tripped over while he/Ø opened the door’

c. Juan saludó a Pedro mientras Ø i/# j/é abría la puerta.
   ‘Juan greeted Pedro while Ø/he opened the door’

Most L2 studies draw on (very) advanced L1 English-L2 Spanish experimental data. The overarching finding is that even advanced and near-native learners have often residual deficits as they occasionally tolerate a pragmatically redundant overt pronoun in topic-continuity contexts where native speakers would prefer a null pronoun (see Quesada, 2014, 2015, for overviews). This has been also reported for advanced L1 Greek-L2 English (Prentza & Tsimpili, 2013) and near-native L1 English-L2 Italian (Sorace & Filiaci, 2006). However, highly advanced learners can attain statistically native-like patterns not only in topic-continuity but also in topic-shift scenarios in several language pairings: L1 English-L2 Spanish (Jegerski, VanPatten, & Keating, 2011; Rothman, 2009), L1 Arabic-L2 Spanish (Bel & García-Alcaraz, 2015), and L1 Greek-L2 Spanish (Lozano, 2002, forthcoming). Some researchers agree that topic-continuity scenarios are indeed persistently problematic, but advanced learners show native-like patterns, e.g. L1 Greek and L1 Farsi learners of L2 Spanish (Judy, 2015; Lozano, forthcoming). These findings are in line with White’s (2011) proposal that the syntax–discourse interface is not monolithic, since some contexts can be more problematic than others. To summarize, results from experimental studies are mixed with respect to native-like attainment.

The above studies are experimental and tested the comprehension/interpretation of pronominals in pragmatically limited (and often unnatural) contexts. A recent series of L2 Spanish corpus studies has revealed that contexts are pragmatically richer in language production. Montrul and Rodríguez-Louro (2006) used an oral retelling task with L1 English-L2 Spanish learners at three proficiency levels (intermediate, advanced, near-native). In topic-continuity contexts, the advanced (7.6%) and near-native learners (0.3%) still produced some residual redundant overt pronouns, in contrast to Spanish native speakers (0%), as previously reported in experimental studies. In topic-shift contexts, they used some null pronouns ambiguously (8.4% advanced, 5.5% near-native speakers vs. 1% native speakers). Similar results are reported by Lozano (2009), who tested lower- and upper-advanced learners in a written film-retelling task from the L1 English-L2 Spanish
subcorpus of CEDEL2 (Corpus Escrito del Español L2; Lozano & Mendikoetxea, 2013). Learners showed deficits only with third-person human contexts, whereas the rest of the pronominal paradigm was intact. Importantly, topic shift was typically realized by a full noun phrase (79% lower-advanced learners, 87% upper-advanced learners, 92% native speakers), and not by an overt pronoun (11% and 13% for learners, 4% for native speakers), contrary to previous experimental research. Production of noun phrases (NP) in topic-shift contexts was also observed in an oral film-retelling task in advanced L1 English-L2 Spanish by Blackwell and Quesada (2012). NP production cannot therefore be a task effect (oral vs. written production). The use of such a prolix form (NP) was subsequently explored by Lozano (2016) in very advanced L1 English-L2 Spanish learners in the CEDEL2 corpus. Results revealed that the nature of the antecedent(s) is a key factor. First, learners’ and native speakers’ use of NP in topic-shift contexts is triggered by the number of potential (i.e. activated) antecedents: when two antecedents are present in prior discourse, an overt pronoun is produced more often than an NP (53% > 29% native speakers, 41% > 36% learners), but when there are three or more potential antecedents, an NP is more often produced than an overt pronoun (97% > 2% natives, 94% > 13% learners). Second, the widely reported overuse of overt pronouns in topic-continuity contexts is also accounted for by the number of antecedents: learners produce more overt pronouns when there are two potential antecedents (67%) than when there is only one (18%). Finally, the corpus data revealed that learners prefer being redundant (i.e. overproducing overt pronouns in topic-shift contexts) than ambiguous (i.e. using null pronouns in topic-shift contexts). This is explained in terms of a new pragmatic proposal (the Pragmatic Principles Violation Hypothesis), which could account for many of the results reported above.

**Subject–verb inversion and IS**

Word order has been also studied within the framework of the IH (see Lozano, 2014, for overviews). One of the challenging learnability tasks in L2 Spanish is the acquisition of subject–verb (SV) and verb–subject (VS) alternations with intransitive verbs: Un niño lloró/Lloró un niño (‘A boy cried/Cried a boy’), Un niño vino/Vino un niño (‘A boy arrived/Arrived a boy’). They are subtly constrained at the interfaces (cf. Table 22.1). Given an adequate preceding context, a question asking about the subject of the sentence (¿Quién gritó/vino? ‘Who shouted/arrived?’) requires the subject to appear postverbally (VS) in the response since focus (i.e. discourse-new) elements are placed sentence–finally in Spanish (e.g. Zubizarreta, 1998). Thus, word order is constrained at the syntax–discourse interface. However, a global, out-of-the-blue question like ¿Qué pasó? ‘What happened?’ triggers an informationally neutral (all-focus) response since no specific part of the sentence is in focus. The neutral word order is canonical SVO with transitives and SV with a subtype of intransitives called unergatives (llorar ‘cry’, gritar ‘shout’, estornudar ‘sneeze’, etc), which typically assign the thematic role of agent to their only argument (the subject). By contrast, with a subtype of intransitives called
unaccusatives (venir ‘to arrive’, llegar ‘to arrive’, salir ‘to leave’, entrar ‘to enter’), the resulting neutral word order is VS because, as widely argued in the literature, these verbs assign the theme/patient role to their argument, and the subject can therefore remain in object (i.e. postverbal) position (Domínguez & Arche, 2014; Lozano, 2006a, for theoretical explanations). In short, SV/VS alternations in Spanish are constrained lexically by the type of predicate (lexicon–syntax interface) and also by IS (syntax–discourse interface). In the remainder of this section we will report only the syntax–discourse results of previous studies.

Hertel (2003) used a contextualized production task (cf. material in (2)) with L1 English-L2 Spanish learners. With focused-subject questions, Spanish native speakers produced VS order slightly more with unaccusatives (36.4%) than with unergatives (32.9%). Advanced learners’ discrimination (53.6% vs. 36.3%, respectively) was not significantly different from natives’ (Hertel, 2003, p. 294), unlike beginner and intermediate learner groups, who mostly produced SV order as a reflection of their L1. Hertel concluded that advanced learners can acquire the discourse-related word order in L2 Spanish.

(2) [Context, in Spanish] You are at a party with your friend Laura. After Laura has left the room, the police arrive because the party is too noisy. When Laura comes back, she asks you: ‘Who arrived?’ You answer:

[Hertel’s (2003) production task] ..........................................
[Lozano’s (2006a, 2006b) acceptability task] a. La policía llegó. -2 -1 0 +1 +2
b. Llegó la policía. -2 -1 0 +1 +2
b. Llegó la policía.
c. Both sentences

Lozano (2006a) tested SV/VS order in highly advanced L2 Spanish learners with L1 Greek and L1 English. In focused-subject contexts the expected order is SV in Greek and English, but VS in Spanish, independent of verb type (unaccusative/unergative). Lozano used a contextualized acceptability judgment task, also reproduced in (2), where the two possible orders (SV and VS) are provided as a reply, each followed by a Likert scale (from -2 = ‘unnatural’ to +2 = ‘natural’). Both groups

<table>
<thead>
<tr>
<th>Lexicon–syntax interface:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global question (What happened?)</td>
</tr>
<tr>
<td>SV: Un niño lloró</td>
</tr>
<tr>
<td>VS: Vino un niño</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntax–discourse interface:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused-subject question (Who cried/arrived?)</td>
</tr>
<tr>
<td>VS: Lloró un niño</td>
</tr>
<tr>
<td>VS: Vino un niño</td>
</tr>
</tbody>
</table>
(see Figure 22.2) showed optionality by simultaneously accepting SV and VS, whereas the native speakers clearly accepted VS but rejected SV with both unergatives and unaccusatives. Similar results are reported by Lozano (2006b) for very advanced learners in a developmental study of L1 Greek-L2 Spanish. However, note that in both studies (Lozano, 2006a, 2006b) very advanced learners can attain native-like competence in global contexts at the lexicon–syntax interface.

In another developmental study, Domínguez and Arche (2014), tested L1 English-L2 Spanish learners at different proficiency levels in a contextualized preference task, cf. (2). Beginner and intermediate learners always preferred SV, in line with Hertel (2003). The advanced learners (Figure 22.2) optionally
preferred SV (52%) and VS (48%) in unergative subject-focused contexts, which is in line with Lozano (2006a, 2006b), but Domínguez and Arche argue that such optionality is a reflection of the L1 Spanish input (cf. their Spanish native speakers’ preference rate: 49%; VS: 51%). By contrast, their advanced learners’ preference for VS (56%) to SV (44%) with unaccusatives was significant (though marginally), in line with natives’ highly significant preference (75% vs. 25%). However, between-group analyses revealed that “the advanced group may not be considered completely native-like even though they prefer the target option and show no optionality” (Domínguez & Arche, 2014, p. 259). This discrepancy in some of the results between Hertel (2003), Lozano (2006a, 2006b), and Domínguez and Arche (2014) probably stems from the differences in the elicitation task used, as evidenced in (2).

Subject–verb inversion (i.e. VS order) has received some attention in the L2 English literature: learners with different L1 backgrounds (Japanese, Arabic, Italian, Spanish, Greek) produce postverbal subjects (shown in italics in (3)).

(3) a) Sometimes comes a good regular wave. (L1 Japanese-L2 English) (Zobl, 1989, p. 204).
   b) … on her face appeared those two red cheeks. (L1 Arabic-L2 English) (Rutherford, 1989, p. 179).
   c) It is almost disappearing the use of writing nice letters to friends. (L1 Italian-L2 English) (Lozano & Mendikoetxea, 2008, p. 106).
   d) In the name of religion it had occurred many important events. (L1 Spanish-L2 English) (Lozano & Mendikoetxea, 2010, p. 486).
   e) There exist about two hundred organizations such as Greenpeace, which have increased the number of its members laterly. (L1 Spanish-L2 English) (Lozano & Mendikoetxea, 2008, p. 487).
   f) … because exist the science technology and the industrialization. (L1 Spanish-L2 English) (Lozano & Mendikoetxea, 2010, p. 487).
   g) It could exist serious problems in everyday life. (L1 Greek-L2 English) (Agathopoulou, 2014, p. 178).

Interestingly, unlike the L2 Spanish experimental studies above, these L2 English studies used contextualized (corpus) data. Rutherford (1989), Zobl (1989), and Oshita (2004) report the production of VS in L2 English by native speakers of different L1s (Japanese, Arabic, Spanish), but their findings say little about advanced learners’ behaviour as their proficiency level was not controlled for. Lozano and Mendikoetxea (2008, 2010) tested advanced learners of English with Italian and Spanish as L1s. Out of all sentences with unaccusative verbs, VS production rates were 3% in the Italian subcorpus (Lozano & Mendikoetxea, 2008), but were higher in the Spanish subcorpus: 8% (Lozano & Mendikoetxea, 2008) and 7% (Lozano & Mendikoetxea, 2010). The Spanish subcorpora rates were relatively high when compared to English native speakers (around 2%), though the learners behaved in a native-like fashion as they were aware of the interface conditions that constrain VS order: (i) lexicon–syntax: the verb must be unaccusative for inversion
to occur, since VS is never produced with unergatives; (ii) syntax–discourse: the subject appears postverbally when it is discourse-new information (focus), but preverbally when it discourse-old information (topic); (iii) syntax–phonology: the subject tends to be heavy when it appears in postverbal position, but light when in preverbal position. Agathopoulou (2014) replicated Lozano and Mendikoetxea’s (2008, 2010) findings in a corpus study of advanced L1 Greek-L2 English learners whose inversions were also constrained at the three interfaces.

Interestingly, in the corpus studies learners often produced certain phrases (XPs for short) in preverbal position (XP-V-S order), as evidenced by (3). XPs were tested in a follow-up experimental study with learners of L2 English at several proficiency levels (A1, A2, B1, B2, C1, C2) in the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001) and a native-speaker control group (total N = 417) (Mendikoetxea & Lozano, submitted 2017). Four main types of XPs were studied: two structurally possible constructions in English (PP locative inversion as in (3b) and there-insertion as in (3e)) versus two structurally impossible constructions (*it-insertion (3c,d,g) and *Ø-insertion (3f)). Results confirmed that (i) VS was significantly preferred with unaccusatives rather than with unergatives at each proficiency level and (ii) *it-V-S and *Ø-V-S were the least accepted constructions at the advanced levels (i.e. C1 and C2), as in native English, though in the corpus study advanced learners overproduced *it-V-S constructions (see Table 22.2). Learners can eventually attain native-like competence of the discourse constraints that regulate word order in English but show some residual optionality in producing ungrammatical preverbal expletives (*it/*Ø), which is argued to be a purely grammatical problem. These results would be in line with Slabakova’s (2014) Bottleneck Hypothesis, since the deficits seem to be morpho-logical (and not discursive) in nature.

To summarize, the experimental studies of L2 Spanish revealed that advanced learners show residual deficits with SV/VS at the syntax–discourse interface, but the corpus studies of L2 English indicate that learners are aware of the syntax–discourse properties of VS, despite some problems with the grammatical encoding of the preverbal XP.

**The left periphery and IS**

Several non-canonical word order phenomena appear in the left periphery of the sentence. The learnability question is whether advanced/near-native learners can eventually acquire not only the non-canonical order, but also its associated topic/focus interpretation.
In many languages, constituents that represent discourse-old (topic) information can be fronted. In languages with clitics (unstressed object pronouns), a clitic (CL) is required after the dislocate, typically after a brief pause (shown by commas in written language), as in (4B) in Spanish. The CL (in bold) doubles the dislocate (in italics), i.e. CL is co-referential with and agrees with the dislocate in gender, person, and number. This is known as clitic left dislocation (CLLD; or just as ‘left dislocation’ in functional approaches). Other Romance languages (e.g. French, Italian) and some Slavic languages (Bulgarian and Macedonian) can mark discourse-old information via CLLD as well (see Bulgarian and Spanish in (5) compared to French and Spanish in (6) (cf. Slabakova, 2015; Slabakov, Kempchinsky, & Rothman, 2012, and references therein for theoretical discussions).

(4) A: ¿Qué hiciste con el libro? / ¿Qué hiciste con la libreta?
‘A: What did you do with the book? / What did you do with the notebook?’
B: El libro, lo tiré a la papelera. / La libreta, la tiré a la papelera.
‘B: The book, I threw it in the bin. / The notebook, I threw it in the bin.’

(5) A: Did you write to Ivan?
B: A Iván, le escribí tres veces. (Spanish)
B’: Na Ivan, mu pisax tri pâti. (Bulgarian; Ivanov, 2012, pp. 354–355)
To Ivan, him I-wrote three times

(6) A: Have you seen Marie?
B: A Marie, la veo a menudo. (Spanish)
To Marie, her I-see often
B’: Marie, je la vois souvent. (French; Donaldson, 2011, p. 403, ex. (2))
Marie, I her see often

In other cases, the dislocate surfaces in left-peripheral position but a clitic/resumptive pronoun is not allowed (unlike in CLLD), cf. (7). This structure is known as focus fronting (FF) in the generative literature (while the term topicalization is typically used in functional approaches). The dislocate bears emphatic stress (shown by capital letters) and has a clear contrastive reading in that it is set apart from another referent (It is the keys that I lost, not the wallet). These structures are formally and functionally similar in these languages.

(7) A: What did you lose yesterday, the keys or the wallet?
B: THE KEYS I lost. (English)
B’: LAS LLAVES perdí. (Spanish)
B”: KLJČOVETE zagubix. (Bulgarian, after Ivanov, 2012, p. 355)

Slabakova, Kempchinsky, and Rothman (2012) examined whether learners can eventually acquire not only the formal aspects of CLLD versus FF but also their
discursive interpretation in L1 English-L2 Spanish at several proficiency levels (intermediate, advanced, and near-native), plus an L1 Spanish control group. All learner groups showed knowledge of the formal syntactic aspects, as reported in previous research. The authors used a written contextualized judgment task accompanied by spoken audio material, where a preceding context biased toward either a CLLD or an FF structure, as in (8a) versus (8b). Learners were presented with both sentences simultaneously, and each had to be judged on a scale (4 = ‘perfect,’ 3, 2, 1 = ‘very strange,’ 0 = ‘don’t know’). Results showed that, while learners progressed toward the L1 norm, only the near-native learners attained native-like knowledge by preferring (i) clitic (topic interpretation) to cliticless (focus interpretation) structures in CLLD-biased contexts, but (ii) cliticless (focus interpretation) to clitic (topic interpretation) structures in FF-biased contexts. The authors conclude that the discursive interpretation of CLLD and FF can be eventually acquired in L2 Spanish, a fact also reported by Slabakova (2015).

(8) [Context] Juan and Mónica invited María for dinner. The dinner was served on the terrace and everything was delicious. María congratulated Juan on the soup he made. When Mónica heard this, she answered:
(a) La carne la preparó Juan, no la sopa. 4 3 2 1 0
(b) La carne preparó Juan, no la sopa. 4 3 2 1 0

The meat (it) prepared Juan, not the soup

Slabakova and Ivanov (2011) investigated CLLD and FF in advanced versus intermediate L1 English-L2 Bulgarian learners and an L1 Bulgarian control group. The task was a contextualized sentence acceptability task. Similar to the examples above, a long context was first introduced, followed by four sentences, each containing a different word order (CLLD, CLRD, FF, SVO). Each sentence was followed by a Likert scale (1 = ‘totally unacceptable’ to 5 = ‘totally acceptable’). There were two test conditions: (i) the topic condition, elicited by a yes-no question as in (5), which requires CLLD and not FF in order to be pragmatically felicitous; and (ii) the focus condition, elicited by a wh-question expressing contrast, as in (7), which requires FF and not CLLD for felicity. Results clearly show that advanced learners attained native-like knowledge by (i) giving high rates to CLLD but low rates to FF in the topic condition, and (ii) showing the opposite behavior in the FF condition, i.e. giving low rates to CLLD but high rates to FF. Ivanov concludes that “advanced learners of Bulgarian ... performed like the native controls ... , which can be interpreted as successful acquisition of the pragmatic meaning of clitic doubling in Bulgarian” (Ivanov, 2012, p. 364). The intermediate learners preferred the canonical SVO, which indicates that they were still unaware of the pragmatic interpretation of CLLD and FF in Bulgarian. Slabakova and Ivanov (2011) also report native-like attainment with CLLD in advanced in L1 English-L2 Bulgarian. All these results run against the IH since native-like attainment is possible at the syntax-discourse interface.
Importantly, most of the evidence presented in this section comes from comprehension/acceptability data. However, Donaldson (2011) used naturalistic oral corpus data to determine whether near-native L1 English-L2 French speakers could acquire the topicalization function of CLLD in French, cf. example (6B'). Results show that near-native learners’ production of CLLD in French is native-like, since CLLD is used by both groups as a topic-marking device in which the dislocate is discourse-old (i.e. topical) information. Donaldson also concludes that his findings force a rejection of Sorace’s IH since native-like attainment is possible.

To summarize, the studies reported in this subsection do not fully support the earlier versions of the IH, though they are silent about the latest version of the IH. More L2 empirical data from on-line processing tasks is needed to obtain a full picture of IS at the left periphery in L2 acquisition.

**Summary: Native-like attainment at the syntax–discourse interface?**

Table 22.3 summarizes the studies reviewed in this section. Overall, native-like attainment is possible in many scenarios at the syntax–discourse interface, though there are certain scenarios that are persistently problematic. Future research will need to address this asymmetry of behavior, as it is clear that not all syntax–discourse interface constructions are monolithic (White, 2011).

**Functional approaches to word order and IS in advanced SLA**

In this section we will review studies that have investigated L2 learners’ use of specific syntactic means of information highlighting located at the interface of lexicogrammar, syntax, and pragmatics. This interface is often referred to as IS or information packaging, i.e. the structuring of sentences by syntactic, prosodic, or morphological means that arises from the need to meet certain communicative demands, e.g. emphasizing a certain point, correcting a misunderstanding, or repairing a communicative breakdown. Compared to their frequency of occurrence and difficulty of acquisition there are still comparatively few studies that have examined the use of linguistic means of information highlighting in advanced interlanguages. L2 learners’ knowledge of discourse organization, including awareness, comprehension, and production, as well as the (contextual) use of linguistic means of information highlighting, is thus still an underexplored research field. The studies to be summarized in this section draw on, broadly speaking, functional linguistic approaches to the study of advanced interlanguage and have used diverse methodological approaches. However, given the importance of context for the study of word order and IS, many studies have made use of learner corpora which provide large amounts of data representing authentic, continuous, and contextualized language use. Given the developments
Table 22.3  Summary of native-like attainment at the syntax–discourse interface.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Study</th>
<th>L1-L2</th>
<th>Proficiency level</th>
<th>Native-like?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic continuity</td>
<td>Several</td>
<td>Several</td>
<td>Advanced and near-native</td>
<td>Yes*, No†</td>
</tr>
<tr>
<td>Topic shift</td>
<td>Several</td>
<td>Several</td>
<td>Advanced and near-native</td>
<td>Yes</td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>Judy (2015)</td>
<td>Farsi-Spanish</td>
<td>Near-native</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Lozano (forthcoming)</td>
<td>Greek-Spanish</td>
<td>Very advanced</td>
<td>Yes</td>
</tr>
<tr>
<td>End focus</td>
<td>Lozano (2006a)</td>
<td>English-Spanish</td>
<td>Very advanced</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Greek-Spanish</td>
<td>Very advanced</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Lozano (2006b)</td>
<td>Greek-Spanish</td>
<td>Very advanced</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Domínguez &amp; Arche (2014)</td>
<td>English-Spanish</td>
<td>Advanced</td>
<td>No</td>
</tr>
<tr>
<td>End focus</td>
<td></td>
<td>Spanish-English</td>
<td>Advanced</td>
<td>Yes</td>
</tr>
<tr>
<td>CLLD</td>
<td>Ivanov (2012)</td>
<td>English-Bulgarian</td>
<td>Advanced</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Slabakova &amp; Ivanov (2011)</td>
<td>English-Bulgarian</td>
<td>Advanced</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Slabakova et al. (2012)</td>
<td>English-Spanish</td>
<td>Near-native</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advance</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advance</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slabakova (2015)</td>
<td>English-Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td></td>
<td>Donaldson (2011)</td>
<td>English-French</td>
<td>Near-native</td>
<td>Yes</td>
</tr>
<tr>
<td>FF</td>
<td>Ivanov (2012)</td>
<td>English-Bulgarian</td>
<td>Advanced</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Slabakova et al. (2012)</td>
<td>English-Spanish</td>
<td>Near-native</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advance</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Slabakova (2015)</td>
<td>English-Spanish</td>
<td>Near-native</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advance</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Note:
* with Ø pronouns  
† with overt pronouns
in the field of corpus linguistics at large, most studies adopt usage-based models of acquisition though not all explicitly situate themselves in this paradigm.

**Transfer of principles of IS in the advanced stages of acquisition**

Numerous studies have demonstrated the relevance of L1 discourse structure for L2 acquisition by examining two typological parameters relating to discourse organization: topic- versus subject-prominence (Li & Thompson, 1976) and pragmatic versus grammatical word order (Thompson, 1978). Research shows two contradictory claims about the role of the topic-/subject-prominence continuum in SLA: (i) irrespective of the learners’ L1, SLA is characterized by an early universal topic-prominent stage, and topic-/subject-prominence is not transferable (e.g. Givón, 1979, 1984); (ii) the learners’ L1 does play a role in SLA in that learners transfer L1 function (IS) to L2 form, and as their proficiency increases, speakers of topic-prominent L1s gradually increase the use of subject-prominent features in their L2 (e.g. Han, 2000; Jung, 2004; Rutherford, 1983; Schachter & Rutherford, 1979). There has been increasing evidence for the second scenario. For instance, L2 English learners with topic-prominent L1s show transfer effects as to (i) overproduction of L2 structures that have no equivalents in the respective L1s (it-extrapolation, there-sentences, and pseudo-passives; see Han, 2000; Schachter & Rutherford, 1979), and (ii) a tendency to use topic-fronting devices and connectors in theme position to introduce new information (e.g. Green, Christopher, Lam, & Mei, 2000).

Several studies investigated the extent to which very advanced L2 learners still apply the principles of IS of their native language as contrasted with the L2. Methodologically, this has typically been done on the basis of open-ended tasks in which learners are made to produce stretches of connected and contextualized discourse such as telling a story, giving route directions, or describing a picture. This research found that core typological principles in IS are retained in advanced interlanguages (Carroll & Lambert, 2006; Carroll, Murcia-Serra, Watorek, & Bendiscioli, 2000; von Stutterheim, 2003). Evidence that transfer continues to occur at the level of discourse is provided by several further studies. Rankin (2012) found that German and Dutch L2 English speakers produced distinct patterns of inversion in declarative clauses, i.e. non-target subject–auxiliary inversion and copula inversion, indicating the transfer of verb-second (V2) syntax from their L1s. Since other reflexes of V2 in interrogatives or with sentential negation were not produced, Rankin concluded that only L1 preferences for topicalization structures continue to transfer at the advanced stages. Similar findings on Dutch L2 English learners’ preference for the initial theme zone are reported by Springer (2012) for topicalized clauses, and by van Vuuren and Laskin (2017) for sentence-initial adverbials. Bohnacker and Rosén (2008) studied IS of V2 declaratives in Swedish, German, and L2 German. Swedish learners of German were compared to native-speaker controls matched for age and genre. These learners were found to have mastered the formal syntactic properties of V2 but to overapply IS
preferences of their L1 Swedish which postpones new (rhematic) information and instead fills the prefield with given (thematic) elements and elements of no or low informational value. This was interpreted as indicating L1 transfer at the interface of syntax and IS, especially for structures that are frequent in the L1.

**Cleft constructions**

Clefts are the most widely studied single-focus construction. Callies’s (2009) study examined advanced German L2 English learners’ use of discourse-pragmatically motivated variations of the basic word order such as locative inversion, topicalization, and *it*- and *wh*-clefts. Triangulation of experimental and learner corpus data provided corroborating evidence. The findings show an over-representation of subject-prominent structures (*it*-clefts, existentials/presentationals, and *it*-extraposition) in the learner data, suggesting that subject-prominence is a factor that constrains productive learner output. Structures without a canonical sentence-initial subject (locative inversion, topicalization) were not used productively despite receiving good acceptability ratings. This is in accordance with markedness assumptions that, in relation to other focus constructions, clefts and other subject-prominent sentence types show a comparatively low degree of structural markedness because SVO/SVC word order is retained. Moreover, the findings confirm that L1 discourse structure continues to play a significant role in advanced L2 acquisition. While beginning and intermediate German learners frequently produce incorrect V2 sentences, this can hardly be observed at the advanced stages. Instead, learners gradually increase the use of subject-prominent features in their L2 production and move from L1 pragmatic word order to L2 grammatical word order.

Swedish L2 English learners were also reported to overuse cleft constructions, especially when thematizing new information and expressing personal opinions (Boström Aronsson, 2003; Herriman & Boström Aronsson, 2009). Clefts have also been studied in the L2 English of Dutch learners, but with inconclusive results. While Springer (2012) found an over-representation, the findings of Verheijen et al. (2013) suggest an under-representation of clefts. They argue that Dutch, a V2 language like German, provides separate positions for discourse linking and aboutness-topics and that advanced learners exploit the pre-subject adverbial position in English to perform the information-structural function of the V2 discourse-linking position. Doval Suárez and González Álvarez (2012) found that Spanish L2 English learners’ texts show an under-representation of contrastive clefts in particular, assuming that these learners prefer to use pseudo-clefts and other lexical or syntactic means for this purpose. Hasselgard (2014) also found an under-representation of clefts for Norwegian L2 English learners.

**Raising and alternations**

Amongst the syntactic phenomena that still represent difficulties even for very advanced and near-native learners are those that exhibit a comparatively large distance or mismatch in the mapping of syntactic surface form to semantic meaning
DeKeyser (2005, p. 7) argues that such form-meaning mappings can still be difficult if the link between form and meaning is not transparent, and that such lack of transparency can be due to at least three factors: redundancy, optionality, or opacity. Opaque syntactic constructions can be argued to represent a form of complexification of form-function mappings in the sense of a ‘hidden’ complexity that, in contrast to overt complexity which is accessible through (morpho-)syntactic surface patterns, must largely be inferred from context (Bisang, 2009). The result is that what looks simple in surface structure is based on a complex background of potential inferences which adds hidden complexity to seemingly simple structures.

A case in point are raising constructions of the type *The paper is easy to read* (object-to-subject-raising or *tough*-movement) and *The reports are claimed to be false* (subject-to-subject-raising). Raising constructions serve IS purposes by moving subjects/objects to topic/focus position if contextually needed. When compared with their underlying, bi-clausal variants, they exhibit a larger distance between syntactic surface form and semantic meaning, since the grammatical subject/object is not the logical or semantic subject/object. This results in ambiguity and potential vagueness of surface forms. The same applies to so-called oblique-subject alternations, which involve verbs that canonically occur with agentive subjects but, if required by IS, can also take as subjects noun phrases that would normally occur as a prepositional phrase. In a series of studies, Callies (2006, 2008a, 2008b, 2009, 2013) showed that non-prototypical subjects, from a semantic and syntactic point of view, remain difficult for advanced learners and are under-represented in their written production, often due to avoidance strategies. In two corpus studies of advanced German and Polish L2 English learners, Callies (2008a, 2008b) found that due to the high degree of typological markedness of English raising constructions these are under-represented in the writing of advanced learners, most likely due to avoidance. A quantitative and qualitative textlinguistic analysis of the corpus data revealed that *tough*-movement in particular is significantly under-represented in advanced learner writing, and that the learners have problems with the appropriate use of all types of raising constructions in written discourse in terms of IS and thematic progression. In an experimental study, Callies (2006) found that advanced German L2 English learners provided comparatively low acceptability rates for non-agentive subject constructions like *This paper discusses...* due to the high degree of typological markedness and interference from L1, which lacks or disprefers those structures. Such non-agentive subject constructions are preferred reporting strategies in the humanities but are significantly underproduced in advanced learners’ academic writing (Callies, 2013). The middle alternation (*These books sell well*) is a construction that poses similar challenges but that has not yet been examined in depth (see Gao, 2015, for an exception).

**Grammatical variation**

A final line of research into advanced SLA comprises studies on lexicogrammatical variation. These studies adopt a variationist perspective on SLA focusing on L2 learners’ acquisition of the influence of several determinants that govern
constituent order and the choice of constructional variants, among them important 
components of IS such as information status, animacy, and the end-weight prin-
ciple. While the many structural, semantic, discourse-motivated, and processing-
related determinants that influence grammatical variation have been widely 
studied in corpus-based research on L1s (especially in English), lexicogrammatical 
variation has not been well researched in L2 to date and is only beginning to attract 
researchers’ attention. Only two more recent studies shall be mentioned here 
briefly. Jäschke and Plag (2016) examined the role of probabilistic grammatical 
constraints on the dative alternation in German EFL and found that L2 learners are 
influenced by the same determinants (e.g. syntactic weight) as L1 speakers but to 
a lesser degree. Their findings also suggest that, initially, the learners do not make 
use of probabilistic constraints in spite of the constraints being influential in the 
L1, and only gradually acquire a sensitivity toward those that govern the choice 
between the two dative constructions. Gries and Wulff (2013) studied the genitive 
alternation in advanced Chinese and German EFL. Their multifactorial study 
showed that the learners rely heavily on processing-related factors, which can be 
overridden by semantic constraints, and that their choices are differentially mod-
ulated by their L1.

Combining experimental and corpus evidence: 
The way forward

(Quasi-)Experimental and introspective data, mostly gained through highly 
controlled elicitation techniques, have traditionally been favored in SLA 
research. Such techniques are typically used when researchers working with an 
analytic-deductive research design intend to test a specific hypothesis. In such 
designs, many learner and contextual variables have to be controlled, which is 
extremely difficult if not impossible in non-experimental settings. However, 
corpora and corpus linguistic tools and methods are also increasingly used for 
the study of SLA, in particular in learner corpus research (LCR), an interdisci- 
plinary field at the crossroads of corpus linguistics, SLA research, and foreign 
language teaching (see Granger, Meunier, & Gilquin, 2015). Learner corpora 
are large systematic collections of authentic, continuous, and contextualized 
language use (spoken or written) produced by L2 learners in less controlled 
tasks. Thus, they can help overcome several problems and limitations posed by 
the dominance of data elicitation techniques in SLA to date. Gilquin and Gries 
(2009) suggest that there is no strict dichotomy between corpora on the one 
hand and experiments on the other. Corpora, just as for linguistic data in 
general, can be located on a continuum of naturalness of production and 
collection.

Still, theoretical and methodological skepticism has been an obstacle to the 
establishment of genuine bi-directional links between SLA and LCR (see Myles, 
2015), which seems especially true for SLA research couched in a generative/
formal framework in which more data-driven, usage-based approaches are not
as appealing. But recently, researchers in both camps have realized the potential that a combination of different types of learner data provides (e.g. Callies, 2009; Gilquin, 2007; Mendikoetxea & Lozano, submitted; Meunier & Littré, 2013; Tracy-Ventura & Myles, 2015). For example, Rankin (2009) argues that the study of interface relations is a field that should be beneficial for both LCR and (generativist) SLA. LCR has made a substantial contribution to the description of advanced written interlanguages with a focus on lexicogrammar. Some corpus-based research that was reviewed in the section “Functional approaches to word order and IS in advanced SLA” showed that advanced learners typically struggle with the acquisition of optional and highly L2-specific linguistic phenomena, often located at the interfaces. This partially mirrors recent developments in generative SLA, where the interfaces between syntax, discourse-pragmatics, and semantics have been at the center of attention.

As of yet, there are only a few studies that have addressed the methodological interface between generative approaches to SLA and LCR (e.g. Lozano & Mendikoetxea, 2008, 2010; Mendikoetxea & Lozano, submitted; Rankin, 2009). These have focused on discourse-conditioned word order alternations such as subject–verb inversion and fronting in the written production of advanced learners of L2 English from various L1 backgrounds. Mendikoetxea and Lozano (submitted) argue for the triangulation of naturalistic (corpus) and controlled (experimental) data in a cyclic fashion: Experimental findings can be the departure point for a corpus-based study, which in turn may reveal new findings that can be tested experimentally. Importantly, such triangulation refers to the study of the same interface phenomenon. If learners show certain kinds of knowledge or deficits at the interfaces, this should be observed in both experimental and contextualized production data. Corpus and experimental data should therefore be combined and contrasted to better account for the observed deficits at the syntax–discourse interface and determine why some interface properties are more problematic than others.

NOTES

1 Sorace’s IH should not be confused with the classic weak versus strong Interface Hypothesis (aka, Interface Position), i.e. whether explicit L2 knowledge can eventually become implicit L2 knowledge.

2 Such native optionality is not reported in previous studies (cf. Lozano, 2006a: SV 58%, VS 84%; Lozano 2006b [on a -2 to +2 scale]: SV 1.25, VS 0.17), probably due to differences in the research method (cf. also results in Figure 22.2).

3 In CLRD (clitic right dislocation) the dislocate appears in the right periphery of the sentence.

4 In the generative literature, these interfaces are known as lexicon–syntax and syntax–discourse interfaces (cf. the section “Subject–verb inversion and IS”).
REFERENCES


23  Advanced-Level Semantics

TANIA IONIN
University of Illinois at Urbana-Champaign

Introduction

One of the central tasks of second language (L2) acquisition is to map form (morphosyntax) to meaning (semantics). For example, an L2 learner of English needs (among many other mappings) to map the articles the and a to the semantics of definiteness and indefiniteness, respectively; to map the simple past versus past progressive forms, such as walked versus was walking to the meaning of complete versus incomplete actions; and to map different anaphoric forms (e.g. she vs. herself) to pronominal versus reflexive interpretation. Acquiring semantics also means learning which meanings are (un)available for complex sentences, a task that may involve the (un)learning of ambiguity. For example, in the domain of scope interpretation, an L2 learner of English needs to learn that a double-quantifier sentence such as A student saw every professor is ambiguous (it could mean that the same student saw all professors, or that different students saw different professors). On the other hand, in the domain of reflexive interpretation, an L2 learner of English needs to learn that certain interpretations are unavailable: e.g. a sentence such as Mary thinks that Kate painted herself is not ambiguous, since herself can only refer to Kate, and not to Mary.

The above examples are just a few of the tasks facing L2 learners as they determine which forms map to which meanings in their L2. The task is particularly challenging because languages differ in how they encode the relationship between semantics and morphosyntax. For instance, with reference to the above examples, some languages lack articles, do not have the distinction between perfective and progressive aspect, treat double-quantifier sentences as obligatorily unambiguous, or allow reflexives to have long-distance antecedents. It is well-known that cross-linguistic influence, or L1 transfer, affects the mapping between morphosyntax and semantics (see Slabakova, 2008, 2016, for an overview). On generative approaches to L2 acquisition, such as the Full Transfer / Full Access model
Advanced-Level Semantics

(Schwartz & Sprouse, 1994, 1996), the initial state of L2 acquisition is hypothesized to be the state of the L1 grammar. However, learners with advanced proficiency may well go beyond L1 transfer, and indeed, a number of studies discussed below have found that advanced learners have native-like sensitivity to semantic contrasts in their L2 that are not encoded in the same way in their L1. As discussed in the subsequent sections, learners’ ability to overcome L1 transfer and acquire subtle aspects of semantic interpretation is often attributed to learners’ access to Universal Grammar (UG).

The debate about whether advanced L2 learners are capable of acquiring subtle semantic contrasts in their L2 goes back to the work of Coppieters (1987) and Birdsong (1992) on near-native learners of French. While Coppieters found that these learners, despite years of immersion in French, were non-native-like on certain semantic phenomena, Birdsong found that about 30% of the near-native learners were fully native-like. These studies tested a variety of linguistic phenomena, whereas more recent work on the L2 acquisition of semantics has tended to focus on specific constructions (e.g. articles, grammatical aspect, scope), in order to test the predictions of particular theories of L2 acquisition.

A common assumption in recent L2 research at the syntax–semantics interface, and one articulated by Slabakova (2008, 2016), is that semantics is universal, while the encoding of semantics in morphosyntax is language-specific. Slabakova argues that inflectional morphology serves as a bottleneck of the acquisition process. There is evidence from many domains of L2 acquisition that learners are sensitive to universal semantic distinctions, and able to make use of semantic computational mechanisms. The reason that learners are not target-like stems from their difficulties in acquiring the corresponding functional morphology.

This chapter gives an overview of three types of linguistic phenomena at advanced L2 proficiency. First, the interpretation of the nominal domain (articles and plural marking) is discussed, with particular reference to learners’ access to semantic universals, and to the persistent difficulty of inflectional morphology (cf. Slabakova, 2008). Second, scope interpretation is discussed, in reference to the Poverty of the Stimulus (PoS) problem for L2 acquisition: this domain provides further evidence that learners are able to access universal semantic principles. And finally, several phenomena falling at the syntax–discourse interface are discussed (anaphor interpretation, topicalization) in light of the Interface Hypothesis (Sorace & Filiaci, 2006), which posits a difference between the acquisition of internal and external interfaces.

These three types of phenomena are chosen because they are the subject of much recent L2 literature, and because together they paint a fairly comprehensive picture of how L1 transfer, semantic universals, and processing affect L2 acquisition at advanced proficiency. A major topic in the L2 acquisition of semantics that is left out of this chapter is the semantics of tense, aspect, and mood; for discussion of these phenomena, please see Chapter 18, Advanced-Level Mood Distinction, and Chapter 19, Advanced Conceptualizations of Tense and Aspect in L2 Acquisition, in this handbook. For information about the kinds of experimental tasks used in much research on the L2 acquisition of semantics (e.g. acceptability judgment tasks, truth-value judgment tasks), please see Ionin (2011) and Ionin and Zyzik (2014).
Semantics of the nominal domain at advanced proficiency: Semantics vs. morphology

As discussed in the introduction, on the view of Slabakova (2008, 2016), semantic distinctions are universal, and learners’ difficulties are traceable to the acquisition of language-specific functional morphology. (Slabakova suggests that the difficulties with morphology are ultimately traceable to processing; see also Lardiere, 2009, on the role of feature reassembly—reassembling the features on functional elements—as the main challenge in L2 acquisition). On this approach, learners with advanced proficiency are expected to show sensitivity to semantic contrasts, but may nevertheless not behave exactly like native speakers due to continuing difficulty with functional morphology. The nominal domain provides a number of examples of this.

The mass/count distinction and plural interpretation

With regard to the mass/count distinction, languages can be divided into two broad classes: languages like English, which have obligatorily plural marking and fully grammaticize the mass/count distinction; and generalized languages like Japanese, Mandarin Chinese, and Korean, in which the mass/count distinction is not (fully) grammaticized (see Chierchia, 1998). Work on the acquisition of the mass/count distinction has focused largely on how speakers of generalized classifier languages acquire the count/mass distinction of a language like English. There have been relatively few studies in this area, and fewer still that specifically focus on performance by learners of advanced proficiency.

Snape (2008), testing L1 Japanese and L1 Spanish L2 English learners, found that by advanced proficiency, the two groups were equally target-like in correctly judging count versus mass English nouns (in contrast, at intermediate proficiency, the Japanese group made more errors). Choi, Ionin, and Zhu (2017) examine the mass count distinction in light of the semantic universal of atomicity (Chierchia, 2010; Kim, 2005): a noun is atomic if and only if there exists the smallest element (the atom) which bears the property denoted by the noun. In English, count nouns always denote atomic entities (e.g. chair), but mass nouns can denote both non-atomic entities (e.g. water) and atomic ones (e.g. furniture). Choi et al. found that advanced L1 Korean and L1 Mandarin L2 English learners correctly used plural marking with count nouns and did not use plural marking with non-atomic mass nouns (*waters), but incorrectly overused plural marking with atomic mass nouns (*furnitures). Choi et al. argue that this provides evidence that learners access the semantic universal of atomicity: i.e. learners assume that all atomic nouns have count morphosyntax, and hence are compatible with plural marking.

Finally, a study by Hwang and Lardiere (2013) examined the acquisition of plural marking in the opposite direction, by L1 English L2 Korean learners. Korean has both intrinsic plural marking, which is fairly similar to plural marking in English (although unlike in English, it is often optional), and extrinsic plural
marking, which is used with expressions other than NPs. Hwang and Lardiere found that at advanced proficiency, learners were quite native-like with intrinsic plural marking, but continued to have difficulty with extrinsic plural marking, which has no equivalent in English. The difficulty in this case is with learning the precise conditioning environment for the use of extrinsic plural marking.

**Articles: Definiteness and specificity**

Much L2 literature has considered how L2 learners acquire the interpretation of English articles *the* and *a* when their L1 lacks articles altogether. The learners’ task in this domain is to learn that *the* encodes definiteness, which involves both uniqueness and hearer knowledge. A number of studies have shown that learners from an L1 with no articles, such as Japanese, Korean, or Chinese, continue to make errors with English articles even at advanced proficiency (e.g. Liu & Gleason, 2002). Snape (2008) shows that errors are more pervasive among advanced L2 English learners whose L1 (Japanese) lacks articles, compared to those whose L1 (Spanish) has articles.

Ionin, Ko, and Wexler (2004) proposed that L2 learners have access, via UG, to different semantic features that articles cross-linguistically may encode: these include definiteness as well as *specificity* (defined as “speaker intent to refer to a particular individual”). Under Ionin et al.’s proposal, L2 learners fluctuate between the two possibilities until the input leads them to determine that English articles encode the definiteness rather than the specificity distinction. Consistent with this hypothesis, Ionin et al. (2004) found that even advanced L2 English learners from article-less L1s (Russian and Korean) overused *the* with specific indefinites (in contexts where the speaker has a specific referent in mind which is unknown to the hearer), while being quite accurate at using *a* with non-specific indefinites.

Trenkic (2008) found a similar pattern of errors among L1 Chinese L2 English learners of high-intermediate to advanced proficiency, and argued that the reason behind the errors was a reliance on explicit strategies, rather than on access to specificity via UG (see Ionin, Zubizarreta, & Philippov, 2009, for a response). Finally, Ko, Ionin, and Wexler (2010) argued that learners also rely on the semantic universal of presuppositionality, in addition to definiteness and specificity: a presuppositional indefinite denotes an individual which is presupposed to exist, for example, through previous mention. This proposal is based on the finding that L1 Korean L2 English learners overused *the* with indefinites that denote members of a previously mentioned set (e.g. producing *the puppy* after prior mention of *five puppies*).

**Articles: Genericity**

Another aspect of article knowledge is knowing how to use articles in *generic* contexts, where reference is to a kind/species as opposed to an individual (as in, *The dodo bird is extinct*), or where a generic statement is made about typical characteristics (as in, *Cats like milk*). Languages differ in how genericity is encoded: for example, English and the standard variety of German use bare
(article-less) plurals for generic statements, while most Romance languages, including Spanish and Italian, use definite plurals (the equivalent of The cats like milk) for generic interpretation. A number of studies have found that L2 learners exhibit L1 transfer in this domain, in both English/German \(\rightarrow\) Spanish/Italian and Spanish/Italian \(\rightarrow\) English/German directions. While transfer is more pervasive at lower proficiency levels, there is evidence that it persists at advanced proficiency as well (e.g. Cuza, Guijarro-Fuentes, Pires, & Rothman, 2013; Ionin & Montrul, 2010; Ionin, Montrul, & Crivos, 2013; Kupisch, 2012; Slabakova, 2006). Transfer in this domain is manifested—depending on the study—as mis-interpretation of English definite plurals as generic, overacceptance of generic bare plurals in Spanish/Italian, or failure to allow definite plurals to be generic in Spanish/Italian.

While plural generics have been much studied in the literature, as discussed above, less is known about how L2 learners acquire singular generics, both indefinite (A cat makes a good pet) and definite (The dodo bird is extinct). Ionin, Montrul, Kim, and Philippov (2011) found that L2 English learners from article-less L1s (Russian and Korean) correctly distinguish between different types of generic environments, allowing indefinite singular generics in generic sentences, but disallowing them as reference to kinds (#A dodo bird is extinct). At the same time, even the most advanced learners failed to exhibit knowledge of definite singular generics. Ionin et al. (2011) argue for the role of UG, which informs learners about different types of generic contexts, but also for the importance of input, in which definite singular generics are particularly rare. Snape, García-Mayo, and Gürel (2013) found a similar difficulty with definite singular generics for L2 English learners whose L1s (Japanese and Turkish) lack the definite article, but not for learners whose L1 (Spanish) has the definite singular generic. Once again, this speaks for the role of L1 transfer in the L2 acquisition of articles.

**Summary: Semantics and morphology in the nominal domain**

To sum up, studies of advanced L2 learners’ knowledge of the nominal domain suggest that L1 transfer continues to affect how learners use articles and plural marking even at advanced levels, as long as the learners’ L1 either lacks articles entirely or assigns a different meaning to articles or plural marking than the L2. The persistent difficulty with the correct interpretation of such morphemes as plural ‐s and the articles the and a, coupled with evidence for learners’ reliance on semantic universals (specificity, atomicity, kind-reference), is consistent with the Bottleneck Hypothesis (Slabakova, 2008), on which semantics is universal but inflectional morphology is the bottleneck of the acquisition process. Similar difficulties with inflectional morphology are attested in the L2 acquisition of the verbal domain, e.g. in the acquisition of aspectual interpretation; for more discussion, see Chapter 20, Inflectional Morphology, in this volume.
Scope interpretation at advanced proficiency: Overcoming PoS

The previous section discussed the learning task at the level of the functional element or phrase. But learners also need to master syntax/semantics at the level of sentence. In particular, sentences that contain quantifiers (such as a, every, most, how many, etc.) and/or operators (such as negation or disjunction) present the challenge of determining the scope of these expressions. Studies of scope in L2 acquisition commonly address learners’ ability to overcome the PoS, as discussed below.

The PoS argument in L2 acquisition

The PoS argument, initially formulated for child L1 acquisition (Chomsky, 1965, 1980, and much subsequent work), has also been applied to L2 acquisition (Schwartz & Sprouse, 2000). The PoS argument is about underdetermination: How does the child come to acquire knowledge of the target language (TL) which is not observable in the input, such as knowledge of which sentences in the TL are ungrammatical? (Hornstein & Lightfoot, 1981). The fact that the child does acquire such knowledge therefore serves as an argument for the involvement of UG in L1 acquisition (see Crain, 1991; Crain & Thornton, 1998, for evidence for children’s knowledge of negative constraints on the grammar).

Schwartz and Sprouse (2000) argue that the PoS argument applies to L2 acquisition as well, even though it is well known that L2 learners do not become as uniformly successful as L1 learners. PoS phenomena can in principle be found in any domain of the grammar, including narrow syntax (see e.g. Schwartz & Sprouse, 1994, on the acquisition of German word order). However, many PoS studies in L2 acquisition address phenomena at the syntax–semantics interface (e.g. Dekydtspotter & Sprouse, 2001; Marsden, 2009; Montrul & Slabakova, 2003; Slabakova & Montrul, 2003), because such phenomena lend themselves particularly well to the logic of the PoS argument (see Slabakova, 2016, chapter 10). For example, if the learners’ L1 and L2 exhibit a mismatch, such that a particular form/meaning mapping is present in the L1 but absent in the L2, this constitutes a PoS situation: the absence of a particular form/meaning mapping cannot be deduced from the input (since absence of evidence is not evidence of absence); and L1 transfer is not helpful, since the form/meaning mapping does exist in the L1. As long as the relevant construction is, furthermore, not subject to explicit classroom instruction, the learners are in a PoS situation. Given the right L1/L2 combination, the acquisition of scope interpretation can meet all of the above criteria, and hence allow for the investigation of the PoS problem.

Schwartz and Sprouse deliberately do not distinguish between end-state and developmental PoS effects: if learners at any developmental stage exhibit knowledge of underdetermined properties of the L2, this serves as evidence for the role of UG. Thus, PoS phenomena are not exclusively a purview of advanced proficiency. At the same time, many PoS studies are conducted with learners at
several different levels of proficiency, with the goal of exploring whether learners at least show evidence of UG access by highest proficiency. Under the Full Transfer / Full Access Model (Schwartz & Sprouse, 1994, 1996), learners at lower proficiency may still be treating their L2 like their L1, due to L1 transfer; learners at higher proficiency may recognize that the L2 is not like the L1, and converge on the target grammar thanks to full access to UG.

Scope of quantifiers and wh-expressions

Sentences which contain scope-bearing expressions such as quantifiers (every, some) or negation have the potential to be scopally ambiguous. When the scope-bearing expressions are interpreted in the same order in which they appear in the surface structure, the sentence is said to have a surface-scope reading; when they are interpreted in the opposite order, the sentence is said to have an inverse-scope reading. This is illustrated in (1) and (2) below. In (1), the surface-scope reading is one on which the indefinite quantifier scopes over the universal quantifier; in (2), the surface-scope reading is one on which the negative operator scopes over the universal quantifier.

(1) Someone read every book.
   surface-scope reading (some > every): There exists some person who read all the books.
   inverse-scope reading (every > some): For every book, some (potentially different) person read it.
(2) Mary didn’t read every book.
   surface-scope reading (neg > every): It is not the case that Mary read all the books (but she may have read some).
   inverse-scope reading (every > neg): For every book, it is not the case that Mary read it.

Languages differ in the availability to surface-scope versus inverse-scope readings for specific constructions. For example, when it comes to the interpretation of two quantifiers inside the same clause, as in (1), English allows for ambiguity: while the surface-scope reading is more readily available, the inverse-scope reading is also accessible to native speakers (see Anderson, 2004, and the references cited therein for experimental studies with adult native English speakers; see e.g. Musolino, 1998, for studies of scope in L1 acquisition). In contrast, languages such as Japanese and Korean have ‘frozen scope’ in the configuration in (1), allowing only the surface-scope but not the inverse-scope reading. In Japanese and Korean, in order for the inverse-scope reading to obtain, it is necessary for the object to scramble (move overtly) to the sentence-initial position in front of the subject. With regard to the configuration in (2), English in principle allows both readings, but the inverse-scope reading is ruled out on pragmatic grounds (Musolino, 2006, among others): this reading can be more directly expressed by a sentence such as
Mary read none of the books, so using Mary didn’t read every book to express the ‘none’ reading is pragmatically infelicitous. On the other hand, in Korean, an SOV language in which the object linearly (and on some accounts, structurally) precedes negation, it is the every > neg reading which is preferred for sentences such as (2).

These cross-linguistic differences pose a learnability puzzle: L1 English L2 learners of Korean or Japanese need to unlearn the inverse-scope reading of (1), while L1 Korean L2 English learners need to unlearn the inverse-scope reading of (2). However, the positive evidence (the input) does not inform learners that such interpretations are missing, and furthermore, scope interpretation is not subject to classroom instruction (so there is no negative evidence). Nevertheless, there is evidence that at advanced proficiency, learners are able to unlearn scope readings that exist in their L1 but not their L2. Marsden (2009) examined sentences such as (1), comparing L1 English and L1 Korean L2 Japanese learners (since Korean, like Japanese, lacks inverse-scope readings of (1), there is no learnability problem for this group). While intermediate-proficiency L1 English L2 Japanese learners showed effects of L1 transfer, overaccepting the inverse-scope readings, advanced learners had recovered from L1 transfer, with at least some learners performing in a fully native-like manner. Marsden (2009) argues that this provides evidence that learners are able to solve the PoS problem via access to UG.

In a similar manner, Chung (2013) found that L1 Korean L2 English learners became target-like with proficiency with respect to the sentence type in (2): while low-proficiency learners treated English as Korean, strongly preferring the inverse-scope reading of (2), advanced learners switched to the target pattern of strongly preferring the surface-scope reading of (2). Chung argues that with advanced proficiency, learners are able to compute the pragmatic implicatures necessary to rule out the inverse-scope reading of (2).

The finding that advanced L2 learners are able to unlearn a particular scope interpretation has also been attested for the scope of wh-phrases. Marsden (2008) examined how speakers of English (and other L1s) acquire the interpretation of Japanese wh-questions like (3), given that English allows both individual and pair-list readings, while Japanese allows only the former. Marsden found that while many learners were non-target-like in their overacceptance of the pair-list reading in Japanese, some advanced learners were able to overcome the PoS problem and successfully unlearn this reading.

(3) What did everyone buy?
individual answer: Everyone bought apples.
pair-list answer: Mary bought bananas, John bought apples, Sue bought pears...

Song and Schwartz (2009) examined a somewhat different phenomenon, namely the interpretation of Korean lexical items which are ambiguous between a wh-word and an indefinite interpretation (e.g. mwues can mean either ‘what’ or ‘something’).
As a result of this ambiguity, questions containing these lexical items are potentially ambiguous between a *wh*-question reading (*What did she buy?*) and a *yes/no*-question reading (*Did she buy something?). However, when the subject of the question is a negative polarity item (NPI) such as ‘anyone,’ the interpretation is determined by the surface-scope relations of the NPI versus the *wh*-word: in SOV order, the NPI scopes over the *wh*-word, so that we get a *yes/no* question (*Didn’t anyone buy something?), but in OSV order, the *wh*-word scopes over the NPI, so that we get a *wh*-question (*What didn’t anyone buy?). Song and Schwartz argue that this is a PoS problem, since the relevant sentence type (containing both an NPI and a *wh*-word) is so extremely rare in the input as to be virtually unattested. Nevertheless, Song and Schwartz found that at the advanced proficiency level, both child and adult L1 English L2 Korean learners have native-like production and interpretation patterns of *wh*-word/NPI interactions. This serves as another example of learners overcoming the PoS problem at advanced proficiency.

**The scope of disjunction**

A somewhat different case of scope ambiguity is presented by the interpretation of negative disjunctive sentences such as (4). Grüter, Lieberman, and Gualmini (2010) examined how such sentences are interpreted by both L1 English L2 Japanese learners and L1 Japanese L2 English learners. In Japanese, only the *or>* *negation* reading is possible, while in English, only the *negation>* *or* reading is possible.

(4) Mary didn’t eat the pizza or the hamburger.

*neg>* *or*: It is not the case that Mary ate either the pizza or the hamburger
(she ate neither one).

*or>* *neg*: Mary didn’t eat the pizza OR Mary didn’t eat the hamburger
(so she may have eaten one but not the other).

Grüter et al. address the learnability of sentences such as (4) within the framework of the Semantic Subset Principle (SSP) proposed by Crain, Ni, and Conway (1994). According to the SSP, a child acquiring her L1 should initially opt for the interpretation that makes the sentence true in the smallest set of contexts. The *negation>* *or* interpretation makes (4) true only in the context in which Mary ate neither the pizza nor the hamburger, and false in any other context (where she ate only the pizza, only the hamburger, or both). In contrast, the *or>* *negation* interpretation makes (4) true in the context in which Mary ate the pizza only, or the hamburger only, or neither. This means that the *negation>* *or* reading is the subset, while the *or>* *negation* reading is the superset. Thus, on the SSP, the child should start out assuming that negative disjunctive sentences have the *negation>* *or* interpretation. If the child is acquiring English, this hypothesis will be correct; if the child is acquiring Japanese, she will hear sentences such as (4) used in a context where Mary ate the pizza but not the hamburger (or vice versa), and will correspondingly shift her interpretation to the *or>* *negation* option. Goro and Akiba (2004) tested
this hypothesis and found support for it: while English-acquiring children had an adult-like negation > or interpretation, Japanese-acquiring children were not fully adult-like.

Turning to L2 acquisition, Grüter et al. (2010) hypothesized that L1 English L2 Japanese learners should have an easier time than L1 Japanese L2 English learners acquiring the interpretation of negative disjunction. L1 English L2 Japanese learners are able to make use of positive evidence: hearing (4) in a context incompatible with the negation > or interpretation (such as Mary eating the pizza but not the hamburger), they can shift to the or > negation interpretation. In contrast, L1 Japanese L2 English learners will not encounter any positive evidence inconsistent with the or > negation interpretation. Grüter et al.’s findings support this hypothesis: while most of the L1 English L2 Japanese learners were target-like on the scope of disjunction, very few of the L1 Japanese L2 English were target-like, even at advanced proficiency. Nevertheless, the fact that at least a few L1 Japanese L2 English were able to overcome the PoS problem and acquire the target interpretation provides evidence for access to UG.

Scope ambiguity and (dis)continuity

Another direction in the study of scope ambiguities in second language acquisition (SLA), pioneered by Dekydtspotter and colleagues, has examined how learners interpret continuous versus discontinuous scope-bearing expressions. An example of this phenomenon in French is given in (5), from Dekydtspotter, Sprouse, and Swanson (2001). Both (5a) and (5b) can be answered by specifying how many books each student bought; however, only (5a), and not (5b), can also be answered by specifying how many books the students bought in common. The former interpretation arises if the indefinite de livres ‘of books’ takes narrow scope relative to the universal tous ‘all’; the latter interpretation arises if de livres takes wide scope. The contrast between (5a) and (5b) is captured if de livres is required to take narrow scope in (5b) (where it is in its base position) but is free to take either wide or narrow scope in (5a), where it has moved. (For details of the analysis, and references to the theoretical proposals on this topic, please see Dekydtspotter et al., 2001).

(5) a. Combien de livres est-ce que les étudiants achètent tous?
    how many of books is it that the students buy all
    ‘How many books are the students all buying?’

b. Combien est-ce que les étudiants achètent tous de livres?
    how many is it that the students buy all of books
    ‘How many books are the students all buying?’

Dekydtspotter et al. (2001) investigated whether L1 English L2 French learners are able to acquire the contrast in (5). They argue that this presents a PoS problem, because (i) discontinuous interrogatives such as (5b) are quite rare in the input; and (ii) even if discontinuous interrogatives were frequent, the learners would
only receive the positive evidence that they have the reading on which *de livres* has 
only receive the positive evidence that they have the reading on which *de livres* has 
narrow scope, but would not be informed about the unavailability of the wide-
narrow scope, but would not be informed about the unavailability of the wide-
scope reading of *de livres* in (5b). Dekydtspotter et al. found that advanced L2 
scope reading of *de livres* in (5b). Dekydtspotter et al. found that advanced L2 French learners exhibited a similar pattern of performance to that of native French 
French learners exhibited a similar pattern of performance to that of native French 
speakers: both groups showed an overall preference for narrow-scope rather than 
speakers: both groups showed an overall preference for narrow-scope rather than 
the wide-scope readings of *de livres*, but, to the extent that the wide-scope reading 
the wide-scope readings of *de livres*, but, to the extent that the wide-scope reading 
was allowed, it was allowed significantly more with continuous interrogatives 
was allowed, it was allowed significantly more with continuous interrogatives 
(5a) than with discontinuous ones (5b). Unlike intermediate L2 French learners, 
(5a) than with discontinuous ones (5b). Unlike intermediate L2 French learners, 
who generally disallowed discontinuous interrogatives (probably due to transfer 
who generally disallowed discontinuous interrogatives (probably due to transfer 
from English), advanced L2 French learners not only recognized that such inter-
from English), advanced L2 French learners not only recognized that such inter-
rogatives are grammatical, but assigned the correct interpretation to them. 
rogatives are grammatical, but assigned the correct interpretation to them. 
Dekydtspotter et al. argue that advanced learners have solved the PoS problem 
Dekydtspotter et al. argue that advanced learners have solved the PoS problem 
thanks to access to UG, which informs them about the possible mappings between 
thanks to access to UG, which informs them about the possible mappings between 
morphosyntax and semantics.

Dekydtspotter and Sprouse (2001) and Dekydtspotter and Hathorn (2005) have 
Dekydtspotter and Sprouse (2001) and Dekydtspotter and Hathorn (2005) have 
found similar results for two other types of discontinuous interrogatives in French. 
found similar results for two other types of discontinuous interrogatives in French. 
In both studies, learners were able to restrict the range of interpretations for dis-
In both studies, learners were able to restrict the range of interpretations for dis-
continuous interrogatives (relative to continuous ones) in a native-like manner, 
continuous interrogatives (relative to continuous ones) in a native-like manner, 
despite no evidence in the input informing them of this restriction. Across studies, 
despite no evidence in the input informing them of this restriction. Across studies, 
learners’ performance became more native-like with advanced (or at least high-
learners’ performance became more native-like with advanced (or at least high-
intermediate) proficiency, relative to lower (intermediate) proficiency. Thus, the 
intermediate) proficiency, relative to lower (intermediate) proficiency. Thus, the 
studies by Dekydtspotter and colleagues provide evidence that advanced learners 
studies by Dekydtspotter and colleagues provide evidence that advanced learners 
are able to overcome the PoS problem in the domain of scope interpretation of dis-
are able to overcome the PoS problem in the domain of scope interpretation of dis-
continuous expressions.

**Summary: Scope and the PoS problem**

To sum up, studies of scope interpretation in L2 acquisition overwhelmingly find 
To sum up, studies of scope interpretation in L2 acquisition overwhelmingly find 
that by advanced proficiency, L2 learners achieve fairly target-like interpretation 
that by advanced proficiency, L2 learners achieve fairly target-like interpretation 
of sentences containing scope operators such as quantifiers, *wh*-words, disjunction, 
of sentences containing scope operators such as quantifiers, *wh*-words, disjunction, 
and negation. Evidence on scope interpretation thus provides a strong argument 
and negation. Evidence on scope interpretation thus provides a strong argument 
that L2 learners are able to overcome the PoS problem, unlearning interpretations 
that L2 learners are able to overcome the PoS problem, unlearning interpretations 
which exist in their L1, as well as acquiring interpretations of sentence types that 
which exist in their L1, as well as acquiring interpretations of sentence types that 
are very rare in the input (see also Chapter 20, Inflectional Morphology, in this 
are very rare in the input (see also Chapter 20, Inflectional Morphology, in this 
volume for more discussion of PoS phenomena).

**Interface with discourse at advanced proficiency**

The preceding sections have discussed what advanced L2 learners know about 
The preceding sections have discussed what advanced L2 learners know about 
some phenomena at the interface between syntax and semantics, but that is of 
some phenomena at the interface between syntax and semantics, but that is of 
course not the only interface that underlies linguistic knowledge. According to the 
course not the only interface that underlies linguistic knowledge. According to the 
Interface Hypothesis (IH), first formulated by Sorace and Filiaci (2006) and Tsimpli 
Interface Hypothesis (IH), first formulated by Sorace and Filiaci (2006) and Tsimpli 
and Sorace (2006), and refined in subsequent work (e.g. Sorace, 2011; Sorace,
Serratrice, Filiaci, & Baldo, 2009), it is important to distinguish between internal and external interfaces in the case of L2 learners and bilinguals. Internal interfaces, such as syntax–semantics as well as syntax–morphology, involve formal features, while external interfaces, such as syntax–pragmatics, involve discourse conditions and pragmatic appropriateness. All of the phenomena discussed so far in this chapter fall at the internal syntax–semantics interface.

On the IH, external interfaces are predicted to be more problematic than internal interfaces. This is not to say that no errors are predicted for internal interfaces: learners in the process of development may certainly have difficulties with a variety of interface phenomena. The IH focuses on end-state, advanced L2 learners, and predicts that such learners should have continued difficulty with external interfaces, but not with internal interfaces. These difficulties are argued to be a result of the processing cost (see Sorace, 2011, for more details).

The IH has been much criticized (e.g. Montrul, 2011; White, 2011), in large part on the grounds that the empirical data do not support the divide into the less problematic internal interfaces and the more problematic external interfaces. As discussed in the preceding sections, advanced learners may continue to have difficulty with some aspects of internal interfaces. As discussed below, it is also controversial whether problems at the external interfaces are unsolvable. This section briefly discusses several different linguistic domains in which pragmatic/discourse knowledge is required for successful acquisition.

**Overt vs. null subject interpretation**

The linguistic phenomenon that gave rise to the Interface Hypothesis is the distribution of overt and null pronouns in null-subject languages like Spanish and Italian. In languages where both null and overt subjects are grammatically licensed, their distribution is largely discourse-governed, so that null subjects are preferred over overt subjects in certain contexts (e.g. with a topical antecedent). Much work by Sorace and colleagues has argued that even highly advanced, near-native L2 learners of Italian face residual difficulties with the interpretation of overt pronouns (see e.g. Belletti, Bennati, & Sorace, 2007; Sorace, 2011; Sorace & Filiaci, 2006). Similar difficulty has been found for advanced L1 English L2 Spanish learners (Jegerski, VanPatten, & Keating, 2011; Keating, VanPatten, & Jegerski, 2011).

For example, Sorace and Filiaci (2006) examined the preferred antecedent for null versus overt subjects in L2 Italian. In the Italian equivalent of a sentence such as While she is wearing her coat, the mother kisses the daughter, the underlined pronoun may be either overt or null. Syntactically, both NPs can potentially serve as antecedents for both overt and null pronouns. However, for discourse reasons, the subject (the mother) is the preferred antecedent for a null pronoun, while the object (the daughter) is the preferred antecedent for an overt pronoun. Sorace and Filiaci found that L1 English near-native L2 Italian learners were target-like on null
subjects but not on overt subjects, indicating that they had not fully acquired the discourse properties of Italian overt pronouns.

The difficulty with the discourse properties of overt pronouns provides much of the motivation for the Interface Hypothesis, on which advanced, near-native L2 learners face particular problems with the external syntax–discourse interface. This difficulty may stem from resource allocation in bilingual processing (Sorace, 2011). However, other authors have argued against persistent difficulty with external interfaces. With regard to anaphor interpretation, Zhao (2012, 2014) finds that advanced L2 Chinese learners are able to acquire the interpretative properties of both overt and null elements in Chinese. Rothman (2009), examining overt and null pronouns in L2 Spanish, argues that while syntactic properties are indeed acquired before discourse ones, consistent with the Interface Hypothesis, advanced learners are able to fully master the syntax–discourse interface.

**Binding and co-reference**

Another aspect of anaphor interpretation that has been the subject of much work in L2 acquisition has to do with binding and co-reference. Much of this work is within the framework of Binding Theory (Chomsky, 1980, 1981) and focuses on Principle A, according to which reflexives like *herself* must be locally bound, and Principle B, according to which pronouns like *her* must not be locally bound. This is illustrated in (6) for English: while *herself* requires a local (same-clause) antecedent ((6a): *herself* must be co-indexed with *Susan*), *her* disallows a local antecedent and is instead compatible with a long-distance antecedent ((6b): *her* may not be co-indexed with *Susan*).

(6) a. Mary₁ thinks that Susan₂ painted herself₂/₁.
    b. Mary₁ thinks that Susan₂ painted her₁/₂.

It is well established that many languages, including Chinese, Korean, and Japanese, have some reflexive types which allow long-distance as well as local antecedents (i.e. in the equivalent of (6a), both *Mary* and *Susan* would be possible antecedents). Assuming L1 transfer, for L1 Chinese/Korean/Japanese L2 English learners, the long-distance interpretation of reflexives must be unlearned. Studies have generally found that by advanced level of proficiency, L2 English learners from these L1s are able to correctly restrict English reflexives to the local interpretation in sentences such as (6a), but difficulties remain with respect to other aspects of reflexive interpretation, such as locality when the indefinite is embedded inside an infinitival clause, as in *Mary asked Susan to paint herself*, as well as allowing reflexives to refer to both object and subject antecedents when the L1 has subject-oriented reflexives. For some sample studies on this topic, please see, among others, Thomas (1991), Matsumura (1994), Akiyama (2002), Jiang (2009), and Domínguez, Hicks, and Song (2012). This research has focused primarily on the
role of L1 transfer, and on whether L2 learners’ interpretation of English reflexives is consistent with the options provided by UG.

Unlike reflexives, pronouns require an understanding of the discourse as well as of syntax. In the reflexivity framework (Reinhart & Reuland, 1993; Reuland, 2011), successful use and interpretation of pronouns requires both syntactic knowledge (Principle B) and pragmatic knowledge (termed Rule I). Principle B ensures that in (6b), the local NP (Susan) and the pronoun (her) receive different indices (2 vs. 1). Normally, different indices pick out different individuals (e.g. 2 picks out Susan while 1 picks out Mary). However, in principle, nothing in the syntax prevents different indices (1 and 2 in (6b)) from picking out the same individual in the discourse, resulting in accidental co-reference (e.g. both 1 and 2 happen to point to Susan). Rule I is a discourse rule that prevents such accidental co-reference except under very specific discourse conditions. This framework has been used to explain why young English-acquiring children have more difficulty with the interpretation of pronouns than with the interpretation of reflexives: the basic proposal is that children have the necessary syntactic knowledge (Principles A and B) but lack the pragmatic/discourse knowledge required for the interpretation of pronouns (Chien & Wexler, 1990, and much following literature; see Conroy, Takahashi, Lidz, & Phillips, 2009, for an overview).

In the case of adult L2 learners, White (1998) found that high-intermediate L2 English learners (with Japanese and French as their L1s) had little or no difficulty with the interpretation of pronouns, suggesting that they have the discourse as well as the syntactic knowledge necessary for pronoun interpretation. However, Kim, Montrul, and Yoon (2015), an eye-tracking study with L1 Korean L2 English learners, found that learners were less target-like in their interpretation of pronouns than of reflexives. Kim et al. attribute the difficulty with pronouns to the processing load imposed by interface principles such as Rule I. At the same time, the advanced learners in Kim et al. were more native-like in their performance than lower-proficiency learners, indicating that native-like attainment of pragmatic as well as syntactic aspects of anaphor interpretation is possible.

**Clitics and topicalization**

Other phenomena at the syntax–discourse interface include clitics and topicalization. Two recent studies have examined how speakers of English, a language with no clitics, acquire the discourse properties of clitics in their L2. Ivanov (2012) examined clitics in L2 Bulgarian, while Slabakova, Kempchinsky, and Rothman (2012) looked at clitics in L2 Spanish. Both studies found that the discourse properties of clitics were difficult for learners at lower proficiency levels, but were ultimately acquired at advanced and/or near-native levels, suggesting that syntax/discourse phenomena are acquirable. The same conclusion was reached by Yuan and Dugarova (2012) with regard to advanced L1 English L2 Chinese learners’ acquisition of wh-word topicalization in Chinese.
Summary: Interface phenomena in SLA

To sum up, there is conflicting evidence on whether advanced L2 learners can become fully target-like with respect to phenomena that require discourse/pragmatic knowledge. At the same time, many studies in this area agree that discourse phenomena are more difficult to acquire than purely syntactic phenomena, and that learners’ difficulties are often due to insufficient processing resources.

Conclusion

On the basis of the studies reviewed in this chapter, the following tentative conclusions can be made about advanced-level semantics. First, advanced learners are able to overcome the PoS problem and exhibit target-like interpretations of sentences containing multiple scope-bearing operators. The area of scope interpretation is probably where the greatest success is attested. Second, advanced learners continue to make errors with regard to the correct use of nominal morphology (plural marking and articles, in particular), despite showing sensitivity to such semantic universals as definiteness, specificity, kind-reference, presuppositionality, and atomicity. These two findings are consistent with the Bottleneck Hypothesis (Slabakova, 2008), on which semantics is universal, yet the acquisition of semantics/morphology mapping is potentially problematic. Finally, we see somewhat conflicting findings with regard to phenomena at the external syntax-discourse interface, so that some findings are consistent with the Interface Hypothesis (Sorace & Filiaci, 2006), but others argue against this.

A number of open questions remain for future research. First, greater clarity is needed with regard to whether phenomena at external interfaces are indeed harder to acquire than phenomena at internal interfaces. Given that the difficulties have been attributed to processing load on some more recent accounts (e.g. Sorace, 2011), it is particularly important to study how advanced learners process phenomena at both internal and external interfaces. Some processing work has indeed been done in this area (e.g. Dekydtspotter & Outcalt, 2005; Felser & Cummings, 2012; Kim et al., 2015), but more needs to be done. It would be worthwhile to examine both on-line (processing) and off-line performance with a variety of interface phenomena, including (but not limited to) the phenomena described in this chapter, and to do so with a variety of L1/L2 combinations. A different direction of investigation would be an examination of how advanced learners are able to recover from transfer and acquire PoS phenomena (such as unlearning a particular scope interpretation). We know that such recovery and unlearning do take place, but the mechanisms for how they take place are not fully known; possibilities include indirect negative evidence, triggers in the input, and universal computational mechanisms.

In sum, the picture of advanced-level knowledge of semantics is quite positive, with advanced learners exhibiting sensitivity to subtle syntax/semantics and syntax/pragmatics phenomena. However, more work remains to be done on the exact mechanisms responsible for learners’ performance.
REFERENCES


Snape, N. (2008). Resetting the nominal mapping parameter in L2 English:


Part V  Advanced-Level Pragmatics, Discourse, and Sociocultural Literacy
Advanced-Level Pragmatics in Instructed SLA

FENG XIAO
Pomona College

Introduction

In his seminal work on communicative competence, Hymes (1972) argues for the importance of social and cultural norms reflected in language use in communication. In other words, pragmatic competence, the ability to use language appropriately according to context, is equally as important as the ability to accurately use grammatical rules (e.g. syntax, vocabulary, and discourse). Following Hymes’s definition of communicative competence, Thomas (1983) makes a distinction between pragmalinguistic failure and sociopragmatic failure. The former refers to the inaccurate use of linguistic forms to perform a pragmatic function, such as using imperatives to form an indirect request, while the latter refers to the inappropriate use of linguistic forms in a given context, such as using the want statement to make a direct request to a person of higher social status. The distinction between pragmalinguistics and sociopragmatics implies that an advanced level of second language (L2) proficiency is needed to achieve an advanced level of pragmatic competence. Empirical studies on the relationship between L2 proficiency and pragmatic competence have found an overall positive effect of proficiency on pragmatic competence, but high proficiency does not automatically transfer to high pragmatic competence (for a review see Bardovi-Harlig, 1999, 2013; Kasper & Rose, 1999, 2002; Taguchi & Roever, 2017; Xiao, 2015), which calls for the role of instruction in the development of L2 pragmatic competence.

Since the 1980s, a collective endeavor has been made to examine the effectiveness of pragmatic instruction, many of which studies used a pre/post-test design to compare different instructional methods, and these studies have found that, in general, explicit instruction is more effective than implicit instruction (e.g. Bardovi-Harlig, 2015, 2017; Cohen, 2017; Glaser, 2013;...
Pragmatic competence and instructed second language acquisition

Corresponding to Thomas’s (1983) dichotomy between pragmalinguistics and sociopragmatics, all models of communicative competence adopt a componential view of language competence (e.g. Bachman & Palmer, 1996, 2010; Canale, 1983; Canale & Swain, 1980). In their latest articulation, Bachman and Palmer (2010) theorize that language knowledge includes organizational knowledge and pragmatic knowledge. Organizational knowledge is the knowledge used to organize sentences and texts (i.e. grammatical and textual knowledge), while pragmatic knowledge refers to functional and sociolinguistic knowledge of language use. Functional knowledge represents the knowledge of linguistic forms that can perform pragmatic functions (e.g. using “would you” to make a request), while sociolinguistic knowledge is the knowledge of appropriateness of using linguistic forms in context (e.g. using indirect forms such as “I was wondering + clause” to make a request to someone of higher social status, and using “Can I + verb” to someone of equal social status). These two subcomponents together represent the nature of pragmatic knowledge. In addition to language knowledge (i.e. organizational knowledge and pragmatic knowledge), Bachman and Palmer’s model includes strategic competence, which is the control function of language knowledge in different situations. Therefore, the theoretical construct of pragmatic competence includes both pragmatic knowledge and the appropriate use of such knowledge in context (control of processing). A recent trend in both first and second language pragmatics advocates for a discursive approach to pragmatic language use in interaction (e.g. Kasper, 2006; Levinson, 2013, 2017; Taguchi & Roever, 2017). This approach focuses on the co-construction of communicative acts in the sequential context of conversation (e.g. how to comprehend the preceding turn...
and produce a sequentially relevant response). To this end, advanced-level pragmatic competence should entail a high level of understanding of pragmatic knowledge as well as a high level of ability to use pragmalinguistic forms to achieve communicative acts in interaction.

Pragmatic competence has been included in the descriptors of proficiency levels in language assessment as seen in the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001), and the Test of English as a Foreign Language (TOEFL) (Chapelle, Grabe, & Berns, 1997), but no standardized pragmatic test has been widely used to determine different levels of L2 pragmatic competence. In other words, descriptions of advanced-level pragmatic competence vary across different studies. In fact, existing instructional studies often created their own outcome measures of pragmatic competence. In the studies that focused on high-proficiency learners, the outcome measures often targeted specific pragmatic features such as speech acts (e.g. Alcón-Soler, 2005; Derakhshan & Esfami, 2015; Félix-Brasdefer, 2008; Ghobadi & Fahim, 2009; House, 1996; Nguyen, Pham, & Pham, 2017); pragmatic routines (e.g. Bardovi-Harlig, Mossman, & Vellenga, 2015); gambits (Wildner-Basset, 1984); hearsay markers (Narita, 2012), and conversational implicatures (Bouton, 1994; Kubota, 1995). Therefore, instructional effects found in these studies need to be interpreted with the pragmatic targets and their measures. In addition to the theoretical construct of pragmatic competence and its measure (what to teach), the generalizability of different teaching methods (how to teach) is also of interest to researchers and language instructors because instructed second language acquisition (ISLA) aims to “understand how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development or acquisition of a language other than one’s first” (Loewen, 2014, p.2). In other words, ISLA aims to bridge the gap between SLA research and classroom instruction, advocating a systematic evidence-based approach to language teaching. With this conviction, the present chapter synthesizes existing instructional studies on pragmatic competence of learners with advanced and high-intermediate levels of proficiency. In particular, it addresses the following two synthesis questions:

1. What pragmatic features have been taught to high-proficiency learners in existing studies, and how have these features been measured?
2. What types of instruction have been examined in these studies, and how effective are they?

Method

A keyword search was conducted to locate relevant studies leading up to May, 2017. The search was done by using the Claremont Colleges Library (CCL) search engine. The CCL bibliographical system covers over 200 databases such as EBSCOhost, ERIC, JSTOR, ProQuest, LLBA, MLA, and Web of Science. The target keywords included proficiency, pragmatics, instruction/teaching, and second/foreign
language. The initial search yielded 677 entries. After a manual check of all entry
descriptions with the following criteria, 29 studies (marked with an asterisk in the
references) were included:

1. The study was a published data-driven study in the English language;
2. The study directly examined instructional effects on pragmatic competence of
   adult second language learners;
3. The study used a pre-post design;
4. The study focused on learners with advanced or high-intermediate levels of
   proficiency;
5. The study either indicated the participants’ proficiency level or included
   sufficient information about the participants’ proficiency (i.e. test score, course
   level, length of study);
6. The study did not treat learners with several proficiency levels as one homoge-
   neous participant group;
7. The study examined effects of a single type of instruction with a control group
   OR examined differential effects of multiple methods with or without a control
   group OR examined differential effects of a single type of instruction across
   proficiency levels;
8. Only one of multiple studies that analyzed the same dataset was included.

In this review, studies that had learners with a high-intermediate level of proficiency
were included for two reasons. First, most of the existing studies did not focus on
advanced-level learners. For example, only seven out of the 29 included studies
explicitly stated that their participants’ proficiency was at the advanced level
while nine of them explicitly stated that they focused on high-intermediate/
upper-intermediate-level learners. Second, no criterion-reference proficiency test
has suggested a clear cut-off point between the high-intermediate and the low-
advanced levels of proficiency. In practice, the high-intermediate level of
proficiency can be considered as the threshold for taking advanced-level courses.
As stated in criterion 5, the 29 included studies either explicitly stated the
participants’ proficiency level or included any of the three types of information
(i.e. test score, course level, length of study) that could be used to judge the partic-
ipants’ proficiency level. As a result, 16 of the included studies had an explicit
statement on their participants’ proficiency with or without a proficiency test
score, while the other 13 focused either on learners enrolled in a third-year lan-
guage course or on those with a minimum length of study (LoS) of five years. The
third-year rule was created because a third-year course is often considered an
advanced-level course at college. The five-year LoS rule was created based on four
studies that had learners from a third-year college-level course and also reported
their LoS ranges (Eslami & Liu, 2013; Félix-Brasdefer, 2008, Nguyen et al., 2012,
2017). Their reported LoSs ranged from 5 to 12 years so participants with a minimal
LoS of five years were considered as high-intermediate/low-advanced-level
learners. Table 24.1 shows the proficiency levels and related indicators of the
included studies.
Table 24.1  Proficiency levels and their indicators of the included studies (N=29).

<table>
<thead>
<tr>
<th>Study</th>
<th>Proficiency level</th>
<th>Proficiency test</th>
<th>Course level</th>
<th>Length of study (year range or average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcón Soler (2005)</td>
<td>H1*</td>
<td>P</td>
<td>-</td>
<td>7–10</td>
</tr>
<tr>
<td>Alcón Soler (2007)</td>
<td>H1</td>
<td>P</td>
<td>-</td>
<td>7–10</td>
</tr>
<tr>
<td>Alcón Soler (2015)</td>
<td>H1</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bardovi-Harlig et al. (2015)</td>
<td>A</td>
<td>S</td>
<td>Level 5 out of 7</td>
<td>-</td>
</tr>
<tr>
<td>Bouton (1994)</td>
<td>H1*</td>
<td>-</td>
<td>6th semester</td>
<td>-</td>
</tr>
<tr>
<td>Derakhshan &amp; Eslami (2015)</td>
<td>H1</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eslami &amp; Liu (2013)</td>
<td>H1*</td>
<td>-</td>
<td>3rd years</td>
<td>7–12</td>
</tr>
<tr>
<td>Eslami, Mirzaei, &amp; Dini (2015)</td>
<td>H1</td>
<td>S</td>
<td>-</td>
<td>7–9</td>
</tr>
<tr>
<td>Félix-Brasdefer (2008)</td>
<td>H1*</td>
<td>-</td>
<td>5th semester</td>
<td>5</td>
</tr>
<tr>
<td>Fordyce (2014)</td>
<td>H1</td>
<td>S</td>
<td>3rd year</td>
<td>-</td>
</tr>
<tr>
<td>Fukuya &amp; Zhang (2002)</td>
<td>H1*</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>Gharibeh, Mirzaee, &amp; Yaghoubi-Notash (2016)</td>
<td>H1</td>
<td>-</td>
<td>3rd year</td>
<td>-</td>
</tr>
<tr>
<td>Ghabadi &amp; Fahim (2009)</td>
<td>H1*</td>
<td>-</td>
<td>-</td>
<td>5 (2+ high school)</td>
</tr>
<tr>
<td>Glaser (2016)</td>
<td>A</td>
<td>S</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gu (2011)</td>
<td>A*</td>
<td>-</td>
<td>-</td>
<td>10–15</td>
</tr>
<tr>
<td>Halenko &amp; Jones (2011)</td>
<td>H1</td>
<td>P</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>House (1996)</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>Narita (2012)</td>
<td>H1*</td>
<td>-</td>
<td>3rd year</td>
<td>-</td>
</tr>
<tr>
<td>Nguyen et al. (2012)</td>
<td>H1*</td>
<td>-</td>
<td>3rd year</td>
<td>6–9</td>
</tr>
<tr>
<td>Nguyen et al. (2017)</td>
<td>H1</td>
<td>-</td>
<td>3rd year</td>
<td>6–9</td>
</tr>
<tr>
<td>Sadeqi &amp; Ghaemi (2016)</td>
<td>A</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(Continued)
In order to answer the synthesis questions, the located studies were coded for substantive features (e.g. participants, target pragmatic features, types of instruction, and features of treatment) and findings (Norris & Ortega, 2000).

**Results and discussion**

*Pragmatic features and measures*

Synthesis question 1 asks what pragmatic features have been taught to high-proficiency learners in existing studies, and how these features have been measured. Table 24.2 shows the pragmatic features and measures of the 29 studies. As shown in Table 24.2, the included studies covered six types of speech acts: request (N=13), apology (N=4), refusal (N=3), compliment (N=2), criticism (N=2), and suggestion (N=1). In addition, two studies focused on comprehension of conversational implicatures (Bouton, 1994; Kubota, 1995), two examined pragmatic routines (Bardovi-Harlig et al., 2015; House, 1996), and one targeted epistemic stance (Fordyce, 2014), hedging devices (Wishnoff, 2000), hearsay markers (Narita, 2012) and gambits (Wildner-Bassett,1984). The speech act of request was the most frequently taught pragmatic feature to high-proficiency learners, and like other speech acts, the production of requests was often measured by a discourse completion task (DCT) (see also Chapter 26, Advanced Second Language Pragmatic Competence). In general, speech act production was measured by both written (e.g. DCT and email) and oral production tasks (e.g. oral DCT, role-play, phone conversation). With regard to functional expressions (i.e. hearsay markers, hedging devices, and epistemic stance), production was measured by both written (Fordyce, 2014; Narita, 2012; Wishnoff, 2000) and oral production tasks (Narita, 2012),
Table 24.2 Pragmatic Targets and Measures of Included Studies ($N = 29$).

<table>
<thead>
<tr>
<th>Study</th>
<th>Pragmatic targets</th>
<th>Pragmatic measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcón Soler (2005)</td>
<td>Request</td>
<td>Written request identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dialogue writing</td>
</tr>
<tr>
<td>Alcón Soler (2007)</td>
<td>Request</td>
<td>Request identification</td>
</tr>
<tr>
<td>Alcón Soler (2015)</td>
<td>Email request</td>
<td>Emails</td>
</tr>
<tr>
<td>Bouton (1994)</td>
<td>Implicature</td>
<td>MCQ</td>
</tr>
<tr>
<td>Eslami &amp; Liu (2013)</td>
<td>Request</td>
<td>DCT</td>
</tr>
<tr>
<td>Eslami et al. (2015)</td>
<td>Request</td>
<td>DCT &amp; free constructed emails</td>
</tr>
<tr>
<td>Félix-Brasdefer (2008)</td>
<td>Refusal</td>
<td>Role-play</td>
</tr>
<tr>
<td>Fordyce (2014)</td>
<td>Epistemic stance</td>
<td>Computer-delivered essay writing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Emails &amp; phone conversations</td>
</tr>
<tr>
<td>Fukuya &amp; Martínez-Flor (2008)</td>
<td>Suggestion</td>
<td>DCT</td>
</tr>
<tr>
<td>Fukuya &amp; Zhang (2002)</td>
<td>Request</td>
<td>DCT</td>
</tr>
<tr>
<td>Gharibeh et al. (2016)</td>
<td>Refusal</td>
<td>DCT</td>
</tr>
<tr>
<td>Ghobadi &amp; Fahim (2009)</td>
<td>Compliment</td>
<td>DCT &amp; role-play</td>
</tr>
<tr>
<td>Glaser (2016)</td>
<td>Refusal</td>
<td>DCT &amp; role-play</td>
</tr>
<tr>
<td>Gu (2011)</td>
<td>Request</td>
<td>DCT</td>
</tr>
<tr>
<td>Halenko &amp; Jones (2011)</td>
<td>Request</td>
<td>DCT</td>
</tr>
<tr>
<td>House (1996)</td>
<td>Routine</td>
<td>Role-play</td>
</tr>
<tr>
<td>Kubota (1995)</td>
<td>Implicature</td>
<td>MCQ &amp; written production of target implicatures</td>
</tr>
<tr>
<td>Narita (2012)</td>
<td>Hearsay marker</td>
<td>2 written comprehension tasks &amp; 1 oral production task</td>
</tr>
<tr>
<td>Nguyen et al. (2012)</td>
<td>Criticism</td>
<td>DCT, role-play, &amp; oral peer feedback on writing</td>
</tr>
<tr>
<td>Nguyen et al. (2017)</td>
<td>Criticism</td>
<td>DCT, role-play, &amp; oral peer feedback on writing</td>
</tr>
</tbody>
</table>

(Continued)
while comprehension was measured by written comprehension tasks (Narita, 2012). Comprehension of implicatures was measured by multiple-choice questionnaires (MCQ; Bouton, 1994; Kubota, 1995). Pragmatic awareness was measured by a speech act identification task (Alcón Soler, 2005, 2007), an error detection task (Eslami & Eslami-Rasekh, 2008), and a dictation task (Takahashi, 2013). In summary, target pragmatic features were assessed by using different measures that reflected pragmatic processing across modalities (i.e. production, comprehension, and awareness).

Table 24.2 also shows that three studies (Derakhshan & Eslami, 2015; Eslami & Eslami-Rasekh, 2008; Eslami-Rasekh et al., 2004) targeted multiple speech acts. In so doing, these studies might reveal potential differences in instructional effects across different speech acts. However, their findings showed the advantage of a certain type of teaching method in teaching all targeted speech acts, implying that high-proficiency learners may be less susceptible to difficulties associated with types of speech acts when an effective instruction is applied. While in a less effective learning condition, high-proficiency learners may still be able to detect and produce some target pragmatic features. For example, Takahashi (2013) found no advantage of an implicit condition over a control condition in the detection or production of bi-clausal request forms in L2 English but a causal effect of L2 listening subskill on the target request forms. Taken together, these studies suggest that high-proficiency learners can be instructed to learn multiple pragmatic features at one time and to learn complex pragmalinguistic forms due to their linguistic readiness.

However, most of the included studies focused on speech acts and conducted a morphosyntactic analysis on L2 data by using the coding scheme of the cross-cultural study of speech act realization patterns (CCSARP; Blum-Kulka, House, & Kasper, 1989). The advantage of using this coding scheme is to reveal the head act (the core lexical and grammatical component that carries the function of a speech act and its modifiers in an utterance), but it was not developed to analyze pragmatic language use in interaction. Pragmatic language use in the sequential context of
interaction represents the control function of pragmatic knowledge, and challenges the traditional view of Speech Act Theory in two aspects (e.g. Levinson, 2017). First, a prescribed coding scheme may not indicate the actual pragmatic function of a linguistic form (e.g. a head act) in interaction. For example, the bi-clausal request form (e.g. I was wondering if + clause; Takahashi, 2010b) can be used as an indirect refusal in the following dialogue:

Seller: Can you pay with cash?
Buyer: I was wondering if I can pay with credit card.

This hypothetical dialogue may happen in the situation where the seller accepted the buyer’s best offer but wanted the buyer to pay with cash, and the buyer did not want to directly reject this request lest he lose the deal. This dialogue shows that a linguistic device (e.g. I was wondering if + clause) with a prescribed pragmatic function (e.g. indirect request) can be used to perform a different pragmatic function in interaction (e.g. indirect refusal). In light of this discursive perspective on speech acts, the understanding of cross-cultural differences in pragmatic language use between the L1 and L2 may not be sufficient for L2 learners to co-construct a communicative act in talk-in-interaction. Therefore, pragmatic adaptation (the control of pragmatic knowledge in the sequential context of interaction) can be considered as an indicator of higher-level pragmatic competence and may be included in pragmatic instruction for high-proficiency learners.

The second challenge to the traditional view of speech acts is that speech act categories may not be finite (Levinson, 2017). The included studies focused on six types of speech acts: request, apology, refusal, compliment, criticism, and suggestion, orienting us to think about other types of speech acts that are under-represented in existing literature (e.g. offering help, inviting, and ordering), and to distinguish between different labels that may overlap in scope (e.g. suggestion, recommendation, and advice). Moreover, different theoretical frameworks may also establish different basic types of speech acts. For example, from a Conversation Analysis perspective, agreement and disagreement are two important speech acts ascribed to responding turns. However, these two types of speech acts are not often discussed in the traditional framework of speech acts (e.g. Austin, 1962; Searle, 1976). The large variation in names and types of speech acts suggest that in L2 pragmatics, it may be less important to include many pragmatic features (e.g. multiple types of speech acts) to represent the construct of pragmatic competence, but more important to select those that can well represent the construct, because the advanced-level pragmatic competence may not be indicated by the breadth of pragmatic knowledge alone, but also by the ability to ascribe the most relevant pragmatic function to a heard utterance and react to it reciprocally (pragmatic adaptation). High-proficiency learners may have the advantage of learning pragmatic knowledge from instruction but they may not necessarily demonstrate a good control of such knowledge in interaction. Therefore, different from most of the existing studies that only focused on utterance-level speech acts and functional words, a comprehensive definition of pragmatic competence needs to include
pragmatic adaptation because it can serve as a good indicator of advanced-level pragmatic competence.

**Pragmatic instructions and their effectiveness**

Synthesis question 2 asks about types of pragmatic instruction and their effectiveness in teaching high-proficiency learners. Of the 29 studies, 25 were framed in the explicit-implicit teaching paradigm, representing the status quo of instructional studies in L2 pragmatics. Table 24.3 shows details of the treatment conditions and results of the 25 studies.

As shown in Table 24.3, the 25 studies can be clustered into three groups: studies that compared an explicit group and a control group (the EC group, \( N = 11 \)), studies comparing an implicit group and a control group (the IC group, \( N = 3 \)), and studies that compared an explicit group and an implicit group (the EI group, \( N = 11 \)).

**Explicit vs. control** There were 11 studies in the EC group. Findings of these studies support the effectiveness of explicit instruction for high-proficiency learners. For example, Bardovi-Harlig et al. (2015) used a computerized oral DCT to examine the effectiveness of explicit instruction on routines in the speech acts of agreement, disagreement, and clarification. The participants were low-advanced-level learners of English in the United States (proficiency indicated by TOEFL). The explicit group had reading materials with target routines highlighted and completed noticing activities such as listening tasks on these targets. The instructor provided explicit metapragmatic information and reviewed answers to different questions in these activities. In the end, the learners completed oral production activities. The control group received no treatment. The findings showed that the explicit group showed a significant pre-post improvement while the control group did not. The positive effect of explicit instruction on the immediate post-test was evidenced in all 11 EC studies.

However, three of the four EC studies that had a delayed post-test showed that this positive instructional effect was not maintained over time. Alcón Soler (2015) found that after receiving explicit instruction, high-intermediate-level Spanish learners of English as a second language (ESL; proficiency determined by a standardized placement test) had a significantly better improvement in their use of mitigators in email requests than those in the control group. The explicit instruction included metapragmatic information, follow-up production activities, and explicit feedback. The positive effect was maintained for three months but not for seven months after the immediate post-test. The weak retention rate of explicit instruction was also reported in Halenko and Jones’s (2011) study on requests. Their participants were high-intermediate-level Chinese learners of English as a foreign language (EFL), whose proficiency level was determined by a placement test. After receiving explicit instruction on sociopragmatics and pragmalinguistics of English requests and completing the follow-up oral production tasks, the explicit group outperformed the control group on the immediate post-DCT, but the instructional effect was not maintained six weeks later. The decrease of this
Table 24.3  Effectiveness of explicit and implicit instruction (N=25).

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment</th>
<th>DoT</th>
<th>Time between IPT &amp; DPT</th>
<th>TL</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcón Soler (2005)</td>
<td>E, I, &amp; C</td>
<td>15×2 h (15 w)</td>
<td>-</td>
<td>-</td>
<td>Mixed</td>
</tr>
<tr>
<td>Alcón Soler (2007)</td>
<td>E &amp; I</td>
<td>15×2 h (15 w)</td>
<td>3 w</td>
<td>-</td>
<td>Mixed</td>
</tr>
<tr>
<td>Alcón Soler (2015)</td>
<td>E &amp; C</td>
<td>4×20 min</td>
<td>DPT 1: 3 m DPT 2: 7 m</td>
<td>+</td>
<td>Mixed</td>
</tr>
<tr>
<td>Bardovi-Harlig et al. (2015)</td>
<td>E &amp; C</td>
<td>4×50 min (2 w)</td>
<td>-</td>
<td>+</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Bouton (1994)</td>
<td>E &amp; C</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Eslami et al. (2015)</td>
<td>E, I, &amp; C</td>
<td>12 w</td>
<td>-</td>
<td>-</td>
<td>E&gt;I&gt;C</td>
</tr>
<tr>
<td>Eslami &amp; Liu (2013)</td>
<td>E &amp; C</td>
<td>10×50 min (10 w)</td>
<td>-</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Eslami-Rasekh et al. (2004)</td>
<td>E &amp; C</td>
<td>12×30 min</td>
<td>-</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Félix-Brasdefer (2008)</td>
<td>E &amp; C</td>
<td>2×75 min (1 w)</td>
<td>1 m</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Fordyce (2014)</td>
<td>E &amp; I</td>
<td>4×45 min (4 w)</td>
<td>5 m</td>
<td>-</td>
<td>E&gt;I</td>
</tr>
<tr>
<td>Fukuya &amp; Martínez-Flor (2008)</td>
<td>E &amp; I</td>
<td>6×2 h (16 w)</td>
<td>-</td>
<td>-</td>
<td>Mixed</td>
</tr>
<tr>
<td>Fukuya &amp; Zhang (2002)</td>
<td>I &amp; C</td>
<td>7×50 min (1 w)</td>
<td>-</td>
<td>-</td>
<td>I&gt;C</td>
</tr>
<tr>
<td>Gharibeh et al. (2016)</td>
<td>E &amp; C</td>
<td>12×90 min (12 w)</td>
<td>-</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Ghobadi &amp; Fahim (2009)</td>
<td>E &amp; I</td>
<td>13×30 min (13 w)</td>
<td>-</td>
<td>-</td>
<td>E&gt;I</td>
</tr>
<tr>
<td>Gu (2011)</td>
<td>E &amp; I</td>
<td>4×20–30 min</td>
<td>-</td>
<td>-</td>
<td>E&gt;I</td>
</tr>
<tr>
<td>Halenko &amp; Jones (2011)</td>
<td>E &amp; C</td>
<td>6 h</td>
<td>6 w</td>
<td>+</td>
<td>Mixed</td>
</tr>
<tr>
<td>House (1996)</td>
<td>E &amp; I</td>
<td>14 w</td>
<td>-</td>
<td>-</td>
<td>Mixed</td>
</tr>
<tr>
<td>Kubota (1995)</td>
<td>E, I, &amp; C</td>
<td>20 min</td>
<td>1 m</td>
<td>-</td>
<td>Mixed</td>
</tr>
<tr>
<td>Narita (2012)</td>
<td>E &amp; C</td>
<td>4×30 min (1 m)</td>
<td>1 m</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Nguyen et al. (2012)</td>
<td>E, I, &amp; C</td>
<td>10×45 min (10 w)</td>
<td>5 m</td>
<td>-</td>
<td>E&gt;I&gt;C</td>
</tr>
<tr>
<td>Nguyen et al. (2017)</td>
<td>I &amp; C</td>
<td>10×45 min (10 w)</td>
<td>5 m</td>
<td>-</td>
<td>I&gt;C</td>
</tr>
<tr>
<td>Sadeqi &amp; Ghaemi (2016)</td>
<td>E &amp; C</td>
<td>10×90 min (10 w)</td>
<td>-</td>
<td>-</td>
<td>E&gt;C</td>
</tr>
<tr>
<td>Shahla et al. (2014)</td>
<td>E &amp; I</td>
<td>4 m</td>
<td>-</td>
<td>-</td>
<td>E&gt;I</td>
</tr>
<tr>
<td>Takahashi (2013)</td>
<td>I &amp; C</td>
<td>3×40 min</td>
<td>-</td>
<td>-</td>
<td>No effect</td>
</tr>
<tr>
<td>Wishnoff (2000)</td>
<td>E &amp; C</td>
<td>2 w</td>
<td>-</td>
<td>+</td>
<td>E&gt;C</td>
</tr>
</tbody>
</table>

positive effect was also reported by Félix-Brasdefer’s (2008) study on L2 Spanish refusals. His study showed a positive effect of explicit instruction on refusal production in high-intermediate-level learners (proficiency determined by course level) in the immediate role-play task, but the difference between the explicit group and the control group was smaller one month later. These findings suggest that although explicit instruction has a positive effect on pragmatic competence of high-proficiency-level learners, retention of this positive effect may need to be enhanced.

**Implicit vs. control** Three included studies compared pragmatic performance between an implicit group and a control group. Fukuya and Zhang (2002) used a pre-post design to examine the effect of recasts on requests produced by high-intermediate-level Chinese EFL learners (proficiency determined by a placement test). Different from direct error correction, a recast only reformulates a learner’s utterance with correct linguistic forms. In this study, a recast was used in a role-play to indirectly raise the learners’ awareness of their inappropriate production of requests. Their findings revealed significant gains in both accuracy and appropriateness of request production on the post-DCT in the implicit group. Likewise, Nguyen et al. (2017) showed a positive effect of implicit instruction with input enhancement and recasts on the production of criticism in high-intermediate-level Vietnamese EFL learners (proficiency measured by course level), and this positive effect was maintained five months after the immediate post-test. In contrast, Takahashi (2013) found no effect of implicit instruction on production of bi-causal request forms in advanced-level Japanese EFL learners (proficiency indicated by TOEFL) after three 40-minute aural input sessions. These findings suggest that even for high-proficiency learners, pure input with no enhancement or follow-up activities is not effective, while input with follow-up activities and feedback (recast) is effective in teaching speech act production to high-proficiency learners.

**Explicit vs. implicit** Eleven included studies compared effectiveness between explicit and implicit instruction. None of them favored implicit instruction over explicit instruction, while six of them showed preference for explicit instruction (Eslami, et al., 2015; Fordyce, 2014; Ghobadi & Fahim, 2009; Gu, 2011; Nguyen et al., 2012; Shahla et al., 2014). For example, Nguyen et al. (2012) analyzed and compared the effectiveness of explicit and implicit instruction on teaching English criticism. The explicit instruction included corrective feedback, whereas the implicit instruction used recast instead. Their findings showed a significant advantage of explicit instruction over implicit instruction in teaching criticism to higher-intermediate-level Vietnamese EFL learners (proficiency measured by course level). Eslami et al. (2015) also contrasted corrective feedback (explicit) with recast (implicit). The explicit instruction provided corrective feedback after consciousness-raising instruction on request forms and strategies, which were followed by production practice and performance discussion. With all other components alike, the implicit instruction provided input enhancement (e.g. bold-faced request forms) and guided reflections on cross-linguistic differences in request
making and recasts. Their data from a DCT and email responses revealed an advantage of explicit instruction over implicit instruction on request production in high-intermediate-level Iranian EFL learners (proficiency indicated by a standardized test). Similar findings were reported by Gu’s (2011) study of the production of English requests (measured by a DCT), Shahla et al.’s (2014) study of English apologies (measured by a DCT), and Fordyce’s (2014) study of the production of English epistemic stance. Findings of these studies suggest the preference of explicit methods in teaching pragmatics to high-proficiency learners.

However, the other five studies in the EI group revealed mixed findings (Alcón Soler, 2005, 2007; Fukuya & Martínez-Flor, 2008; House, 1996; Kubota, 1995). Alcón Soler (2005) examined instructional effects on the awareness and production of L2 English requests. The participants were high-intermediate-level Spanish EFL learners (proficiency indicated by a placement test) in three groups: the explicit, implicit, and control group. The explicit instruction included watching authentic film excerpts on requests, explicit metapragmatic information, and follow-up production activities with self-correction, while the implicit included scripts of the same film excerpts with target forms in bold and follow-up production practice with self-correction. The control group received no instruction. Pragmatic awareness was indicated by the number of request strategies identified from the film excerpts while request production was measured by a DCT. The findings showed that the two experimental groups outperformed the control group on both the pragmatic measures. The explicit group outperformed the implicit group in gains of request production but not in pragmatic awareness. In her follow-up study (Alcón Soler, 2007), the data of a delayed post-test were added into the analysis. The findings showed that, three weeks after the immediate post-test, the explicit group outperformed the implicit group on both pragmatic awareness and request production, showing a higher retention rate of explicit instruction. These findings suggest that, compared to implicit instruction, explicit instruction may not show an immediate advantage on the improvement of both pragmatic awareness and production, but pragmatic awareness and production improved through the explicit condition can be maintained longer. In other words, the positive effect of explicit instruction is affected by modality. This claim is also supported by Fukuya and Martínez-Flor’s (2008) study on English suggestions. Their findings showed that both explicit and implicit instruction improved suggestion production in high-intermediate-level Spanish EFL learners (proficiency measured by a placement test). The explicit group had more gains in oral production (phone conversation) but not in written production (email). Mixed findings due to modality were also found in the comprehension and production of English implicatures. Kubota (1995) found that high-intermediate Japanese EFL learners (proficiency indicated by length of study) in the explicit and implicit group both outperformed their counterparts in the control group on the immediate post-test but the instructional effect was only maintained on pragmatic production for one month. There was no significant difference between explicit and implicit instruction. In addition to the mediating effect of modality, House (1996) found that advanced-level German EFL learners (proficiency indicated by length of
study) in the explicit group had higher gains in the token and type frequency in
the use of routines (gambits) and strategies of requests (assessed by a role-play
task), but they did not outperform the implicit group in the production of requests
as a communicative act in interaction over the course of 14 weeks of instruction.

In summary, studies that compared explicit and implicit instruction have indicated
an overall preference of explicit pragmatic instruction for high-proficiency
learners. However, the positive effect of explicit instruction may be mediated by
the modality of pragmatic processing required for the outcome measure (e.g.
Alcón Soler, 2005, 2007; Fukuya & Martínez-Flor, 2008; Kubota, 1995). Moreover,
compared to the use of routines and strategies in utterances, production of com-
municative acts in interaction may be more difficult for high-proficiency level
learners to improve through explicit instruction (e.g. House, 1996).

However, explicit and implicit instruction may not be considered as a dichotomy
but a continuum (e.g. Jeon & Kaya, 2006; Takahashi, 2010b) because both types of
instruction can include identical practices (e.g. Alcón Soler, 2005, Eslami et al.,
2015, Nguyen et al., 2012) and the difference between explicit and implicit
instruction is often determined by whether or not the treatment includes explicit
instruction or feedback on metapragmatic information (e.g. Kasper, 2001;
Taguchi, 2015).

The argument for a continuum of the explicit-implicit paradigm needs empirical
evidence. Four included studies that did not fit into the traditional explicit-implicit
mode can offer some insight on this issue (Derakhshan & Eslami, 2015; Eslami &
pared effectiveness between explicit inductive and explicit deductive instruction
on refusal production in advanced-level EFL learners (proficiency determined by
TOEFL). The two teaching methods differ in the delivery of metapragmatic
information (i.e. teaching-fronted lecture vs. guided self-discovery). Explicit
deductive instruction introduces metapragmatic rules first followed by examples
while explicit inductive instruction provides examples first and guides learners to
discover metapragmatic rules. In both conditions, the learners also did follow-up
activities and had individual meetings with the instructor to discuss videotaped
role-plays. During the individual meetings, the explicit deductive group received
explicit feedback on their performance in the role-plays. The explicit inductive
group received no feedback but guidance to find more appropriate refusal strat-
egies prior to the instructor’s input. The outcome measures include a DCT and a
role-play task. The findings showed that both types of instruction had a positive
pre-post effect while the explicit inductive group had more gains with regard to
token frequency and appropriateness of the use of refusals. This advantage was
larger for appropriateness than for token frequency.

The effectiveness of self-discovered metapragmatic rules was also reported by
Eslami and Eslami-Rasekh’s (2008) study on identification of pragmatic errors in
requests and apologies (measured by an error identification task) and production
of these two speech acts (measured by a DCT). In the experimental group, advanced-
level EFL learners (proficiency determined by LoS) were guided to discover the
appropriate use of target speech acts followed by oral production but with no
corrective feedback, while in the control group, the learners received no treatment. The findings showed that the experimental group outperformed the control group in pragmatic error identification and production of requests and apologies. Derakhshan and Eslami (2015), on the other hand, compared three different types of self-discovery instruction: peer discussion, role-play, and interactive translation. Participants were high-intermediate-level ELF learners whose proficiency was determined by a placement test. All three participant groups received the same film excerpts and scripts on requests and apologies. The peer discussion group discussed different speech act forms and appropriateness of these forms in different situations. The role-play group acted out by using different target forms. The interactive translation group did L2-L1 translation on the given materials and verbalized their thoughts during the process. The outcome measure was a multiple-choice questionnaire. The findings showed that peer discussion was the most effective method, followed by role-play, and then interactive translation. These findings showed that self-discovery methods were effective for teaching pragmatics to high-proficiency learners but the written-production-based self-discovery (e.g. translation) was not as effective as oral-production-based (e.g. role-play and peer discussion) self-discovery because translation (written production) may not directly reflect the pragmatic language use (oral production) elicited in the film excerpts. Different from the three studies reviewed above, Wildner-Bassett (1984) compared a traditional teaching method used in an intensive English program and the suggestive accelerative learning and teaching (SALT) method on gambit production in advanced-level German learners from a company (proficiency level judged by the author). The traditional method could be considered more explicit because it presented a list of target gambits with a discussion on their functions in responding turns (i.e. agreement and disagreement, p.143). The follow-up activities included explicit feedback while reviewing gambits used in the given video excerpts and a follow-up role-play task. The SALT method could be considered less explicit because prior to introduction of the target gambits, the SALT treatment started with a presentation of a dialogue with nonverbal cues (e.g. gesture and intonation) to engage the learners. In the follow-up practices where feedback was given, music and relaxation methods were used along with nonverbal cues to facilitate learners’ oral production. The outcome measure was a role-play task with a native speaker. One of the findings showed that both groups had significant pre-post gains but there was no between-group difference in gambit production in terms of token and type frequency (indicated by the richness ratio).

The findings of these four studies suggest that in an effective instructional condition (e.g. the explicit condition), the positive instructional effect can be mediated by different methods of metapragmatic information delivery (Eslami & Eslami-Rasekh’s 2008; Glaser, 2016; Wildner-Bassett, 1984) and modalities of pragmatic processing embedded in instructional activities (Derakhshan & Eslami, 2015).

In essence, the answer to the second synthesis question is that high-proficiency learners can benefit from both explicit and implicit instruction but explicit instruction is more effective than implicit instruction due to direct introduction of metapragmatic information. In an explicit condition, the way of delivering
metapragmatic information (e.g. teaching-fronted lecture vs. guided self-discovery) can make the positive effect more or less significant, and the learning outcome is also affected by modality. In an implicit condition, input with enhancement and implicit feedback (e.g. recast; Fukuya & Zhang, 2002; Nguyen et al., 2012) are effective while pure input exposure is not (Takahashi, 2013).

However, contradictory results were found in two studies that directly examined the possible mediating effect of proficiency on the relationship between pragmatic instruction and learning outcomes. Fordyce (2014) treated general proficiency (indicated by TOEFL) as a categorical variable. He conducted a repeated measures analysis of variance (ANOVA) analysis and found no proficiency effect on written production of L2 epistemic stance. Takahashi (2013) treated proficiency subskill (listening ability, measured by a standardized English test) as a covariance in her structural equation model and found large contributions of listening ability to the learners’ awareness and oral production of bi-clausal request forms. Given the scarce number of related studies, no conclusion can be made at this time. However, one possible explanation for the contradictory results is that due to the involvement of different modalities of information processing, proficiency subskills (e.g. listening, speaking, reading, and writing) may have a different relationship with pragmatic performance. Listening and speaking subskills are directly related to the process of aural information while reading and writing subskills are directly related to the process of visual information. Takahashi (2013) used a computerized oral DCT (with aural prompts) to measure pragmatic production and found a direct contribution of listening subskill to pragmatic production. The positive proficiency effect found in her study may be due to the direct link between listening ability and the processing of aural information in her oral DCT. In Fordyce’s (2014) study, the learners’ writing subscore but not the composite TOEFL score for general proficiency might have a positive effect on the targeted written pragmatic production.

Conclusion and future directions

This chapter synthesized 29 previous studies that examined instructional effects on pragmatic competence of high-proficiency learners and found that these instructional studies focused on speech acts, conversational implicatures, and functional expressions (e.g. hedging devices and hearsay markers), and they used receptive tasks (e.g. MCQ) to assess pragmatic comprehension and awareness and productive tasks (e.g. DCT and role-play) to measure pragmatic production. An overall positive instructional effect was found to teach these pragmatic features to high-proficiency learners. In a more effective instructional condition (e.g. an explicit condition), high-proficiency learners’ pragmatic gains were less affected by types of pragmatic features but were affected by modality (i.e. awareness, comprehension, and production), while in a less effective instructional condition (e.g. an implicit condition with no input enhancement or recasts), high-proficiency learners were still able to learn some complex pragmalinguistic forms (e.g. bi-clausal request forms). Three areas in this camp of research can be improved in future research.
First, most of the reviewed studies focused on utterance-level pragmalinguistic forms (e.g. routines and speech acts) and conducted morphosyntactic analyses by using the CCSARP coding framework. The current trend in pragmatics advocates the control function of pragmatic processing in the sequential context of interaction, which underscores the importance of pragmatic adaptation (the ability to comprehend the preceding turn and respond to it reciprocally; e.g. Levinson, 2017; Taguchi & Roever, 2017). Future studies should examine instructional effects on pragmatic adaptation because pragmatic adaptation taps into the processing aspect of pragmatic competence in the sequential context of interaction, and relies on the spontaneous use of pragmatic knowledge in the co-construction of communicative acts. This ability of discursive pragmatic language use may be more difficult to develop because knowing the basic form-function-situation mapping (pragmalinguistics and sociopragmatics) cannot guarantee appropriateness in the ongoing co-construction of talk-in-interaction. Therefore, pragmatic adaptation can be considered an indicator of advanced-level pragmatic competence, and empirical evidence is needed to reveal which type of instruction is effective in teaching this aspect of pragmatic competence.

Second, more studies need to examine the possible mediating effects of proficiency subskills (e.g. listening, speaking, reading, and writing) on the instruction of different aspects of pragmatic competence (e.g. pragmatic knowledge and pragmatic processing) because both proficiency and pragmatic competence are multifaceted constructs. Proficiency subskills represent learners’ ability to process information in aural and visual forms while many reviewed studies in this chapter suggest that the learning outcome of pragmatic instruction is affected by the modalities of pragmatic processing involved in instructional activities (e.g. Derakhshan & Eslami, 2015) and outcome measures (e.g. Alcón Soler, 2005, 2007; Fukuya & Martinez-Flor, 2008). Therefore, listening and speaking subskills may be more correlated with pragmatic tasks that require the processing of aural information (e.g. oral production) while reading and writing subskills may be more correlated with pragmatic tasks that require the processing of visual information (e.g. written production and comprehension). In order to examine these complex relationships between proficiency subskills and different aspects of pragmatic competence along with instructional effects, future studies need to shift from univariate analysis (e.g. t-test and ANOVA) to multivariate analysis (e.g. multiple analyses of variance, MANOVA) and analysis of variance and covariance structures (e.g. structural equation modeling). In this way, the complexity of constructs under examination (i.e. proficiency and pragmatic competence) can be accurately represented in inferential statistical analysis.

Finally, replication research is needed to extrapolate the findings of existing studies. Since the majority of the reviewed studies used a quasi-experimental design, the instructional effectiveness should be replicable if future studies adopt the same research design and use the same outcome measure. For example, a study can replicate an existing instructional method by using learners with different L2 proficiency levels so that the findings of this study could be compared with the findings of all studies that examine the same instructional method. This type of
replication is especially useful in pragmatic instruction because very few studies
to date have tested the effectiveness of a certain type of instruction across different
proficiency levels (e.g. Fordyce, 2014). Future studies should do so because their
findings can shed light on proficiency-level-specific pragmatic instruction. In
addition to the replication of instructional methods, future studies should also
replicate the outcome measures, in order to find the most reliable pragmatic task
that can be developed into a standardized pragmatic test for classroom assessment.

REFERENCES

*Alcón Soler, E. (2005). Does instruction work for learning pragmatics in the EFL
*Alcón Soler, E. (2007). Fostering EFL
learners’ awareness of requesting through explicit and implicit consciousness-raising
Clevedon, UK: Multilingual Matters.
interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. Language learning, 49(4), 677–713.
(Eds.), The Routledge handbook of instructed second language acquisition (pp. 224–225).
New York: Routledge.
*Bouton, L. (1994). Can NNS skill in interpreting implicature in American English be improved through explicit
Canale, M. (1983). From communicative competence to communicative language
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to
Cohen, A. (2017). Teaching and learning
Council of Europe. (2001). Common European framework of reference for languages:
Learning, teaching, assessment. Cambridge: Cambridge University Press.
instruction on the pragmatic development of apology and request. 


Advanced Reading Proficiency in Collegiate Foreign Language Learners

KEIKO KODA\textsuperscript{1} AND SIHUI ECHO KE\textsuperscript{2}

\textsuperscript{1} Carnegie Mellon University
\textsuperscript{2} University of Kentucky

Introduction

In response to the global demands for individuals with high levels of proficiency in a foreign language, the language teaching community has strongly endorsed intellectually challenging, content-driven approaches to foreign language instruction. The endorsement has resulted in various forms of content-driven pedagogy. As a tool for gathering content information, reading ability is integral to all forms of content-driven instruction. Despite its centrality, however, it remains uncertain as to what constitutes reading proficiency in a foreign language, how it develops, how its development relates to linguistic knowledge, and how best we can promote it in language classrooms.

To address these questions, we first describe the major components of reading ability and explain how each component relies on a distinct facet of linguistic knowledge. Based on the description, we conceptualize advanced reading proficiency from multiple perspectives, including language learning, reading, and knowledge acquisition. We then demonstrate how advanced proficiency is viewed in language proficiency guidelines and advanced-level foreign language textbooks.
Language proficiency and reading ability

The term reading proficiency implies that the competence it refers to comprises two partially overlapping constructs: reading ability and language proficiency. In conceptualizing advanced reading proficiency, we must first clarify how the two constituent constructs have been viewed in their respective fields.

Changing views of language proficiency in second language research

Over the past three decades, the definition of language proficiency has been expanded and refined. In the early years, language proficiency was seen as knowledge of the language code (e.g. grammar and vocabulary) and a set of discrete skills in isolation (e.g. spelling and pronunciation). The first wave of refinement occurred in the 1970s through the incorporation of appropriate language use as a core component of the construct (Canale & Swain, 1980; Hymes, 1972). In the subsequent wave, the construct was expanded to include strategic competence that manages language use behaviors as an additional core dimension of the construct (Bachman & Palmer, 1996, 2011). In recent years, the notion of communicative competence has been extended to language use for learning and communicating disciplinary content in academic contexts (Nagy & Townsend, 2012; Schleppegrell, 2004; Snow & Uccelli, 2009; Zwiers, 2007). This wave of extension highlights the importance of genre-specific language use in conceptualizing the construct.

From a cognitive perspective, Hulstijn (2011, 2015) defines language proficiency as consisting of two discontinuous domains—that is, basic and higher language cognition (BLC and HLC, respectively). He describes BLC as the domain that “all native speakers have in common” (p. 21), including tacit grammatical knowledge and phonological perception, irrespective of educational background. In contrast, HLC pertains to the dimension “where differences between native speakers can be observed” (p. 21) as a result of education and other types of learning and training experience. He contends that while not all aspects of BLC may be easily attained in a foreign language, high levels of HLC could be achieved by well-educated adult language learners. In fact, recent studies have demonstrated that HLC is attainable for adult foreign language learners (Dewey, Clifford, & Cox, 2015; Leaver & Campbell, 2015).

The distinction between BLC and HLC provides a conceptual basis for analyzing how linguistic knowledge upholds reading development in the native language and how reading ability in the native language supports foreign language learning. To understand the functional and developmental reciprocity between language proficiency and reading ability, it is useful to clarify how reading ability has been conceptualized in the reading literature.

Changing views of reading ability

Traditionally, two diametrically opposing views have dominated reading research. One regards reading as an indivisible whole while the other sees it as a constellation of distinct competencies. Goodman (1967, 1969), as a strong proponent of the
holistic view, contended that learning to read is a natural process that occurs during the course of human development. Language is learned as a whole through communication. Because reading involves a communicative use of language, it is inseparable from language and should also be learned as a whole. In contrast, the componential view posits that reading is a constellation of distinct capabilities, which can be isolated for inspection either individually or in tandem. Because individual differences exist in virtually all facets of reading ability, it is essential to determine which particular variations are centrally related to comprehension performance. The componential view thus is useful in isolating the sources of reading problems attributable to a deficiency in a single skill or a combination of multiple skills.

Models of reading processes have also changed dramatically. In early studies, reading comprehension was described as a top-down process, in which the reader draws on prior knowledge to guide comprehension. For example, the “psycholinguistic guessing game” model (Goodman, 1973) describes the reader’s primary task as generating a hypothesis regarding the forthcoming text content. The reader pays attention to words selectively only for confirming the hypothesis.

In the subsequent years, however, the hypothesis received little support from empirical studies. Eye movement research, as an illustration, repeatedly showed that most content words received direct visual fixation (Balota, Pollasek, & Rayner, 1985; Just & Carpenter, 1980, 1987); that the absence of even a single letter was disruptive, heavily diminishing reading efficiency (McConkie & Zola, 1981; Rayner & Bertera, 1979); and that text comprehension was impeded by a single anomalous word (Kintsch, 1998). Contrary to the predictions from the top-down conceptualization, these findings indicate that the majority of words in a text are thoroughly processed during text reading, and that the emerging text interpretation does not override incoming information extracted from individual words.

The current literature uniformly describes reading as continual interactions between information explicitly presented in a text and stored knowledge of the reader (Adams, 1990; Nassaji, 2014). Such a view implies that reading entails, at minimum, two sets of skills involved in linguistic processing for text meaning construction and conceptual maneuvers for connecting text information with prior knowledge. The componential view serves well in our quest for a better understanding of complex text-reader interactions. A clear grasp of the component skills of reading and their disparate linguistic demands is vital for identifying language-specific problems associated with reading development in a foreign language.

**Contributions of linguistic knowledge to reading**

**Component skills of reading**

Reading entails three interlinked operations, including text-meaning building, personal-meaning construction, and knowledge refinement (Koda, 2016). Text-meaning building involves (i) analyzing word forms, (ii) retrieving word meanings, and (iii) assembling word meanings into larger text segments, such as sentences and
paragraphs. The operation entails a set of skills that are necessary for converting graphic symbols to meaningful text segments that correspond to the learner’s real-life experiences and background knowledge stored in memory. Those skills include:

- segmenting the graphic form of a word into its phonological and morphological constituents;
- identifying the word;
- retrieving its meaning from memory;
- integrating word meanings into the emerging interpretation of the text;
- inferring the meaning of an unfamiliar word or phrase based on local text meanings and prior knowledge;
- integrating local text segments into a coherent text meaning.

Personal-meaning construction is important for achieving a deeper text understanding. To do so, constructed text meanings must be personalized at both the local and the global levels. At the local level, the learner must draw on her prior knowledge to infer what is implied by unstated information. At the global level, personalization is needed for the learner to make out how text information fits with her real-life experiences. Through this process, locally assembled text meanings are connected with relevant information stored in long-term memory. The operation involves the following skills:

- comparing text information with the learner’s real-life experiences;
- comparing text information with what the learner knows about the topic;
- comparing the view presented or implied in the text with the learner’s own view on the topic.

The third operation, knowledge refinement, involves the incorporation of personal text meanings into the learner’s knowledge bases. According to Britton (1994), a new understanding emerges when text information induces a change in a relationship, or a set of relationships, among stored concepts. Learning during reading only occurs when the learner recognizes such conceptual restructuring in the existing knowledge bases as a result of reading. Thus, knowledge acquisition is dependent on and restricted to what the learner already knows. The operation relies on an additional set of skills:

- reflecting on similarities between text information and what the learner knows about the topic;
- reflecting on differences between text information and what the learner knows about the topic;
- reflecting on a change, however subtle it might be, in the learner’s knowledge of the topic.

In short, reading involves the simultaneous utilization of the learner’s knowledge of the language and that of the world for the purpose of constructing the meaning of a text and personalizing the constructed meaning to generate a new meaning.
Contributions of linguistic knowledge to reading operations

As described above, text-meaning building heavily relies on linguistic knowledge. This section describes the specific facets of linguistic knowledge contributing to the acquisition and use of the skills involved in text-meaning building.

Orthographic knowledge

Successful comprehension requires rapid and effortless access to word meanings. Good readers recognize many words instantly and holistically. They are also adept at analyzing the graphic form of an unfamiliar word, such as letters and letter clusters, to infer its sound and meaning (e.g. Ehri, 1998, 2014; Hogaboam & Perfetti, 1978; Shankweiler & Liberman, 1972; Share, 2008). Seidenberg and McClelland (1989) define orthographic knowledge as “an elaborate matrix of correlations among letter patterns, phonemes, syllables, and morphemes” (pp. 525), contending that the knowledge evolves through forming inter-letter associations through the cumulative experience of decoding and encoding word meanings in print. The more frequently a particular pattern of letter sequences is experienced, the stronger the associations that hold them together. In essence, effortless word recognition occurs as a result of a durable representation of a word in memory (e.g. Adams, 1990; Ehri, 1994, 1998, 2014; Seidenberg & McClelland, 1989). Once formed, orthographic knowledge becomes a powerful mnemonic device that bond the spellings, pronunciations, and meanings of specific words in memory (Ehri, 2014, p. 5).

Phonological knowledge

Phonological decoding refers to the processes involved in accessing, storing, and manipulating phonological information (Torgesen & Burgess, 1998). Studies consistently document that poor readers are handicapped in a variety of phonological tasks. Their deficiencies tend to be “domain-specific, longitudinally predictive, and relatively unaffected by non-phonological factors – such as general intelligence, semantic, or visual processing” (Share & Stanovich, 1995, p. 9). The primary function of phonological decoding is to offer quick access to stored meanings of familiar spoken words (Frost, 1998). Efficient phonological decoding also enhances the functioning of working memory (Kleiman, 1975; Levy, 1975). Because phonologically encoded information is more durable in working memory than visually encoded information (Gathecote & Baddeley, 1993), efficient phonological decoding generates stronger representations and thereby assists the integration of encoded information into larger linguistic units. It is agreed that efficiency in phonological decoding is causally related to word reading, learning, and comprehension.
Morphological knowledge

As the smallest functioning unit in the composition of words, morphemes serve as the basis for word formation. Knowledge of morphemes thus is centrally related to all skills involved in text-meaning building. In learning to read, children rely on their emerging understanding of a word’s internal structure in figuring out how orthographic patterns correspond to recurring morphological constituents. Such insight plays a vital role in reading acquisition in English and other languages whose writing systems directly encode morphological information (Ehri, 2014; Frost, 2012; Nunes & Bryant, 2006). In all languages, morphological knowledge becomes critical in later stages of reading development in which word learning and content learning become inseparable. According to Nagy and Anderson (1984), roughly 60% of the new words children encounter in printed school materials are structurally transparent multi-morphemic words, such as “fire-fight-er” and “un-lady-like.” This implies that the meaning of at least half of the new words could be inferred through morphological decomposition. Morphological knowledge thus bolsters the capacity for identifying familiar elements in an unfamiliar word, and in so doing, enables children to extract partial information from the unknown word. Without such capacity, it is virtually impossible to fill semantic gaps created by unfamiliar words during text comprehension.

Vocabulary knowledge

The ability to retrieve word meanings interacts directly and reciprocally with every one of the other operations in reading. As an illustration, it depends on accurate and speedy analysis of word forms (orthography, phonology, and morphology). At the same time, it relies on local text meanings for deriving the context-appropriate meaning of individual words. The knowledge connects the graphic form of words in a text with the learner’s knowledge stored in long-term memory. Such mediation is necessary because stored knowledge of word forms has an arbitrary relation to meaning representations (Schreuder & Flores d’Arcais, 1992). Word meanings in a way serve as passcodes to stored knowledge bases as they include “information about the things to which words refer – be they related to the external world or internal states of the mind” (p. 422).

As a complex construct, vocabulary knowledge emerges gradually through repeated encounters with a word referring to a particular object, event, or property in particular situations. Because words display different meanings in different contexts, retrieval of a word’s meaning must include the selection of the sense that best fits the context in which the word appears. For example, the word ‘car’ can evoke all the different images of cars stored in memory—ranging, perhaps, from a sleek convertible car to a wrecked car in a scrap yard. The selection of the context-appropriate meaning of the word depends on the emerging interpretation of the local text into which its meaning is incorporated. Anderson and Nagy (1991) underscore the importance of flexibility in word meaning selection during reading.
comprehension by stating that “really knowing a word ... always means being able to apply it flexibly but accurately in a range of new contexts and situations” (p. 721).

In recent years, academic vocabulary has attracted considerable attention. The central concern of this research is how best instruction can foster the genre-specific domain of this knowledge. Using the metaphor of “words as tools,” Nagy and Townsend (2012) explain that knowing words implies the ability to use the knowledge as tools for communicating and thinking about disciplinary content. As noted earlier, academic language proficiency fundamentally differs from social communicative competence. It is essential that reading instruction take into account the genre-specific purposes for which academic words are used and provide the learner with opportunities to use the instructed words for those purposes.

**Syntactic knowledge**

Sentence comprehension entails the incremental integration of individual word meanings in such a way that an integrated ‘chunk’ reflects the overall meaning of larger linguistic units, such as phrases and clauses. The integration process, often referred to as ‘syntactic parsing,’ involves two major operations: phrase construction through lexical-information integration, and case assignments to the constructed phrases. To illustrate, the sentence ‘Nancy tapped the man with the cane’ allows two interpretations regarding the cane holder. If the phrase ‘with the cane’ is taken as a modifier of the verb ‘tapped,’ Nancy is the cane holder. If, on the other hand, the phrase is interpreted to modify ‘the man,’ the cane should be in his hand. Hence, decisions regarding phrase attachment have major semantic consequences, and syntactic knowledge is integral to this process.

Syntactic parsing varies from one language to another, so second language learners must learn how phrases are constructed and how cases are assigned to the phrases in a new language. It has been reported that syntactic knowledge significantly contributes to reading performance among school-age second language learners (e.g. Nagy, McClure, & Mir, 1997; Verhoeven, 2000). More recent studies have shown that syntactic knowledge is a strong predictor of reading comprehension among collegiate foreign language learners (Jeon, 2011; Kim & Cho, 2015; Shin & Kim, 2012; Shiotsu, 2010).

**Discourse knowledge**

To build coherent text representations, locally constructed text meanings must be integrated across sentences and paragraphs. A text’s surface structure offers a variety of reliable clues signaling coherence relations among text elements. Significant elements often are placed in prominent text locations to highlight their relative weight, and they are connected with other text segments in detectable ways (Goldman & Rakestraw, 2000). Linguistic devices are also used to achieve text coherence, such as connectives and co-references. Studies have demonstrated
that knowledge of coherence devices differs considerably among native English-speaking children (e.g. Garner et al., 1986); that explicit training on coherence awareness tends to improve text comprehension and memory (e.g. Pearson & Fielding, 1991); that explicit demonstrations of text organization generally improve text comprehension (e.g. Baumann & Bergeron, 1993; Buss, Ratliff, & Irion, 1985); and that efforts to increase the structural salience of a text generally facilitate comprehension (e.g. Anderson & Davison, 1988; Beck & Dole, 1992). These findings suggest that while knowledge of discourse structure and of devices of signaling text coherence affect development discourse knowledge, its acquisition occurs only through substantial reading experience, and formal training can expedite the process by directing attention to specific text features.

To sum up, reading entails an assortment of diverse skills, each necessitating a distinct facet of linguistic knowledge. Without sufficient linguistic knowledge, text-meaning building is impossible. Conversely, reading provides opportunities to use linguistic knowledge and related skills for analyzing and reflecting on text meanings. Thus, in short, reading ability and language proficiency are functionally and developmentally interdependent, mutually enhancing their development and refinement. Importantly, advanced levels of academic language proficiency are recognized as the ability to use language for learning and thinking about disciplinary contents (Nagy & Townsend, 2012). Similarly, higher levels of reading ability are described as the capacity for using decoding skills to make connections between text information and stored knowledge of the reader (Adams, 1990).

Reading proficiency in a foreign language builds on reading ability acquired in the native language and emerging knowledge of the target language. To achieve advanced proficiency, the learner must acquire (i) sufficient knowledge of the target language to build text meanings on her own using available language aids, such as a bilingual dictionary and vocabulary glossary (Koda, 2016) and (ii) decoding skills in the target language to gain access to relevant information stored in memory (Favreau & Segalowitz, 1982; Haynes & Carr, 1990). These findings indicate that advanced proficiency cannot be achieved until the learner acquires sufficient linguistic foundations that support the constructive and autonomous use of the ability in question for augmenting and refining her knowledge. Following this logic, advanced reading proficiency in second language acquisition can be defined as the ability to use available resources in two languages to construct, analyze, and reflect on text meanings in the target language.

**Advanced L2 reading proficiency in practice**

To promote higher levels of reading proficiency, it is essential that language instruction foster the autonomous and constructive use of language for knowledge expansion and refinement. It is not clear, however, how reading proficiency is viewed in the language teaching community and how their views are reflected in reading instruction in collegiate foreign language classrooms. To address these
issues, we analyzed widely circulated language proficiency guidelines and advanced-level textbooks widely used in Chinese programs in US colleges and universities. Two questions guided our analysis: (i) To what extent is reading ability distinguished from linguistic knowledge? (ii) To what extent is knowledge expansion incorporated in reading instruction?

**Analysis of proficiency guidelines**

We compared three proficiency guidelines, including *Performance Descriptors for Language Learners* (referred to as ACTFL; the American Council on the Teaching of Foreign Languages, 2015), *Common European Framework of Reference for Languages* (referred to as CEFR; Council of Europe, 2011), and *Chinese Language Proficiency Scales for Speakers of Other Languages* (referred to as CLPS; the Office of Chinese Language Council International, 2007). Table 25.1 presents an overview of the guidelines, the descriptions of reading proficiency, and linguistic and non-linguistic knowledge associated with reading proficiency.

Notably, none of the guidelines use the term *reading*. Reflecting the current conceptualizations of language proficiency, all three guidelines describe language ability as a complex multimodal competence that includes responsiveness to contextual variations, genre specificity, and communication modality. In the ACTFL guidelines, the term *interpretive* is used to refer to a set of competencies that correspond to reading. Similarly, in the CEFR, the term *understanding* replaces reading, and in the CLPS, reading is described as an aspect of *written language ability*.

To examine the extent to which reading ability is distinguished from linguistic knowledge, we compared how the construct referred to by the alternate terms (*interpretive* in ACTFL, *understanding* in CEFR, and *written* in CLPS) are characterized. As shown in Table 25.1, all three guidelines describe reading as the skills to detect main ideas and supporting details of a written text. The guidelines list various facets of linguistic knowledge, ranging from knowledge of vocabulary and grammar to text type and text structure, as the major components that support reading performance. Similarly, reading difficulties are estimated by the linguistic demands and structural complexity of a text. Advanced proficiency is characterized as the ability to handle linguistically demanding texts. The ACTFL guidelines include cultural awareness and use of strategies as components of advanced proficiency.

We also compared the extent to which knowledge expansion is incorporated in the descriptions of advanced proficiency. All three guidelines, some more explicitly than others, state that the contents of input source materials are generally familiar to learners. Thus, the guidelines show no expectations that advanced proficiency include the constructive use of language for knowledge expansion. It should be noted that the ACTFL guidelines include the strategy to use background knowledge to solve comprehension problems as a component of advanced proficiency. This reflects the link between the guidelines and the definition of language ability by Bachman and Palmer (2011).
Table 25.1  Summary of performance descriptors of three foreign language proficiency guideline/frameworks.

<table>
<thead>
<tr>
<th>Proficiency guideline/framework</th>
<th>Overview</th>
<th>Proficiency level</th>
<th>Overall performance/global scale/general description</th>
<th>Linguistic knowledge in reading performance</th>
<th>Communication skills in reading performance</th>
<th>Content knowledge in reading performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTFL</td>
<td>3 modes of communication: interpersonal, interpretive, presentational</td>
<td>Advanced range</td>
<td>Interpersonal: Expresses self fully to maintain conversations on familiar topics and new concrete social, academic, and work-related topics, can communicate in paragraph-length conversation about events with detail and organization, and confidently handles situations with unexpected complications, shares point of view; Interpretive: Understands main ideas and supporting details on familiar and some new, concrete topics from a variety of more complex texts that have a clear, organized structure; Presentational: Communicates information and expresses self with detail and organization on familiar and some new concrete topics using paragraphs</td>
<td>Language control, Vocabulary, Accuracy</td>
<td>Communication strategies, Cultural awareness</td>
<td>Uses knowledge of cultural differences</td>
</tr>
<tr>
<td>CEFR</td>
<td>3 modes of communication: understanding (listening and reading), speaking, and writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 modes of domains in relation to language use contexts: personal, public, occupational, and educational</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 types of users (basic, independent, and proficient users) &amp; 6 proficiency levels (A1, the lowest, A2, B1, B2, C1, C2, the highest)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Can understand with ease virtually everything heard or read; Can summarize information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation; Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knowledge of text structure and linguistic complexity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uses text genre knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Proficiency guideline/framework</th>
<th>Overview</th>
<th>Proficiency level</th>
<th>Overall performance/global scale/general description</th>
<th>Linguistic knowledge in reading performance</th>
<th>Communication skills in reading performance</th>
<th>Content knowledge in reading performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLPS</td>
<td>2 modes of communication: spoken (listening comprehension and oral ability); written (reading comprehension and writing ability); No description of “domains,” yet with exemplifications of tasks; 5 bands: 1-5, Band 1 being the lowest, Band 5 being the highest</td>
<td>Band 5</td>
<td>Able to comprehend general language material encountered on a variety of occasions in a variety of fields (including his/her own field of specialty); Able to grasp the important points and synthesize and analyze them; Able to take part fairly competently in communications and discussions on a wide variety of topics, including general topics in specialized fields, using a wide variety of communicative strategies to express his/her opinion and attitude; Able to explain various opinions in a coherent and fairly appropriate manner</td>
<td>Accuracy</td>
<td>Communication strategies (indirect description, e.g. able to grasp their gist and find specific information that one needs; able to grasp the facts and important points, read between the lines, and understand the viewpoints or intention of the author)</td>
<td>Uses text genre knowledge</td>
</tr>
</tbody>
</table>

Analysis of advanced-level textbooks

To examine how the descriptions of reading proficiency are reflected in instruction, we examined several features of three widely used Chinese textbooks, including *The Routledge Advanced Chinese Multimedia Course: Crossing Cultural Boundaries* (referred to as CCB; Lee, Liang, Jiao, & Wheatley, 2014), *Reading into a New China: Integrated Skills for Advanced Chinese* (referred to as RNC; Li & Liu, 2010), and *A Changing China: Advanced Chinese* (referred to as ACC; Wu & Yu, 2014). The features we compared include the learning objectives and expected outcomes, the knowledge and skills that are assumed to underlie advanced-level reading, passage types, post-reading exercises, comprehension questions, and instructional activities.

As shown in Table 25.2, the primary objectives of the three textbooks include fostering culturally appropriate communication skills and helping students consolidate the basic knowledge of the language they have previously learned. These objectives clearly reflect the emphasis on linguistic knowledge as the core constituent of advanced proficiency in the guidelines reviewed above. All three textbooks have devoted major efforts to helping students expand their knowledge of vocabulary and syntax and helping instructors design relevant instructional tasks. For instance, in RNC, vocabulary learning is a specified goal in every lesson. Students are expected to complete a word self-study section by going through a vocabulary list with the parts of speech and English translation of new words, and participating in an activity of guessing unknown compound word meanings based on characters learned previously. In ACC, for the purpose of facilitating students' learning of Chinese-specific grammatical patterns (e.g. zero anaphora), extensive exercises have been implemented in each lesson (e.g. sentence translation, cloze). Of the three, only RNC explicitly specifies reading as the primary focus in advanced courses, although such prominence given to reading is implied in instructional tasks and activities.

Another feature commonly incorporated in the three textbooks includes strategy instruction and exercises. This clearly reflects the ACTFL guidelines. The three textbooks similarly underscore the importance of helping students make a smooth transition from informal conversation skills to formal literacy competence. All three textbooks achieve the objective by reducing the gaps between the two discontinuous proficiency domains. As noted above, advanced proficiency in L2 reading is characterized as the ability to use language for achieving genre-specific purposes in diverse contexts. Nonetheless, the three textbooks aim to support the continuous development of linguistic knowledge (e.g. understanding the differences between colloquial and literary vocabulary in Chinese). None of the tasks and activities in the textbooks is designed to develop the reading skills specifically for analyzing and reflecting on text meanings. This mirrors the emphasis on linguistic knowledge as the central constituent of advanced reading ability in the proficiency guidelines. In the strict sense, there are no real gaps in the instructional foci at the basic and advanced levels.
Table 25.2  A summary of the analysis of three advanced CFL textbooks.

<table>
<thead>
<tr>
<th>CFL textbook acronym</th>
<th>CCB</th>
<th>RNC</th>
<th>ACC</th>
</tr>
</thead>
</table>
| Textbook title in Chinese and English | 文化纵横观  
*The Routledge Advanced Chinese Multimedia Course: Crossing Cultural Boundaries* | 变化中的中国   
*Reading into a New China: Integrated Skills for Advanced Chinese* | 中国新象   
*A Changing China: Advanced Chinese* |
<p>| Targeted language proficiency level | Advanced-low to advanced-mid according to the ACTFL standard; For 6th to 8th semester of Chinese learning | Intermediate-high to advanced-low according to ACTFL standard; For 5th to 6th semester of Chinese learning | Intermediate range according to ACTFL standard; For 5th to 6th semester of Chinese learning |
| Description of learning outcomes | To consolidate their foundation in the language; To improve language skills and cultural literacy; To begin a transition to authentic Chinese literary texts | To develop both fluency and accuracy in Chinese through a topic-based syllabus; Skill integration with a special focus on reading | To develop their language process competency in all four skills to a more advanced degree; To expand explicit knowledge of sociocultural influences on Chinese language use; To apply the knowledge to conduct culturally appropriate spoken and written communication across various social domains and genres |</p>
<table>
<thead>
<tr>
<th>Relevancy to advanced L2 reading competency</th>
<th>Linguistic knowledge</th>
<th>Communication skills</th>
<th>Content knowledge</th>
<th>Reading texts (genres &amp; topics)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grammar and sentence patterns; Character writing system; Written and colloquial vocabulary</td>
<td>Reading comprehension strategies (e.g. dealing with new characters, comprehending the meaning of each sentence, and building up the links between sentences and text coherence)</td>
<td>Cultural literacy; Genre knowledge of Chinese literary texts</td>
<td>Informational and literary texts on popular culture, social change, cultural traditions and history, and politics</td>
</tr>
<tr>
<td></td>
<td>Vocabulary; Word usage and sentence pattern; Grammar accuracy and fluency</td>
<td>Abilities to read independently in the future (e.g. guessing meaning from context, understanding written structures, identifying text organization, making inferences, distinguishing main ideas from supporting facts, and distinguishing facts from speculation)</td>
<td>Differences between oral and formal written language use</td>
<td>Narrative, expository, and new features on population, education, family, gender, environment, business, and technology in the rapidly changing China</td>
</tr>
<tr>
<td></td>
<td>Genre-specific vocabulary and expressions</td>
<td>Relating to one’s own understanding and analysis of a certain topic</td>
<td>Formal language use</td>
<td>Formal and literary texts on current social issues in China such as population, youth, love, and marriage</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th><strong>CFL textbook acronym</strong></th>
<th><strong>CCB</strong></th>
<th><strong>RNC</strong></th>
<th><strong>ACC</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of post-reading exercises</td>
<td>7 types: Main text reading comprehension; Vocabulary and grammar exercises (e.g. fill-in-the-blank, sentence translation, character component analysis); Composition of 100–150 characters; Oral language practicum (e.g. role-play, interview); Follow-up discussion questions for review; Cultural notes; Supplementary text reading</td>
<td>6 types: Pronunciation (e.g. providing the pinyin of target characters); Words and sentence patterns (e.g. collocation, matching, fill-in-the-blank, paraphrasing, translation); Grammar (e.g. cloze); Reading comprehension (e.g. scanning for details; scanning with questions); Speaking (e.g. presenting one’s opinion or attitude toward the readings and searching for supporting evidence); Writing (e.g. writing an essay and presenting one’s opinion or attitude toward a certain topic)</td>
<td>4 types: Vocabulary (e.g. collocation, character analysis, dictionary use, fill-in-the-blank); Grammar (e.g. fill-in-the-blank, translation, paraphrase); Cloze; Writing (text summary, and 500-character argumentative essay about one’s opinion or attitude toward a topic)</td>
</tr>
</tbody>
</table>

*Note.* In the composition/essay writing exercises, instructions are given regarding length and the general topic. No detailed scaffolding is provided. 

*Note.* In the essay writing exercise, only limited scaffolding is provided explicitly, that is, a list of useful sentence patterns and grammatical structures. 

*Note.* In the essay writing exercise, students are given two topics and asked to choose one. No explicit scaffolding is provided.
<table>
<thead>
<tr>
<th>Types of reading questions</th>
<th>Instructional activities</th>
<th>Pre-reading</th>
<th>During reading</th>
<th>Post-reading</th>
<th>Post-reading activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open-ended and multiple-choice questions; scanning for main ideas or local information extraction; author’s intent</td>
<td>Warm-up (e.g. watching video clips and searching websites relevant to the text topic)</td>
<td>Discussion with guided questions</td>
<td>Skimming for main text ideas; Comprehending and searching details</td>
<td>Role-play; Interview; Whole-class discussion</td>
<td>Different reading techniques (e.g. skimming and scanning, learning the style of formal writing in newspapers)</td>
</tr>
</tbody>
</table>
What is of interest is how the stated learning objectives are reflected in passage selection, comprehension questions, and instructional activities. As illustrated in Table 25.2, the three textbooks include similar types of exercises, such as vocabulary and grammar knowledge enhancement, speaking practicum, and essay writing. Evidently, the vast majority of exercises in the textbooks are devoted to the improvement of linguistic knowledge, rather than enhancing reading skills for knowledge refinement. Similarly, reading questions probe the learner’s ability to grasp main ideas that are expressed in the text used as an input source. Only CCB includes reading questions that ask the learner to express her opinions about the main theme of the text.

All three textbooks aim to promote the deployment of reading strategies to enhance text comprehension. For example, all of them incorporate metacognitive activities designed to help the learner activate relevant knowledge before reading, learn new words during reading, and create questions to guide comprehension. Also, RNC devotes a section to introducing different modes of reading, such as skimming and scanning. Thus, the textbooks accentuate reading strategies as a core component of advanced proficiency.

In sum, our analyses revealed that the proficiency guidelines and textbooks we reviewed implicitly assume that (i) reading entails detecting main ideas and important details presented in a written text; (ii) the nature of reading ability is essentially the same across proficiency levels; and (iii) linguistic complexity and text features largely determine reading difficulty. These assumptions indicate that L2 reading proficiency is largely determined by the learner’s knowledge of grammar and vocabulary in the target language. Thus, both proficiency guidelines and textbooks give a privileged role to linguistic knowledge in describing and fostering advanced reading proficiency. It is of interest to explore to what extent these assumptions are held in advanced-level textbooks in European languages.

**Summary and conclusions**

In this chapter, we attempted to conceptualize advanced reading proficiency by dissecting the construct into two major constituents, language proficiency and reading ability. We described how diverse operations of reading are intertwined with distinct facets of linguistic knowledge. Based on the current conceptualizations of language proficiency and reading, we defined *advanced reading proficiency* as the ability to use available resources in two languages to construct, analyze, and reflect on text meanings in the target language. In contrast, lower-proficiency readers can be described as those who have yet to develop sufficient linguistic knowledge that supports text-based meaning construction.

Our analyses of proficiency guidelines and advanced-level foreign language textbooks revealed that reading development is viewed as a gradual increase in the ability to detect main ideas and significant details of written texts that demand progressively more sophisticated knowledge of the target language. Such a view assumes, though implicitly, that linguistic knowledge is primarily responsible for
text understanding. As such, the prevailing practice provides limited opportunities for using the cognitive skills and conceptual knowledge available in the native language as resources to enhance reading comprehension in a foreign language.

Although the notion of linguistic thresholds has been acknowledged in the L2 reading literature, the precise role of linguistic knowledge in the development of diverse reading subskills remains largely unexplored. To gain further insight into L2 reading proficiency and its development, we should consider expanding current research in three particular directions. First, it is important to have a clear grasp of the extent and manner in which linguistic knowledge restrains the use of non-linguistic resources, such as conceptual knowledge and cognitive skills, during reading among adult foreign language learners. Similarly, it is crucial to examine how such linguistic constraints vary across learners of different languages. Finally, from a pedagogical perspective, the language community would be better served by promoting evidence-based practice through seamless integration of L2 reading research and practice. To our knowledge, such integration has just begun to emerge in the area of classroom-based language assessment (Koda & Yamashita, 2017).

The hallmark of advanced proficiency is the autonomous and constructive use of the ability in question for expanding and refining that very ability. It is essential that foreign language pedagogy exploit all of the resources available to collegiate learners to help them achieve advanced reading proficiency. Once reached, it allows them to expand their knowledge, both linguistic and conceptual, exponentially through self-teaching.

REFERENCES


Introduction

LoCastro (2003) defines pragmatics as “the study of speaker and hearer meaning created in their joint actions that include both linguistic and non-linguistic signals in the context of socioculturally organized activities” (p. 15). This definition underscores the key area of study in pragmatics, i.e. a communicative act in a socially situated interaction.

Pragmatic competence requires control of a complex interplay of language, language users, and context of language use. The complexity of pragmatic competence is reflected in the two-layered construct of pragmatics: pragmalinguistics and sociopragmatics (Leech, 1983; Thomas, 1983). Pragmalinguistics refers to the available linguistic resources to perform a communicative function (e.g. complimenting), while sociopragmatics refers to the speaker’s assessment of the context in which those linguistic resources are implemented (e.g. how to compliment whom on what occasion). To become pragmatically competent, second language (L2) learners need a range of linguistic resources (grammar and lexis). At the same time, they need the ability to evaluate layers of contextual information, select appropriate linguistic tools, and use them effectively. Given the combination of linguistic ability and contextual sensitivity involved, pragmatic competence is typically considered as an aspect of advanced L2 competence that takes a long time to acquire.

This chapter presents what advanced-level pragmatic competence looks like by reviewing cross-sectional studies that compare pragmatic performance across proficiency levels. In the 1980s and 1990s, comparative cross-linguistic studies dominated the field of L2 pragmatics research. This trend later expanded to include a comparison of L2 groups coming from different proficiency levels, length of stay
in the target culture, and duration of formal study. The trend still continues today.
Cross-sectional studies, by design, compare data from two or more groups of
varying proficiency or length of study. Any between-group differences are consid-
ered to represent different levels of pragmatic competence. Hence, by comparing
performance between higher- and lower-proficiency groups, we can discern the
areas of pragmatics that advanced learners are capable of handling, which, in turn,
informs our knowledge of the nature of advanced pragmatic competence. Using a
systematic literature review of studies published up to 2016, I will discuss the
characteristics of more advanced learners’ pragmatic competence as opposed to
that of less advanced learners. These characteristics are presented in two of the
areas of pragmatics that have been most commonly studied: comprehension of
indirect meaning and production of speech acts.

Comprehension of indirect meaning

Indirect communication such as sarcasm, irony, humor, and understatement occurs
so commonly that it is hard to imagine everyday interactions without it. Grice (1975)
originally explained the process of indirectness comprehension in his Cooperative
Principles. Consider the following example, adapted from the Santa Barbara Corpus
of Spoken American English (Du Bois, Chafe, Meyer, & Thomson, 2000).

(1) A: So you and your husband just moved to New York?
(2) B: Yes, last year.
(3) A: Do you like living in New York?
(4) B: We went all over the United States, and we didn’t find any place we liked better.

B’s response to A’s question in line 4 is indirect. Although there is no explicit
response (yes or no), A assumes that B is being relevant to the ongoing discourse
and has made an appropriate response to A’s question in line 3. Based on this
assumption of relevance, A draws the conclusion that B likes living in New York.

Comprehension of indirect meaning entails processing linguistic input and con-
textual information, and using them to infer the speaker’s implied intention
behind the utterance. This is a complex task for L2 learners because they have to
first fully understand the surface meaning of an utterance. Then they have to rec-
ognize the mismatch between the utterance meaning and the proceeding discourse,
and resolve the mismatch by inferring the unspoken message. Because of the
multiple levels of processing involved, both at a linguistic and at a non-linguistic
level, comprehension of indirect meaning requires advanced-level proficiency.
Advanced-level learners are equipped with a wide range of knowledge and skills
prerequisite to indirectness comprehension: linguistic knowledge (knowledge of
grammar and vocabulary), comprehension skills (listening and reading abilities),
knowledge of Grice’s (1975) maxims (presumption of relevance), and knowledge
of norms and conventions of interaction in the target language culture.
Indeed, existing findings almost unanimously support that proficiency plays a decisive role in indirectness comprehension (see Taguchi & Roever, 2017, for a review). Cross-sectional studies found that higher-proficiency learners outperformed their lower-proficiency counterparts on comprehension of implicatures and indirect speech acts (e.g., Cook & Liddicoat, 2002; Garcia, 2004; Taguchi, 2011a). Studies using regression-based analyses revealed a significant main effect of proficiency on implicature comprehension (Roever, Wang, & Brophy, 2014; Taguchi, 2005). Longitudinal studies also documented a naturalistic development of implicature comprehension as learners’ proficiency matures with time (e.g., Taguchi, 2007a, 2012).

The positive effect of proficiency found across research designs, L2 groups, and learning environments (e.g., study abroad vs. domestic instructional context) suggests that learners are able to apply their L1-based inferential skills to L2 as long as they have sufficient linguistic abilities to support inference. L1 transfer of the inferential process is possible because the ability to seek relevance of meaning is part of normal human cognition, which develops naturally through the process of socialization in L1 (Baron-Cohen, Leslie, & Frith, 1985; Hala & Carpendale, 1997). What L2 learners need is the threshold language ability that facilitates the transfer of their L1-based inferential practice to L2 comprehension. Advanced proficiency is an advantage in this process.

While the proficiency effect in indirectness comprehension is unquestionable, literature is limited when it comes to identifying the precise aspects of indirectness comprehension that can discriminate between more and less proficient learners. In other words, what aspects of indirectness make comprehension difficult for less proficient learners? What types of implicatures, indirect speech acts, irony, or humor can more advanced learners comprehend, and in which ones do beginners still fall behind? I will address these questions by reviewing existing findings from cross-sectional studies that compared comprehension across two or more L2 groups of different proficiency.

Methods of review

Using ProQuest, which contains over 30 databases including linguistic and educational archives, I conducted bibliographic searches to locate all studies dealing with L2 indirectness comprehension published up to 2016 when I was writing this chapter. I used these subject terms in my searches: ‘pragmatics,’ ‘second language,’ ‘comprehension,’ and ‘proficiency.’ I also searched review articles and handbooks in pragmatics for relevant studies. This search process yielded 176 studies. Then, I analyzed each study according to these eligibility criteria:

1. The study is a data-driven empirical study that examines comprehension of indirect meaning among L2 learners in different proficiency groups.
2. The study has sufficient information about the instrument (with sample items) used to assess comprehension.
3. The study has a clear, objective basis on which proficiency was determined (e.g. standardized exam scores, course/grade levels, and/or length of formal study).
4. Studies that used a correlational or regression analysis between proficiency levels and indirect meaning comprehension were excluded.

This screening process resulted in 11 primary studies (marked with * in the references). I proceeded to code each study for target language, sample size, participants’ L1s, proficiency measures, indirect meaning types, instrument characteristics, and findings (differences between high- and low-proficiency groups).

Of 11 studies, seven studies targeted L2 English, while four studies examined comprehension in L2 Spanish, Japanese, and Chinese. A majority of the studies used audio input with multiple-choice questions, while several studies used written dialogues followed by multiple-choice questions. Three studies incorporated visual input using film or TV excerpts. Studies varied in the target item types: implicatures, indirect speech acts, and sarcasm/irony. About half of these studies determined proficiency with standardized exam scores, while others used course levels to operationalize proficiency. Only two studies had a clearly advanced group of learners (i.e. graduate students in the TESOL/applied linguistics track in a US university; Garcia, 2004; Taguchi, 2002). The remaining studies involved undergraduate participants in upper-level language courses. Since the actual ‘advancedness’ of these participants is uncertain, the following discussion about characteristics of advanced comprehension is only relative, based on the comparison between more and less advanced learners’ performance in each study.

Comprehension of advanced competence in comprehension of indirect meaning

My review of 11 primary studies is guided by what these studies tell us about more advanced learners’ indirectness comprehension. After analyzing study findings for commonalities and discrepancies, several trends emerged. One trend is that more advanced learners (higher-proficiency learners) are able to comprehend a wider range of indirect utterances. This generalization is closely related to the ‘all-around proficiency advantage’ found in many studies. In Roever’s (2005) study, English as a foreign language (EFL) learners from the third to ninth grade in Germany completed a written test assessing recognition of formulaic and idiosyncratic implicatures. Formulaic implicatures were marked by conventional structures (e.g. showing agreement by saying, ‘Is the Pope Catholic?’), whereas idiosyncratic implicatures were conversational implicatures that flout the Gricean maxim of relevance. There was a strong proficiency effect on both types.

Similarly, a series of studies by Taguchi (2009, 2011a) and Taguchi, Li, and Liu (2013) in English, Japanese, and Chinese found that more proficient learners (based on Test of English as a Foreign Language, TOEFL, scores or course levels) performed better on all item types assessed in the studies. The primary focus of these studies was to examine whether conventionality encoded in implicatures can assist comprehension, resulting in higher accuracy scores and shorter response
times. Using an on-line listening test, L2 learners’ comprehension was compared across three item categories: conventional indirect refusals, conventional indirect opinions, and non-conventional indirect opinions. The first category involved convention of discourse, i.e. a conventional pattern of giving an excuse for a refusal (A: ‘Do you want to take a break and go for coffee?’ B: ‘I need to go home pretty soon, so let’s finish it up.’). The second category, conventional indirect opinions, involved convention of form, i.e. conventional, fixed linguistic forms attached to certain meaning. For example, Japanese quantifier adverbs such as chotto (‘a little’) and amari (‘not very’) often convey the speaker’s hesitation and thus mark negative feelings or dislike. The last category, non-conventional indirect opinions, did not involve conventional forms of expression. An example of this type is to convey a negative evaluation of a paper by saying, ‘It is difficult to write an essay in Japanese, isn’t it?’ As this example suggests, the linguistic options for expressing opinion are almost without limit, and they are more idiosyncratic and freer than for other implicature types. Irrespective of target languages, all three studies found that more advanced learners—those who were in the third or fourth year of language courses in a US university or had over 550 TOEFL scores in an EFL university setting—scored significantly higher on all item categories than their lower-proficiency counterparts. Taguchi’s (2011a) L2 English study also found the same proficiency effect on comprehension speed (measured as response times).

As described above, these studies found that proficiency advantage was all-encompassing across all indirect meaning types tested. Other studies also found this distinct proficiency advantage (Carrell, 1979; Cook & Liddicoat, 2002; Garcia, 2004; Koike, 1996; Yamanaka, 2003), but these studies also revealed the interaction between proficiency and indirectness categories. More proficient learners were able to comprehend certain indirect utterances without any problem, but less proficient learners struggled with those utterances. In Carrell’s (1979) classic study, English as a second language (ESL) learners of different course levels in a US university were presented with 10 pairs of request utterances via audio. Each pair had two request utterances, which differed in the polarity of literal and conveyed meaning. See the example below.

(a) Why not color the circle blue? (negative literal meaning/positive conveyed meaning)
(b) Why color the circle blue? (positive literal meaning/negative conveyed meaning)

In (a), literal meaning is negative with a negation marker ‘not,’ but conveyed meaning is positive because the speaker is requesting the hearer to color the picture with the named color. Utterance (b) is the opposite: the literal meaning is positive (without the negation ‘not’), whereas the conveyed meaning is negative (asking the hearer not to color the picture with the named color).

Participants were asked to listen to request utterances followed by a color (blue or other colors) and judge whether the displayed color was appropriate. Results showed that lower-level learners had difficulty judging utterance items in which
the literal polarity is positive while the conveyed polarity is negative (Type B), while higher-level learners were able to judge both types of utterances more correctly.

The difficulty of Type B utterances follows Clark and Lucy’s (1975) findings that native English speakers took a longer time to process an utterance the intended meaning of which was negative (Please don’t do X) than when it was comprehended as positive (Please do X). This is because the negative intended meaning presents a larger mismatch between the literal meaning, the context of request, and the intended meaning. A Type B utterance presents a contradiction between the utterance and context because the speaker questions the hearer’s willingness to do something (Why color the circle blue?) when in fact such a question does not match with the request context where the hearer’s willingness is presupposed. In Clark and Lucy’s study, native speakers were equally accurate with positive and negative intended meaning (both Type A and Type B utterances) and differed only in response times. In the context of L2, Carrell found that learners’ accuracy was lower for negative intention items (Type B), but more advanced learners did not have much difficulty. Hence, this study revealed a feature of more advanced-level indirectness comprehension, i.e. ability to comprehend a negative intention in a positively cued sentence, which presents a larger mismatch between the intended meaning and its context.

Other studies also suggest that the degree of distance between the literal meaning, the intended meaning, and the context of an utterance is a factor distinguishing between more and less proficient L2 listeners. In Cook and Liddicoat’s (2002) study, ESL learners of two proficiency levels (TOEFL scores of 550 as a cut-off) responded to a written questionnaire that contained scenarios followed by a request-making expression. The expressions had three directness levels: direct (e.g. Pass me the salt), conventional indirect (e.g. Can you pass me the salt?), and non-conventional indirect (hinting; e.g. Are you putting salt on my meat?). Although the low-proficiency group had difficulty with both indirect forms, the high-proficiency group only had difficulty with non-conventional indirect forms. Hence, more advanced proficiency is indicated in the ability to deal with indirect forms (albeit conventional), which present a mismatch between the surface meaning and intended meaning. In contrast to high-proficiency learners who can draw on both linguistic and contextual knowledge in disambiguating indirect messages, lower-proficiency learners only have access to linguistic knowledge due to their processing constraints. As a result, they can comprehend only directly expressed meaning.

So far I have discussed advanced learners’ competence in dealing with a greater degree of indirectness, namely the ability to deal with a larger distance between the propositional meaning and intended meaning. There are other studies supporting this generalization, but these studies revealed what features actually increase the degree of indirectness, which in turn cause comprehension difficulty for lower-proficiency learners (Garcia, 2004; Koike, 1996). I will highlight Garcia’s (2004) study here.

Garcia (2004) compared high- and low-proficiency L2 English speakers in their comprehension of indirect speech acts (proficiency based on TOEFL scores). Comprehension of indirect requests, suggestions, corrections, and offers was
assessed using a multiple-choice listening test developed from corpora of naturalistic conversations (four items in each category). Participants listened to a series of dialogues and selected the speech act occurring in each dialogue. High-proficiency learners in this study (graduate students in the TESOL program) were clearly advanced learners, to the level that their comprehension showed no difference from that of native English speakers. Garcia found a significant difference between the high- and low-proficiency groups in all speech acts except indirect requests. One linguistic feature that hindered learners’ comprehension was an indirect agent. Lower-proficiency learners were less competent in identifying the correct illocution when the agent of an action and its recipient were ambiguous, while higher-proficiency learners were still capable of dealing with such indirect agency. See the sample utterance in the speech act of correction (p. 105).

(8) Office Worker: Hmm, he might be out today, or teaching, or
Student: Do you know if he teaches a class today? ’Cause that’s not on his schedule.

In (8), the student complains that the professor is not available during office hours. She accuses the office worker of being wrong about her information about the professor. She does so without establishing the secretary as a specific agent, i.e. the one who is wrong. Only about 40% of the low-proficiency group identified this speech act as correction, while twice as many higher-proficiency speakers and native speakers recognized the correct intention.

In addition to an unclear or implicit agent, false starts and hesitations and lack of conventional modals also influenced speech act recognition. In the sample item below (p. 109), the professor’s offer of help is not very clear because she hesitates several times with the sound ‘uh.’ She restates her offer with ‘If you want me to …,’ but leaves the statement incomplete. Probably because the offer was never realized in a conventional, explicit manner, lower-level students achieved a low accuracy rate of 50%, while that of the high-proficiency group was 80%.

(11) Professor: Actually, the other thing I was gonna recommend too is to uh, give me a, if you want me to look at it sometime, your uh, your cover letter, or your statement.

Use of unconventional modals also adds to ambiguity. As shown in the following example (p. 108), the server uses the modal ‘need to’ instead of the more conventional ‘should’, which seems to have caused difficulty for lower-proficiency learners whose accuracy rate was about 60%, in contrast to 100% in the higher-proficiency group.

(13) Server: But you’re going to need to check with them and make sure that they remove the rest of that.
These findings are consistent with Shively, Menke, and Manzón-Omudson’s (2008) findings that linguistic knowledge greatly affected comprehension of irony. L2 Spanish learners enrolled in the second, fourth, and sixth semester Spanish courses in a US university were asked to interpret the meaning of ironic utterances in films. Higher-proficiency learners recognized irony with higher accuracy than lower-proficiency learners. Due to the lack of knowledge of syntax and vocabulary, lower-level learners were not able to comprehend the meaning of ironic statements. As a result, they were not able to detect a clear opposition between the utterance meaning and the situation, and they comprehended the ironic comments literally.

As shown above, various linguistic features contribute to the implicitness of meaning, including an unclear agent and its recipient, hesitations, false starts, incomplete sentences, and lack of conventional linguistic cues. These features cause difficulty for lower-proficiency learners in their process of identifying the speaker’s implied intention. Hence, more advanced proficiency in indirectness comprehension can be characterized as the ability to block ambiguity coming from these features of implicitness and still draw inferences of meaning by making full use of various resources (e.g. linguistic knowledge, contextual cues, background knowledge and schema, sequential organization of discourse).

Summary

I have reviewed existing findings on L2 comprehension of indirect meaning. Previous studies revealed a clear proficiency benefit for indirectness comprehension. Because relevance-seeking is a general cognitive mechanism that can transfer from L1 to L2, advanced proficiency, exemplified in features such as strong linguistic knowledge, large vocabulary size, and fluent information processing, facilitates such transfer. Drawing on this proficiency advantage, advanced learners demonstrate several characteristics in their comprehension. They can handle a wider range of indirect utterances, including the utterances that involve a greater mismatch between the propositional meaning and intended meaning. They can also cope with various linguistic and non-linguistic features that make meaning implicit and comprehension effortful (indirect agency, unclear utterances, and non-conventional linguistic forms).

Advanced competence in indirectness comprehension corresponds to a greater degree of inference that learners can deal with. The degree of indirectness is closely related to the amount of effort that the listener has to put in to interpret the message, which is affected by the number of contextual cues that must be processed (Sperber & Wilson, 1995). The more indirect the utterance is, the greater the number of linguistic and contextual cues to process. As the number of cues increases, the search for meaning becomes more extensive, leading to greater comprehension effort. More advanced-level learners can process a larger number of cues simultaneously and go through a greater degree of inference to derive a correct interpretation.
Production of speech acts

Speech acts have been the most thoroughly examined construct in L2 pragmatics research. The popularity of speech acts can be explained by the clear connection among form, function, and context of use that a speech act presents, which is useful to operationalize pragmatic competence. The core of Speech Act Theory (Austin, 1962; Searle, 1969) is that we use words to achieve a real-world function. The utterance ‘It’s hot in here’ can be a statement about the temperature of a room, but it can also perform a function of request, hinting that someone should open the windows. Whether it serves as a statement or request depends on the context of the utterance, i.e. the setting, topic, and shared goals between the interlocutors. Hence, context impacts the meaning and consequent action of a speech act.

Like Grice’s (1975) maxims, Speech Act Theory is grounded on the speaker’s intention behind an utterance, but the theory takes a speaker-oriented view of intention by focusing on what speakers do by producing an utterance, rather than how listeners comprehend the intention. Speech Act Theory distinguishes between locution (what is said), illocution (what is meant), and perlocution (effect of the utterance on the listener). Illocution can be conveyed directly or indirectly. Direct speech acts show a clear correspondence between the linguistic form and underlying intention (e.g. using an imperative form in a request), whereas indirect speech acts show no such direct match: intention is conveyed with covert forms (e.g. using a hint in a request). The speaker’s choice of the form (direct or indirect) is based on his/her understanding of contextual factors. Power differences and social distance between the interlocutors, as well as the degree of imposition involved in the given act, are evaluated to guide selection of appropriate speech act strategies, or what Brown and Levinson (1987) call politeness strategies.

Unlike these traditional theories that consider meaning as static, pre-planned, and inherent in an utterance, recent discursive approaches to pragmatics (e.g. Eelen, 2001; Locher & Watts, 2005) contend that meaning is contextually contingent and discursively emergent. Under this approach, speech acts or politeness strategies are not inherent in particular linguistic forms. Illocution or politeness is jointly constructed among participants and is contingent upon the unfolding course of interaction. This view is fundamentally different from traditional pragmatics theories, which view context as involving fixed factors (power, distance, and imposition) and the speaker’s linguistic choice as predetermined corresponding to these contextual specifics.

The impact of discursive approaches is beginning to be revealed in L2 speech act research (Kasper, 2006). We have some empirical data on how a speech act is interactionally constructed in a collaborative manner in a dynamic context (e.g. Al-Gahtani & Roever, 2012, 2014a, 2014b, 2015). However, this trend is not yet prevalent in cross-sectional studies. This is because cross-sectional research typically takes a quantitative research design with the goal of eliciting language samples and comparing them across groups.
This section reviews existing findings from cross-sectional studies that compared production of speech acts between two or more learner groups of different proficiency. The guiding question in this review is the following: What are the linguistic strategies—both utterance-level and discourse-level—that distinguish between more and less proficient learners’ speech act performance? Similar to the previous section on indirectness comprehension, between-group comparisons will be discussed only in relative terms by comparing more and less proficient learners, with proficiency operationalized according to the researcher’s criteria (e.g. standardized test scores, course levels). Although my analysis will not reveal what superior-level L2 speakers can do, it will shed light on the characteristics of more advanced-learners’ speech act performance, as opposed to that of less proficient learners.

Methods of review

I followed the same search procedures as for the previous section on comprehension studies. Using over 30 databases in ProQuest, I identified published studies (up to 2016) in journals, book chapters, and conference proceedings that examined speech acts in a cross-sectional design. I used the following search terms in combination: ‘pragmatics,’ ‘speech acts,’ ‘proficiency,’ and ‘cross-sectional.’ I also searched review articles and handbooks for relevant studies. This search process came up with 153 studies. Then, I analyzed each study based on these eligibility criteria:

1. The study was a data-driven empirical study that assessed production of speech acts among L2 learners in different proficiency groups.
2. The study had sufficient information about the instrument (with sample items) or data collection methods used to assess speech acts.
3. The study had a clear and objective basis on which proficiency was determined (e.g. standardized exam scores, course/grade levels, and/or length of formal study).
4. Studies that used a correlational or regression analysis between proficiency levels and speech act performance were excluded.

This screening process resulted in 28 primary studies (marked with ** in the references). I proceeded to code each study for target language, sample size, participants’ L1s, proficiency measures, speech act types, instrument characteristics, and findings (differences between more and less proficient groups). Several studies employed a mixed cross-sectional and longitudinal design (Al-Gahtani & Roever, 2015; Li, 2014). In those studies, the cross-sectional parts of the findings were used in this review.

Several studies examined multiple types of speech acts. Requests outnumbered all other speech acts, yielding 18 studies in total, followed by apologies (7 studies), refusals (6 studies), compliments (3 studies), and other expressive speech acts (e.g. giving opinions, disagreement). The majority of studies examined L2 English
(15 out of 28 studies). There are three studies in L2 Arabic, three in Chinese, two in Spanish, two in Greek, and one each in German, Russian, and French. Most studies used course levels and age to operationalize proficiency.

**Characteristics of advanced competence in speech act production**

Due to the space limit, I will only discuss the speech act of request for the characteristics of more advanced learners’ performance because most cross-sectional studies are focused on this speech act. I will synthesize findings in two areas of speech act analysis: linguistic strategies for formulating a request and a sequential organization of a request.

*Linguistic strategies of requests* Of 18 studies on requests, 16 studies used either a discourse completion test (DCT; written or spoken) or a role-play task in which participants are presented with a situational scenario and produce the request in one turn or act out the request scenario with an interlocutor. Two studies examined institutional settings in which learners interacted with their instructors during office hours and their requests were recorded (Al-Gahtani & Roever, 2014a, 2014b). All studies except three (Al-Gahtani & Roever, 2012, 2014a, 2014b) used a coding framework adapted from Blum-Kulka, House, and Kasper (1989) to categorize request-making strategies and compare frequency of different strategies across L2 groups. The coding framework helps us categorize request head acts according to three degrees of directness: direct (e.g. imperative forms), conventional indirect (e.g. query preparatory questions with ‘can you’ and ‘may I’), and non-conventional indirect strategies (e.g. hinting). In addition, the framework provides a classification for two types of modification: (i) external modification (supportive moves such as preparing the listener for the upcoming request and giving a reason for a request), and (ii) internal modification (syntactic and lexical devices such as conditional forms or politeness markers that mitigate the potential face threat).

Comparison across studies has revealed a remarkable consistency in the differences between more and less proficient learners’ speech acts. Irrespective of target languages, L1 backgrounds, instrument types, and contexts of learning, more proficient learners’ requests are always marked by a high frequency of conventional indirect strategies, accompanied by a variety of external modification devices, although the use of internal modification (e.g. syntactic and lexical mitigations) is not consistent in their requests. In contrast, less proficient learners strongly favor direct request strategies with a minimum use of external modification. They also overuse one type of modification. These are consistent findings from all 15 studies that used Blum-Kulka et al.’s coding framework, with one exception (Pinto, 2005).1

This general trend is illustrated in Bella’s (2012) study. She investigated requests by L2 learners of Greek at three proficiency levels as assigned by an institutional placement test oriented toward the Common European Framework of Reference
for Languages (Council of Europe, 2001): lower-intermediate (around B1), intermediate (around B2), and advanced (around C1). There were 50 participants in each group, in addition to 50 native speakers of Greek. She used a DCT including situations of varying power and social distance: asking a roommate to clean the kitchen (−Power, −Distance), asking a friend to pay for your meal (−Power, −Distance), and asking a professor for an assignment extension (+Power, +Distance).

Results showed that the lower-intermediate group used direct strategies most frequently (59%). In contrast, the intermediate and advanced learners used conventional indirect strategies most often, at 58% and 74%, respectively, with the latter approximating the native speakers’ frequency of 70%. The lower-intermediate group also overused non-conventional indirect strategies (hinting), employing them almost 10 times more than other groups. The intermediate and advanced groups used twice as many external modifications as the lower-intermediate group in a wider range. While the lower-intermediate group used ‘grounds’ (giving a reason for a request) and ‘promise of reward’ strategies at about equal frequency, more advanced groups also incorporated different strategies, such as ‘preparator’ (preparing the listener for the upcoming request), ‘imposition minimizer’ (e.g. I will definitely pay you back tomorrow), and ‘considerator’ (e.g. if you have money). In terms of internal modifications, there was a tendency for the high-intermediate and advanced groups to use more ‘downtoner’ strategies (e.g. just, perhaps), while the low-intermediate group overused the politeness marker ‘please’ in all situations. These two more advanced groups also used more syntactic mitigations, particularly conditional clauses, while the frequency of syntactic modification was less than half in the lower-intermediate group.

These patterns (conventional indirect strategies and a greater range of external modification in higher-proficiency learners) were found repeatedly in other studies, including Rose’s (2000, 2009) studies on English requests by L1 Cantonese learners in primary and secondary schools, Félix-Brasdefer’s (2007) study on Spanish requests by American learners in a US university, and Taguchi’s (2006, 2011b) studies on English requests by L1 Japanese learners in an immersion setting. In Taguchi’s studies, higher-proficiency learners (TOEFL scores of 550 or above) were also more fluent when producing speech acts (faster speech rate), indicating that efficient processing is a characteristic of more advanced learners’ speech acts.

Another indication of more advanced speech act production is the emergence of complex syntactic structures for mitigating request intentions. Although not frequent, Rose (2009) found that only the highest proficiency group (upper-level secondary school students) were able to produce a gerund complement structure (e.g. ‘Would you mind’ + a gerund). In Savic’s (2015) study on Norwegian learners of L2 English, the sixth graders in a primary school used a complex, bi-clausal structure (e.g. ‘Do you have … that I can borrow?’, ‘If’ + clause), while these forms were absent from the second and fourth graders’ productions. These findings add to the generalization that more advanced learners’ speech acts are marked by pragmalinguistic sophistication, as seen in the use of a variety of complex and compound structures, supportive moves, and syntactic/lexical mitigations.
These generalizations are understandable considering that higher-proficiency learners are likely to have a large repertoire of linguistic resources to draw on. The type of tasks used to examine speech acts also add to the generalizations. Most studies used production measures such as DCT and role-play. Unlike a comprehension task, a production task requires skilled linguistic processing. Although advanced linguistic competence does not automatically lead to sophisticated speech acts, learners’ lexis and morphosyntax must be accurate so the illocution conveyed through the forms is understood correctly. For instance, learners need to know that lexis such as ‘just’ and ‘perhaps’ work as downgraders and mitigate the tone of a request. They also need to know that non-conventional forms with modals or complement clauses (e.g. ‘Could you’ + verb, ‘Would you mind’ + verb) could soften the request intention, as opposed to more direct forms such as ‘please’ + verb. Incorrect linguistic forms, as seen in wrong word order or word choice, may obscure meaning or lead to misunderstanding. Because of this precise linguistic processing involved in production, speech acts are constrained by learners’ grammar. Higher-proficiency learners, who are likely to possess higher-level grammatical competence, have an advantage because of their greater availability of vocabulary and grammatical forms, as well as their greater processing capacity for accessing these forms. Processing capacity of component knowledge (grammar and lexis) is particularly helpful as it leaves more processing space for fine-tuning utterances and producing internal and external modifications, which lower-proficiency learners find especially difficult.

Although differences in pragmalinguistics across proficiency levels are clear, there is little evidence of a proficiency impact on the sociopragmatic aspect of speech acts. In some cases, higher-proficiency learners showed sociopragmatic sensitivity by using more supportive moves in a situation involving unequal power relationship and large social distance (e.g. asking a teacher to extend the assignment deadline), but variation across situations was generally negligible. A notable exception is Rose’s (2009) study, in which the upper-grade L2 English learners produced the ‘alerter’ (e.g. ‘excuse me’) three times more in a situation involving a request to a teacher compared with a situation involving a peer. However, this type of large discrepancy across situations was rare. Similarly, analysis of request head acts revealed little variation across situations. The more advanced group used conventional indirect strategies with modals (e.g. ‘Could you’ + verb) regardless of situation type. They did not use other indirect forms (e.g. bi-clausal structures such as ‘I wonder if’ + verb) or internal modification (e.g. conditional clause). These findings indicate that the ability to diverge linguistic means corresponding to different contextual specifics is an indicator of advanced speech act production, which is not yet revealed in existing studies. A study that involves a typically advanced group of learners (e.g. graduate students in the target-language country) may reveal such sociopragmatic sensitivity and varied pragmalinguistic repertoires for implementing the sensitivity.

Sequential organization of requests

Higher-proficiency learners’ use of external modification was also found in a few studies that applied Conversation
Analysis (CA) to examine requests in a situated interaction (Al-Gahtani & Roever, 2012, 2014a, 2014b, 2015). Unlike other studies that used a pre-existing coding taxonomy of speech act strategies, these studies took the emic approach, revealing how participants co-construct a speech act sequentially turn-by-turn. Hence, the focus of CA-based analysis on external modification is not about the presence or absence of the modification itself. Rather, the analysis attends to the location of modification devices, which reveals how these devices are contingent upon the interlocutor’s reactions and are embedded in the sequential organization of talk. Existing studies found that more proficient learners were capable of responding to the interlocutor’s utterances in a sequentially relevant manner by using appropriate external modification at the right time. Hence, from the CA point of view, more advanced speech acts are characterized as learners’ ability to monitor the progress of unfolding discourse and make appropriate contributions so they can jointly accomplish the act with their interlocutors.

Al-Gahtani and Roever’s (2012) study illustrates this characteristic of advanced learners. They investigated how 26 Saudi ESL learners co-produced request sequences with their interlocutor in English. Participants were recruited from four proficiency levels in an ESL program in Australia: beginning, lower-intermediate, upper-intermediate, and advanced, based on their course levels, cloze test scores, and self-reported International English Language Testing System (IELTS) scores. They completed three role-play tasks involving request-making scenarios with varying power relationships between interlocutors (e.g., asking a housemate to buy bread, asking a professor for lecture notes). A marked difference between proficiency levels was found in the use of pre-expansions, namely sequences preceding the request, including moves such as ‘preliminaries to preliminaries’ (e.g., ‘I have a favor to ask you’) (Schegloff, 2007). Data showed that upper-level learners used pre-expansions frequently, whereas lower-level learners proceeded straight to the request without any pre-expansions. See the excerpt below from an upper-intermediate participant for illustration (adapted from Al-Gahtani & Roever, 2012, pp. 51–52). This is from a role-play situation of asking a roommate to get some bread.

1. P: hi ((name))
2. I: hi ((name))
3. P: hhh> actually < I wanna ask you something?
4. I: $su::re.$
5. P: $rightarrow$ .hhh today I have too many (. ) assignments to do=
6. I: =Yeah
7. P: $rightarrow$ ↑so I have no:: more time (.1) to do my shopp[ing
8. I: [.
hh
9. P: $rightarrow$ for today (. ) a::nd I’m running out (. ) the bread so could you (.3) buy
10. some bread for me?
11. I: su::re (. ) yeah (. ) but> you know < right now I’m wa::nting this match so (. )
12. do you wa::nt it at the moment (. ) or:: I can buy it later on?
Al-Gahtani and Roever explain that this learner first engages in a greeting sequence with the interlocutor (lines 1–2). Then he produces the pre-expansion in line 3 to check the interlocutor’s availability for the request. In the following turn (line 4), the interlocutor provides a ‘go-ahead’ cue (‘Sure’). Acknowledging this, in the next few lines, the learner produces another pre-expansion, i.e. giving a reason for his request. The interlocutor’s response token in line 6 (e.g. ‘Yeah’) is another ‘go-ahead’ cue. As we can see in this excerpt, the learner’s pre-expansions occur in synergy with the interlocutor’s response tokens. By attending to these cues coming from the interlocutor, the learner evaluates the odds of his request being accepted. Observing that the chance is high, he proceeds to the request using the conventional indirect form (‘could you’) in lines 9–10. Hence, the learner’s request is co-constructed with the interlocutor in a conversational sequence over multiple turns.

The increased use of pre-expansions in higher-proficiency groups was found in other studies on L2 Arabic learners in an intensive Arabic program (Al-Gahtani & Roever, 2014b, 2015). The use of preliminary moves to request increased according to proficiency levels: 21% frequency for beginners, 46% for low-intermediate, 78% for high-intermediate, and 96% for advanced learners.

Al-Gahtani and Roever interpreted the use of pre-expansions as learners’ ability to organize their requests as dispreferred social actions. Some actions are designed as preferred (e.g. agreements), whereas others are designed as dispreferred (e.g. disagreements; Pomerantz, 1984; Sacks, 1992; Schegloff, 2007). A request is a dispreferred action because a certain degree of imposition is put on the listener. Hence, a request is often marked by hesitations, delays, mitigations, and various preliminary moves. Pre-expansions (e.g. providing explanations, checking for availability) are a type of preliminary move that preface a dispreferred act of request. Proficient learners’ use of pre-expansions reflects their orientation toward the dispreferred nature of a request. They are capable of arranging their preceding moves through pre-expansions according to this preference structure. By doing so, they can allow the requestee to signal an upcoming rejection indirectly, or in the case of acceptance, the requestee can pre-empt the request with an offer without waiting for the request to be verbalized.

Advanced learners’ attentiveness to the dispreferred action was also found in other cross-sectional studies dealing with different speech acts (i.e. disagreements; Dippold, 2011; Pekarek Doehler & Pochon-Berger, 2011). In Dippold’s study, L2 learners of German in their first, second, and final year of university study in the United Kingdom engaged in a peer discussion on a controversial topic. Upper-level learners showed a tendency to delay their disagreement or provide a partial agreement before presenting an argument. Pekarek Doehler and Pochon-Berger (2011) compared disagreements between the fourth- and eighth-year L2 French learners in a naturalistic classroom setting. The upper-level group typically delayed their disagreements by using a token agreement (e.g. ‘Yes, but …’) and avoided the initial-turn disagreement.

In addition to pre-expansions, higher-proficiency learners also showed a tendency to use longer post-expansions to organize a dispreferred action (Al-Gahtani & Roever, 2014a, 2015). Post-expansions are sequences that follow the core
adjacency pair (e.g. question-answer, request-acceptance/refusal). When the second part of the pair is a dispreferred action (e.g. a refusal after a request), post-expansions can take a long sequence (e.g. asking for explanations or negotiating further). Al-Gahtani and Roever (2014a) found that longer post-expansions in requests only occurred with learners at the highest proficiency level. Even when the interlocutor’s response to a request was non-committal, lower-proficiency learners typically moved to a closing sequence of a conversation without using even a minimal form of post-expansion (e.g. acknowledging the acceptance of the request). This was in stark contrast to the most advanced learners, who provided extended post-expansions by elaborating on the reason for the request or insisting on their request. Learners’ proficiency also impacted their interlocutor’s reactions. High-level learners’ requests were followed by the interlocutor’s lengthy post-expansions, indicating that their requests were questioned and negotiated by the interlocutor. In contrast, lower-level learners’ requests were often accepted promptly because the interlocutor saw that the learner’s ability to engage in extended conversation was limited.

These analyses further confirm that proficient learners are able to engage in an extended discourse and take control of the direction of their request by implementing the preference structure skillfully. Pre- and post-expansions embedded in a sequential organization of talk are the characteristics of more advanced-level requests in interaction. Unlike studies in the previous section that focused on request-making strategies in isolation, studies discussed here show how a request is collaboratively constructed between learners and their interlocutors as both parties attend to mutual goals. Studies in this section illustrate that linguistic strategies do not occur in isolation. They are contingent upon the unfolding of discourse and occur progressively corresponding to the interlocutor’s reactions turn-by-turn. Hence, advanced proficiency in making an appropriate request (or other speech acts) is part of learners’ interactional competence—the ability to use a variety of linguistic and interactional resources skillfully for the joint construction of a communicative act (Young, 2011). Knowledge of speech act strategies, supportive moves, and internal modifications is part of L2 learners’ linguistic resources, but learners also need interactional resources to implement such knowledge at the appropriate time in a real-time interaction. These interactional resources are an area that can discriminate between less and more advanced learners, as shown in the studies here. More advanced speech act performance is most revealed in learners’ ability to co-construct a speech act in a harmonious manner with other participants.

Summary

This section has focused on the production of requests and reviewed cross-sectional studies involving L2 groups of different proficiency. There was a consistency with existing findings that, as their proficiency level increases, learners shift from direct to conventional indirect strategies involving various modal forms, with the emergence of longer, more complex forms such as bi-clausal or complement structures.
Higher-proficiency learners also use a wide variety of modifications in and around the request head act. While these are all at the utterance level of analysis, discourse-level analysis showed that more advanced learners have the ability to organize their requests as dispreferred social actions by using pre-expansions (e.g. preliminaries such as explanations and availability checks), as well as post-expansions to account for negative reactions coming from the interlocutor. Combining these two types of analysis, we can conclude that more advanced learners can produce speech acts that are linguistically elaborate and furnished with a range of supportive moves and mitigation devices. They can also use these resources in contingency turn-by-turn, depending on the direction of real-time interaction. With greater linguistic knowledge and processing efficiency, more proficient learners are able to produce complex forms and supplement them with external modification with ease, as well as control the progress of their requests in interaction.

While the pragmalinguistics of advanced-level requests is rather clear, less is known about higher-proficiency learners’ sociopragmatic knowledge, namely their understanding of social relationships, cultural conventions, and norms of interaction that affect speech acts. Variations in linguistic strategies across different social situations found in several studies indicate that some advanced-level learners show their incipient sociopragmatic sensitivity, but clearly more cross-sectional research is needed to reveal advanced sociopragmatic knowledge.

**Conclusion and future directions**

I have presented a review of 11 cross-sectional studies on comprehension of indirect meaning and 17 studies on production of requests. My review was guided by the question: What do the studies tell us about the characteristics of more advanced learners’ comprehension and production of pragmatic meaning? After analyzing study findings for commonalities and discrepancies, several generalizations emerged:

1. More advanced learners (compared with their less advanced counterparts) are able to comprehend a wider range of indirect utterances that involve greater distance between the propositional meaning and intended meaning.
2. More advanced learners are able to comprehend meaning without being distracted by linguistic and non-linguistic cues that add to the implicitness of information (e.g. unclear agency, incomplete sentences, and hesitation markers).
3. More advanced learners go beyond the one-to-one mapping between form and force and use a greater range of linguistic repertoire to construct the speech act of request. Their linguistic repertoire involves forms that are syntactically and lexically complex.
4. More advanced learners are sensitive to the dispreferred nature of a request and construct the request by using various discourse moves in a sequentially appropriate manner to mitigate the dispreferred act.
After surveying the literature I see several directions for future research. First, instruments and tasks used in the previous studies need to move beyond the level of DCT and multiple-choice questions and instead incorporate tasks that elicit more advanced-level skills. For example, authentic recordings of conversations or audio-visual input can be used to assess advanced learners’ implicature comprehension. In addition to comprehension of implicature, researchers can be creative in designing a task that can assess learners’ production of implicature or disambiguation of implied meaning in interaction.

Second, although we have a number of cross-sectional studies comparing more and less advanced learners, the actual ‘advancedness’ of pragmatic competence has been underexplored. This is because the more advanced learners in these studies are restricted to learners who have three to four years of formal language study in a university setting. These learners are not bilinguals or superior-level expert speakers. Because of this, the level of pragmatic competence targeted in these studies is limited to intermediate-level pragmatic abilities, which are also representative in the current review. As a result, we do not know what kind of pragmatic abilities superior-level L2 speakers possess. Obviously one future direction is to include truly advanced learners (e.g. graduate students or working professionals in a target-language country) in the cross-sectional design and compare their performance with undergraduate-level students. Based on the findings in this review, we can design a task tapping into the aspects of pragmatics that upper-level undergraduate students struggle with, for example internal modification (syntactic and lexical mitigations) and sociopragmatic variations. We can examine whether superior-level learners are able to deal with these aspects with ease, or what types of challenges still remain in their development of pragmatic competence.

We can also operationalize advanced pragmatic competence differently by moving away from the popular constructs of implicature and speech acts and instead incorporating pragmatic features that are effortful even for adult native speakers to master. There are several such features, including honorifics (Brown, 2011; Ikeda, 2009), style shifting (Taguchi, 2015), argumentative discourse (Dippold, 2011), job interview skills (Louw, Derwing, & Abbott, 2010), and comprehension and production of humor (Bell & Attardo, 2010). For example, honorifics systems in Japanese and Korean are complex linguistically and socioculturally because they are double-layered. The speaker has to have accurate knowledge of honorifics and social situations in which honorifics are called for. At the same time, they have to know that honorifics use is variable and subjective because they project the speaker’s desired identity (Okamoto, 2011). Honorifics are a reflection of the social self that the speaker wants to project—how formal or informal he/she wants to appear in a particular context. Because of this subjectivity, there is immense situational and individual variation in the use of honorifics. When there are no agreed-upon norms or when norms vary across situations, how do expert speakers decide what forms to use when and to what extent? The ability to navigate contextual dynamics and adapt to situations with appropriate linguistic choices is certainly an aspect of advanced pragmatic competence worth investigating in the future.
**NOTE**

1 This generalization needs to be interpreted with caution because different languages have different norms in terms of the appropriate directness levels of requests. Al-Gahtani and Roever’s (2015) study on L2 Arabic showed that higher-proficiency learners used more conventional indirect strategies than lower-proficiency learners, who favored direct strategies, conforming to the general patterns. However, the use of direct strategies (not conventional indirect strategies) was the shared norm among native Arabic speakers. In fact, when the same learners’ requests were traced longitudinally over time, learners showed a U-shaped pattern where they started out using direct requests, then reduced them and increased their use of indirect requests, and finally reverted to direct requests, which conform to the native-like sociopragmatic norms and cultural expectations. Hence, the shift from more simple direct forms to more complex indirect forms does not apply universally across languages.

**REFERENCES**


**Advanced Rhetoric and Socially Situated Writing**

GREGG FIELDS AND PAUL KEI MATSUDA
Arizona State University

**Introduction**

This chapter addresses rhetoric and writing as two important considerations of advanced second language acquisition (SLA). We will specifically elaborate on how these two considerations influence the development of socially situated language acquisition. Brief overviews will be provided for both, and existing models will be explored to affirm the unique contributions of rhetoric and writing to a socially situated understanding of advanced SLA as they shift the focus of acquisition from target-oriented competence, accuracy, and fluency constructs to an efficacy-oriented multiple knowledges framework.

**A brief overview of rhetoric**

Rhetoric has a very long history in western thought. However, in second language studies, the term *rhetoric* is often associated with scholarly work on contrastive rhetoric (Connor, 1996; Connor & Kaplan, 1987; Kaplan, 1966) or intercultural rhetoric (Belcher & Nelson, 2013; Connor, 2011) from the second half of the twentieth century, and these approaches to language and writing development influenced multiple generations of scholars and teachers who hoped to better understand and support the development of effective writing and rhetorical practices. Over the past decade, publications have begun to wain in this area as researchers challenge the approach for needing to address process (Liebman, 1992), culture (Atkinson, 2004), genre (Scollon, 1997), and writerly intention more directly and with a less Anglocentric model (Kubota & Lehner, 2004). Connor (2011) and others have
maintained the value of contrastive and intercultural rhetoric while also trying to address some of these concerns and shift the conversation toward the latter. More recently, Belcher (2014) has written a substantive review, “What we need and don’t need intercultural rhetoric for: A retrospective and prospective look at an evolving research area.” While contrastive and intercultural rhetoric have been significant in developing an understanding of language development through writing, these more recent forms of rhetorical study tend to focus primarily on developing appropriate argumentation or arrangement (Liebman, 1992), the underlying structure of a single piece of writing, by comparing different structures across languages often with native-speaker English as the target, though intercultural rhetoric is an attempt to shift away from this Anglocentricity.

However, as one part of the classical western trivium of higher education (Leki & Silva, 2004, p. 3; Liebman, 1992), rhetoric has much more depth and breadth than the structure of an utterance or writing alone. Classical definitions describe rhetoric in an Aristotelian form as the “art [of] discern[ing] the real and the apparent means of persuasion” (Aristotle, 2004, 1355b). Rhetoric in its oldest western tradition is a practical work of discovery that is based on situated argumentation adapted to specific occasions or exigencies (Bitzer, 1968). In the Greek tradition, rhetoric was primarily associated with spoken civic discourse. As it has developed over millennia, rhetoric has been used to describe oral, written, visual, and multimodal forms of communication and language use in a variety of contexts. No matter which mode of communication, rhetorically oriented advanced language users address the context for a piece of communication and break it down into its multiple rhetorical components—audience, topic (scope), communicative purpose or intent, exigency, genre, constraints/limitations, cultural context, and design—some of which will be discussed in more depth throughout this chapter.

Further, it is not uncommon to associate studies of rhetoric with discourse, but the relationship between these two can get a bit muddy depending on one’s disciplinary background. Applied linguists may focus on any number of micro or macro discourse considerations in a variety of contexts; and because both discourse and rhetoric deal with language in use, they can possess heavily overlapping definitions and functions in some contexts. Within this chapter and to draw a soft border to distinguish one from the other, while discourse is often given the “important role of shaping reality [or] creating patterns of understanding” (Mayr, 2015, p. 756), rhetoric is the leveraging of these patterns of understanding to accomplish a socially impactful communicative purpose.

**A brief overview of writing**

Writing—the characteristics and uses of written discourse as well as its production processes and strategies—became an important area of teaching and research in the mid-twentieth century as the need for advanced literacy for academic and professional contexts became clear. The field of writing studies, also known as rhetoric and composition or composition studies, began to emerge in North American higher education in the 1960s where the ubiquitous first-year writing
requirements and the increasing emphasis on writing as a mode of knowledge production and dissemination made the need for advanced knowledge of writing and the teaching of writing clear. Research on writing encompassed a wide range of issues—including both the process and the product of writing, as well as various issues in teaching and learning writing. More recently, with the advancement of technology, writing came to be conceived more broadly to include not only orthographic representation of ideas but also multimodal resources including audio and visual elements (Knoblauch & Matsuda, 2008).

In second language studies, research on writing began in the 1960s as the college writing requirement made it abundantly clear that oral proficiency and the knowledge of orthography was not sufficient preparation for college-level writing, as it was previously supposed. In addition to the studies of written discourse in the applied linguistic tradition, the development of L2 writing research was heavily influenced by the development in writing studies in general. Around 1990, research on L2 writing reached a critical mass and, with the publication of Second Language Writing (Kroll, 1990) and the creation of the Journal of Second Language Writing in 1992, L2 writing came to be recognized as an interdisciplinary field of research on L2 writing, writers, and writing instruction (Leki & Silva, 2004; Matsuda, 2003).

In terms of SLA, writing often is linked to alphabetic print considerations like orthography or spelling and second language writing systems (Cook & Bassetti, 2005), and even with this aspect of writing systems present as a related field, writing is often given only cursory attention or even excluded as a major focus from some texts in SLA (Archibald, 2000; Ellis, 2015; Robinson, 2001). Unfortunately, despite the presence of scholarship on second language writing systems in SLA, this more limited attention to writing tends to diminish the importance of advanced writing practices which can display a writer’s rhetorical and social awareness and indicate advanced language efficacy.

Advanced writing has a variety of definitions and seeming contradictions which benefit and complicate the process of SLA. First, being defined as both a process and a product, advanced writing goes beyond the single text or textual artifact. Writing can be considered the process of producing textual artifacts which are situated in a specific sociocultural context and which embody rhetorical as well as linguistic aspects. The individual textual artifacts are written products, but again these are not always purely alphabetic or character-based products. Second, because of this dual process-product existence, writing is something that happens as a process over time, is published as a product into a specific context, and is often read asynchronously in a variety of social contexts and temporal moments. For this reason, one’s context for writing a text is just as important as one’s readers’ contexts while attempting to interpret that text. As a final note, writing, and particularly advanced writing, is a unique form of anthropological development allowing communication to stretch beyond the moment of a single utterance, and therefore, writing as a form of language development is an advanced practice in and of itself. As a result of the process, product, and practice of writing in socially situated contexts, in this chapter, we will elaborate on the unique affordances of writing and rhetoric for enhancing advanced SLA.
Complexity, accuracy, and fluency constructs

To begin highlighting the unique affordances that rhetoric and writing bring to advanced SLA, we might examine how they overlap and extend beyond popular approaches to SLA that emphasize complexity, accuracy, and fluency (CAF) constructs. As a point of orientation, in 2009, a special issue of *Applied Linguistics* highlighted the CAF constructs prevalent for approaching language acquisition. According to the editors (Housen & Kuiken, 2009), “Complexity has … been commonly characterized as ‘[t]he extent to which the language produced in performing a task is elaborate and varied’ (Ellis 2003, p. 340), accuracy as the ability to produce error-free speech, and fluency as the ability to process the L2 with ‘native-like rapidity’ (Lennon 1990, p. 390).” Tracing the historical lineage of CAF, authors in the special issue provide substantial background for these SLA constructs and emphasize areas where further developments are needed. As one author notes (Larsen-Freeman, 2009), the CAF constructs are helpful for examining performance but not proficiency because they only measure the product. Measuring language acquisition through a product orientation in this way can provide valuable information about the capabilities of a speaker *under significant time constraints*, i.e. where communication must be constructed immediately in the case of speech or perhaps live chat. Focusing on product may also be beneficial for providing immediate situated feedback through teacher- or peer-correction strategies like repair (Lyster, Saito, & Sato, 2013). When time constraints are more relaxed and provide greater allowance for revision—i.e. a writing task crafted over multiple weeks—focusing solely on product can negate process-based proficiency measures which become more apparent through a rhetorically based model.

Further, the CAF constructs can be limited and even problematic at times by focusing on microstructures and idealized standards. First, while effective for providing measurement of word- or sentence-level languaging, the CAF constructs are not as effective for addressing larger macro structuring of a text or the process of developing that text, and these constructs have little to no ability to gauge situated rhetorical knowledge based on intention or interpretation. Second, CAF constructs tend to emphasize target-oriented, linear progression: Such an emphasis privileges a particular target, which is often the idealized native speaker of a language. Unfortunately, this ideal can create substantial challenges for language learners by defining advanced learners in connection to a controversial standard of the “sociolinguistic” and “developmental” construct of the native speaker (Davies, 2003, p. 185; Reid, 2008) to the detriment of the learner. In many institutional settings and in target-based acquisition models, second language learners, even advanced ones, are often given incongruous standards to meet for complexity, accuracy, and fluency that negate everyday discourse turns seen even in the practices of native speakers, like self- and other-initiated repairs for example. In these ways, the CAF constructs can present challenges to educators, researchers, and learners by focusing on micro rather than macro structures, minimizing the process for developing a piece of communication, and establishing standards based on an idealized native speaker.
A rhetorically based framework

To address some of these challenges, one can shift emphasis to, or at least supplement CAF with, an efficacy orientation to acquisition borrowed from rhetoric and socially situated writing. By focusing on what is effective in a particular context and for a particular audience, one can shift the target emphasis from the idealized native speaker to the socially situated communicative need. (It is not unlike what World Englishes scholars like C. L. Nelson (2011) describe when addressing “intelligibility, comprehensibility, and interpretability.”) One of the benefits of this sort of rhetorically based model is that it is does not fully conflict with the CAF constructs; instead, a rhetorically based model can shift the emphasis or the focus of learners, teachers, and researchers to the situated contextually relevant language need. A rhetorical orientation to language acquisition makes the target realistic and attainable by connecting to a specific rhetorical situation, audience, and context.

As one rhetorically based option, one might borrow from Christine Tardy’s articulation of a four-knowledges framework (2009). Tardy’s study focuses on the development of genre awareness in a cohort of international graduate students in the United States across multiple semesters. Within her manuscript, she elaborates on a rhetorically based four-knowledge framework which can help learners, teachers, and researchers to better understand, on the one hand, what advanced acquisition might look like if conceptualized more broadly and, on the other, what advanced writers know or do in socially situated contexts. Tardy’s framework provides an effective way to define advanced writing by focusing on the learner’s development of not only formal knowledge addressed by CAF but also process knowledge, rhetorical knowledge, and subject-matter knowledge. Tardy explains her framework as “a confluence of these four dimensions” (p. 20) and elaborates on each of these four knowledge categories in the following ways:

- **Formal knowledge**—“the more structural elements ... prototypical form(s), discourse or lexicogrammatical conventions ... common ... content or structural moves, and the various modes and media” of communication as well as the linguistic code. (p. 21)
- **Process knowledge**—“the procedural practices ...” processes of writers, processes of distribution, and processes of the audience for consuming the text, as well as the larger network of writing “chains, sets, or systems.” (p. 21)
- **Rhetorical knowledge**—“an understanding of the ... intended purposes ... within a local context” or system and an “awareness of the dynamics of persuasion within a socio-rhetorical context,” as well as one’s “own position[ality]” within that context. (p. 21)
- **Subject matter knowledge**—“knowledge of the relevant[, context-specific] content” as perceived by readers within that context. (p. 22)

By addressing formal knowledge which includes CAF and other formal knowledge as well as the other knowledge dimensions, this rhetorically based framework provides a way to understand how one comes to write or communicate in a
rhetorically advanced way. Because this framework is multidimensional, it allows one to view advancedness either holistically, based on the overlap of the four types of knowledge, or compartmentally by focusing on individual types of knowledge. Thus, one might draw from her framework that a writer can be defined as advanced if he or she is advanced in at least one of the knowledge areas above, expanding the definition of what an advanced writer might be perceived to be. Still, as writers become more advanced, Tardy finds that these knowledges “become increasingly integrated with growing expertise” and a holistic advancedness develops (p. 22). For the most advanced writers, not only has a certain level of expertise developed based on these overlapping knowledges, but the writer is also capable of enacting this knowledge in practicable and effective ways based on his or her awareness of the situated context.

Using the four-knowledges framework allows for a focus on what types of knowledge might be necessary for effective communication in a situated context. This shift of emphasis moves the measure of acquisition away from the imagined universal target of the native speaker toward a contextualized target of a rhetorically situated writer. Even if focusing solely on the textual product, the writer may be seen as advanced if he or she succeeds in accomplishing a rhetorical purpose within the specific context by enacting at least one of the four knowledges in an advanced way, not simply depending on whether he or she has native-like or flawless grammatical accuracy or native speaker fluency; in certain contexts, grammatical accuracy may even hinder rhetorical efficacy. And through the inclusion of process knowledge, this model also invigorates the discussion of advanced acquisition from a proficiency standpoint which focuses attention on how an utterance or text is being produced and not just what has been produced.

**Progressively scaffolded models of rhetorical knowledge**

Modern models for defining rhetorical knowledge have been developed and expanded on substantially by rhetoric and writing scholars throughout the last half of the twentieth century. Effective models of the rhetorical situation have been formed, from simpler models like the communication triangle to more complex mappings of rhetorical situation, exigency (Bitzer, 1968), audience (Ede & Lunsford, 1984; Ong, 1975), and stakeholders, to substantially complex ecological rhetorical models (Cooper & Holzman, 1989; Edbauer, 2005) which view interlocutors and their utterances or texts as socially, temporally, and spatially situated as agentive organisms in interactive communicative ecosystems (Kramsch, 2003; van Lier, 1997, 2004, 2010). By working from the simpler to the more complex models in this section, we develop a scaffolded structure into which learners or teachers can incorporate the various rhetorical components, while also providing some background into the development of the various models to aid researchers.
To begin with the simpler model, the communication triangle establishes four seemingly straightforward components of a rhetorical situation—sender, message, receiver, and context—which rhetoricians have adapted to writing situations by shifting these labels to writer, text, reader/audience, and context. These components of the rhetorical situation establish a minimum level of socially situated rhetorical knowledge, and developing an understanding of the relationships between these components is an effective first step toward advanced knowledge. We can take a key principle of understanding from each of the four component parts. First, pieces of communication in all modalities have creators, senders, or writers: Acknowledging this point requires that we recognize the influences upon the creator of the text. Writers do not develop texts in a vacuum. They have unique social positionings and experiences that influence the messages that they present. Second, messages have underlying rhetorical purposes; as we stated previously, rhetoric is the leveraging of patterns of understanding to accomplish a socially impactful communicative purpose. Because communicative messages have an underlying purpose, perception of one’s own purpose when writing and of another’s purpose when responding is a key indicator of advanced rhetorical knowledge. Third, messages in communicative situations always have an audience—a listener, a reader, or a viewer. Having an audience requires communicating in a way that is linguistically accessible to that audience as well as rhetorically effective for that audience. Fourth, each of these first three components exist within a context. The context may affect how the writer writes, how the message is sent, or how the audience receives the message. Context may impact the sender or receiver physically, emotionally, mentally, or socially. In this way, the communicative triangle acts as a teaching tool while it also gives insight into the sender’s situatedness, the message’s purpose, the audience’s reception, and the impact of the larger context on the rhetorical situation, and we will expand on each of these areas of insight in a more focused way later in the chapter. However, for those with more advanced rhetorical knowledge, this simpler model acts as more of a stepping stone to progressively complex models.

In communication scholarship, more elaborate models have been developed and can be used by teachers or language learners to create a complex and nuanced awareness of the rhetorical situation, aiding writers in acquiring deeper rhetorical knowledge. For instance, the interactive and transactional models of communication acknowledge the possibility of dissonance as well as feedback loops within systems of communication, indicating that rhetorically oriented language learners would not only be aware of the existence of a complex audience or a situated context, but they would incorporate processes which allow them to anticipate potential dissonance and shift the discourse to assuage that potential dissonance. These processes might incorporate various feedback loops that are common in instances where time is less of a constraint. For example, a student might incorporate feedback loops when they develop final written products through a process that involves instructor and peer feedback, or writers’ processes may involve time to research key concerns or semiotic triggers to avoid within a specific rhetorical situation when dealing with a particular audience or variety of audiences. These more elaborate communication models begin touching on the interactive
rhetorical process, but other complex models are also available for understanding advanced rhetorical knowledge.

More complex rhetorically oriented models for understanding writing and discourse have been introduced by multiple scholars (Britton, Burgess, Martin, McLeod, & Rosen, 1975; D’Angelo, 1975; Kinneavy, 1971; Moffett, 1968). Each model is “rooted in the semiotic structure of the … communication triangle” (Kinneavy, 1983, p. 123), and each has received varying degrees of uptake among teachers and scholars over the past few decades. These models are seminal for many who study rhetoric and writing, but they have not been disseminated as substantially in language studies more broadly. Understanding some of these more elaborate models can help scholars and teachers more effectively study and develop rhetorical efficacy in language learners. About a decade after introducing his own model, Kinneavy (1983) synthesized these four models of rhetoric and discourse in an attempt to articulate a larger “meta-system,” hoping to show how each model had overlapping common ground for developing advanced rhetorical awareness. He shares that Moffett’s model introduces modes of discourse or “what we discourse about” (p. 129). Britton’s model introduces the aims of discourse or “why we discourse” (p. 129), and this connects to a language user’s purpose behind a particular message. He also shares that both Moffett and Britton overlap in trying to “determine the kinds of audiences addressed in discourse” (p. 129). Adding to these points, while in alignment with the others but not fully synoptic, D’Angelo’s model (p. 130) elaborates on the various modes of discourse established by Moffett and holds indirect links to various parts of the other models. As a final note, Kinneavy’s earlier model provides an articulation not only of the semiotics of the triangle, but expands to include semantic and pragmatic aspects that flesh out the other scholars’ works. Without providing much further detail on the models themselves, it is important to note that these sorts of rhetorical models can act to enhance a writer’s ability to analyze and understand particular rhetorical situations, and by developing this meta-awareness of the particular situation, an advanced writer can better address situated needs through a well-crafted product developed through an analytical process over time. Further, as rhetorical models, they can be more broadly applied to various modalities of communication beyond writing as well. Suffice to say, these sometimes complex models of the rhetorical situation can be drawn upon as part of an advanced writer’s process and act as a background to guide textual and rhetorical decisions.

And while, for many rhetoric and writing scholars and teachers, the communication triangle or the aims and modes of discourse are the bread and butter of their classrooms or professional contexts, a third track of rhetorical awareness is also often effective for developing advanced rhetorical knowledge. This third trajectory for developing rhetorical knowledge builds on three key pieces of scholarship which we will detail below: Bitzer’s (1968) “Rhetorical situation,” Cooper’s (1986) “Ecologies of writing,” and Edbauer’s (2005) “Unframing models of public distribution: From rhetorical situation to rhetorical ecologies.” Together, these scholars establish an increasingly complex ecological model which views interlocutors and their utterances in speech or text as socially, temporally, and spatially situated as
agentive organisms in a communicative ecosystem. In educational linguistics, this idea of a language ecology has also been attended to substantially (Kramsch, 2003; van Lier, 2004, 2010). From a language acquisition perspective, this ecological model can help language learners position themselves effectively within an environment through intentional responses to authentic situations. Bitzer wades into this approach by initially focusing on the

rhetorical situation as a natural context of persons, events, objects, relations, and an exigence which strongly invites utterance; this invited utterance participates naturally in the situation, is in many instances necessary to the completion of situational activity, and by means of its participation with situation obtains its meaning and its rhetorical character. (p. 5)

The rhetorical situation is introduced here as a complex interaction of not only persons but also events, objects, and relations. Exigency is also established as an invited response. As a key component to a rhetorically based efficacy orientation, exigency is what provides a communicator a specific invitation for communicating at a particular time and in a particular way to a particular audience. Exigency aligns with classroom activities that provide realistic or task-based learning. These activities support learners by providing an authentic invitation to communicate. By responding to these authentic and rhetorically situated invitations, the learners have motivation for more specific, situated, and unique audience-oriented responses. Outside the classroom, exigency applies to language learning on a rhetorical level by reinforcing the need for situated, efficacy-oriented communication because exigency is the motivation or need to speak, write, or communicate in another language with a purpose and into a specific context. It is beneficial as a mechanism for learning, but it is also beneficial as a way of reorienting what counts as advanced. In the rhetorically based model, if learners respond to situated exigencies and if the communicative purpose is accomplished, learners prove themselves advanced by being rhetorically effective through this accomplishment. By showing herself capable of rhetorical efficacy, the learner establishes that she has sufficient rhetorical knowledge to be considered advanced. Thus, defining and recognizing the rhetorical situation and exigency as key to the complex contexts of meaningful interlocution supports well-established pedagogies in second language studies and provides meaningful theoretical anchoring for an efficacy orientation to advanced language acquisition.

Building on the rhetorical situation, Cooper (1986) uses a rhetorically based model to counter a purely cognitive approach to language use by emphasizing not only the rhetorically situated nature of writers and texts, but their situatedness within an ecology. Ecologies of writing treat the writer as an organism capable of not just responding to an invitation but adapting or developing that response through dynamic interaction, receiving iterative feedback from others in the writing ecology (p. 368). Cooper focuses on establishing a real relationship between the audience and the writer as organic, ongoing, and ecological. When writers see their audiences as real co-participants in an ecology, they progress beyond what Piaget
established as an egocentric, less-advanced writing based solely on self-focused cognitive or behavioral approaches. Through this rhetorical model, ecologies of writing establish the development of audience awareness as an organic interactive process leading to the assertion of rhetorical knowledge through the interaction.

Adding depth to the rhetorical situation, writing ecologies, and audience interaction, Edbauer (2005) has expressed concerns that these commonly understood conceptions of the rhetorical situation are still moderately static, limiting writerly awareness. Edbauer has suggested a model of the “rhetorical ecology” that is in full and constant flux which reinforces the messy, flexible, and interactive nature of any communicative situation while also opening up the realm of contextual influences that a writer might need to be aware of during the process of writing. She suggests also that, rather than seeing various forms of traditionally perfunctory steps in a writing process as a means to an end, we should see the steps on the trajectory as things to aim for, produce, and engage with (pp. 21–22). In these steps, not only does feedback with real audiences occur, but advanced writers see the steps of writing along the way as progressive products each documenting a critical communicative and rhetorical moment. By establishing an iterative process of rhetorical development, we move learners beyond the target-oriented mentality because, by acknowledging that the process is messy, the goals of language acquisition change. Possessing the various types of knowledge necessary to be considered advanced is no longer the target, but by engaging in the organic process of developing effective rhetorical knowledge intentionally through iterative steps, one might be considered advanced in this rhetorically based model.

To close up this section on rhetorical knowledge development, some of these models and their respective elements at times may feel more tangible than others, but they each provide valuable insight into rhetorically based and efficacy-oriented models of advanced SLA. The various rhetorical models presented in this section provide an efficacy orientation to language acquisition based on the needs of the specific audiences in specific rhetorical situations. First, simple communication models establish key principles of advanced rhetorical knowledge and writerly awareness. Then, models based on the rhetorical situation and writing ecologies help more clearly situate language learners and writers in real, authentic, communicative situations. These rhetorically based models elaborate on both process knowledge and rhetorical knowledge from Tardy’s framework. For the next portion of this chapter, we will focus more directly on writing and push more deeply into key components mentioned from the various rhetorical models above: the situated writer, the purpose behind the message, approaches to audience, and the impact of context. Beyond these, we will also discuss genre knowledge as a key component to a rhetorically based model of advanced acquisition.

**Rhetorically situating the writer**

As the primary component, writers have unique social positionings and experiences that influence the message that they present. Basically, writers do not develop texts in a vacuum, though less advanced writers may isolate into a cognitive vacuum of sorts, disregarding influences or failing to use existing feedback
mechanisms effectively, as Kesselring and Müller (2011) indicate in their depiction of Piaget’s three phases of egocentrism. However, advanced writers are often aware of influencing factors and develop processes to manage or leverage these influences. On an emotional level, an advanced writer may incorporate into the writing process coping mechanisms for anxiety, stress, or emotional reactions. In this case, being effective may be less connected to formal knowledge acquisition and more connected to a writer’s psychological self-awareness and social awareness, and to an advanced process that provides support or mechanisms for resolving these influences while determining whether they will affect the final written product (Atkinson, 2002). This perspective on the writer links to considerations of a writer’s identity (Lave & Wenger, 1991; Matsuda, 2015), emotions, self-perception (Pavlenko, 2006), and stance (Aull & Lancaster, 2014; Dobbs, 2014; Hyland, 2005; Hyland & Sancho Guinda, 2012). As a result of this advanced process, writers may use various textual features to control their responses.

For instance, in his substantial corpus analysis of dissertations in Hong Kong, Hyland (2004) found that, “while … rhetorical decisions may sometimes reflect either conscious choices or unreflective practices, the analysis of metadiscourse patterns … indicates that effective argument involves a community-oriented deployment of appropriate linguistic resources to represent writers, their texts, and their readers” (p. 148). Here, metadiscursive rhetorical choices are established as ways that writers position themselves within a larger community. As part of the initial coding of this study, Hyland notes nearly a dozen different textual measures of the metadiscourse from these graduate students. The list identifies two categories, “interactive” and “interactional” (p. 139), and emphasizes textual features like transitions, frame markers, endophoric markers, evidentials, and code glosses in relation to the former while presenting hedges, boosters, attitude markers, engagement markers, and self-mentions in connection with the latter. Aull and Lancaster (2014) further support Hyland’s assertion in their examination of stance markers in undergraduate student writing, sharing that “advanced academic writing privileges caution, possibility, and delimited claims over certainty, while incoming FY [first-year] writers tend to use stance features that achieve the opposite kind of stance” (p. 164). Their findings suggest that advanced writers have developed rhetorical practices that use strategic linguistic markers to curb a claim’s intensity and display a writer’s awareness of his or her own socially situated presence in the writing. These textual features display advanced rhetorical choices on the part of a writer. While one might measure the linguistic markers on their own as an indicator of advanced acquisition, the use of these markers as rhetorical tools to accomplish a larger purpose indicates not only advanced formal linguistic knowledge but advanced rhetorical knowledge as well.

**Communicative purpose**

As a second component, messages have underlying rhetorical purposes; as we stated previously, rhetoric is the leveraging of patterns of understanding to accomplish a socially impactful communicative purpose. There are a variety of situated rhetorical purposes, but these often get pared down to three primary options: attempts
to persuade, inform, or entertain. It is often more effective to think of these in an overlapping Venn diagram of varying proportions rather than as mutually exclusive. For example, a children’s book may be simultaneously entertaining through visual design or character hijinks, informative in describing how to brush one’s teeth, and persuasive through a variety of child-oriented logical appeals. Adding complexity to this triad of purposes, one should also consider more complex purposes a writer might pursue: building rapport, establishing credibility, intentionally alienating, misdirecting, foregrounding, backgrounding, and other strategic, socially impactful purposes. Because any message has an underlying purpose, perception of one’s own purpose when writing and perception of another’s purpose when responding are key indicators of advanced rhetorical knowledge.

As a key rhetorical component, the ability of writers to both set out and accomplish rhetorical purposes provides an appropriate measure of rhetorical effectiveness. Advanced writers in this scenario work toward clear rhetorical purposes and also succeed in accomplishing those purposes within the specific context for situated audiences. Using purpose and success to approach advanced writing can allow for more variability in the rhetorical contexts. Writers who successfully accomplish their rhetorical purpose would be considered advanced, because they effectively read the rhetorical situation to know which conventions are needed or helpful to follow and which might be waived. Yet, even if writers accomplish their rhetorical purposes, they may not have needed to do so using conventional methods, standardized grammar, or normalized genre—once again reinforcing an efficacy orientation rather than target-orientation.

**Audience considerations**

As a third component, a writer’s audience—a listener, a reader, a viewer—impacts how the message is crafted and what linguistic knowledge is requisite for rhetorically effective communication. Having an audience requires communicating in a way that is rhetorically effective as well as linguistically accessible to that audience. There are multiple ways to consider this rhetorical component, and they shift on a spectrum from a writer-created fictional audience (Ong, 1975) to an invoked/addressed audience (Ede & Lunsford, 1984) and finally to a real, tangible, immediate audience (Cooper, 1986). This spectrum moves from an awareness of one’s own intentions on the less advanced side to addressing complex audiences on the more advanced side. Below, we include this spectrum of audience understanding, starting with the self-focused perspective (1) and ending with an awareness of real, complex, multifaceted audiences (5).

1. Writer’s awareness of self and one’s own purpose, egocentrism phase 1 (Kesselring & Müller, 2011);
2. Writer’s basic perception of an audience’s existence, egocentrism phase 2;
3. Writer’s understanding of an imagined primary audience (Ong, 1975);
4. Writer’s ability to address specific contextual audiences (Ede & Lunsford, 1984);
5. Writer’s ability to address multiple real specific audiences at once (Cooper, 1986).
Less advanced writers may be able to express themselves and their ideas, but they may do so in a way that audiences or readers perceive as ineffective or inappropriate for the situated context. So even if linguistically advanced or fluent, in a rhetorically based model, they are ineffective. More advanced writers are capable of navigating complex issues while maintaining strategic awareness of multiple audiences. This audience-awareness component supports writers’ development of more situated, complex rhetorical knowledge. As writers shift from an awareness of self and purpose to an awareness of a complex audience and its needs, they can correlate their linguistic and rhetorical development to the conventions, constraints, norms, and requirements of specific situated contexts, if given sufficient time to do so. It may even be easier to focus on developing context-specific formal knowledge or competence for some second language writers, rather than attempting to first acquire the broader sociocultural competence. For example, a writer may prefer to focus on developing an understanding of the IMRD (Introduction, Methods, Results, and Discussion) format for lab reports to match the disciplinary conventions or jargon while intentionally avoiding a pop culture genre or tourist vocabulary. Some might then consider this writer as more rhetorically aware because of agentive choices to prioritize the specialized discourse of a discipline over the ability to socialize more effectively when traveling abroad.

This perception of positionality for the writer might be understood as knowing or even discovering one’s role in an “ecology of writing” (Cooper, 1986):

The perspective of the ecological model offers a salutary correction of vision on the question of audience. By focusing our attention on the real social context of writing, it enables us to see that writers not only analyze or invent audiences, they, more significantly, communicate with and know their audience … Just as the ecological model transforms authors (people who have produced texts) into writers (people engaged in writing), it transforms the abstract “general audience” into real readers. (pp. 371–372)

In developing this awareness of the social situation, the writer ascertains not just the possibility of an audience but also the reality of an audience, a recognition of specific persons or readers included in that audience. Further, by “transforming” one’s sense of self as positioned within a rhetorical ecology of writing, the writer actually begins to shift his or her positionality within that context because developing this awareness displays greater ownership of and, more importantly, membership within a particular social context.

For some writers in certain contexts, advanced rhetorical awareness can be indicated by little more than the writer’s ability to recognize that a receiver, an audience for the written message, exists. The writer in this scenario may still be struggling with lexical limitations or struggling to internalize a standardized grammar, but acknowledging an audience may help a writer in this situation to spend time more effectively focusing on the lexicogrammatical concerns (among others) most salient to a singular reader or audience. In other contexts, the writer may not be considered as possessing advanced rhetorical awareness until he or
she displays clear recognition of and adaptation to a complex audience or readership. As part of a rhetorically based model, audience awareness has particularly salient ramifications because a writer’s awareness of and orientation to his or her audience need not have direct correlation to his or her formal linguistic knowledge. Yet, an awareness of audience may act to focus the writer’s attention on the linguistic norms and conventions of that particular audience, and depending on the time available for accomplishing the task, a writer may be able to use the rhetorical awareness of audience to add agentive purpose to his or her linguistic development.

**Concerning context**

Much of what has been previously discussed in this chapter directly speaks to the context component of a rhetorically based model. The discussion of audiences and readers, the situated role of writers, the ecological social positions that they hold: each of these point to part of the context of a particular piece of communication. Even as a component of rhetoric, there are still a variety of elements that can be put under the umbrella of context which we will only briefly touch upon. The prior conversational context, physical context, and technological context—to name a few—can each impact what is considered rhetorically effective for a particular audience at a particular time.

Prior conversation influences how an audience interprets a piece of writing. More so, in a rhetorically based model, advanced writers must often position themselves within this prior conversation, and to do so, must know the socially situated context of what was previously stated and who stated it. In academia, this is the core purpose of a literature review or survey of the literature (Swales & Feak, 2009), but it is just as relevant in many non-academic contexts as well. Positioning oneself in the context of prior conversation allows the writer to establish credibility with the audience, avoiding unnecessary repetition and keying in to established forms of topical reference. For some conversational contexts, this rhetorically intentional development of credibility requires both time to seek out the contextualizing past conversations of one’s discourse and those elements of process knowledge evident in Tardy’s aforementioned framework.

Beyond the prior conversational context, in more recent rhetorical theory, scholars have begun to develop a more robust understanding of the physicality of various contexts through the continuing conversations of ambient rhetoric (Rickert, 2013) and material ecologies (Jordan, 2015). These understandings link to actor network theory (ANT) which represents a “disparate family of material-semiotic tools, sensibilities and methods of analysis that treat everything in the social and natural worlds as a continuously generated effect of the webs of relations within which they are located” (Law, 2008, p. 2). For second language learners, this development may allow for an additional index of acquisition once fully understood and harnessed. Aspects of this physical locality have already been broached in conversations about the whole language learner (Schwarzer, 2009, p. 30) where adapting the learning environment to “feel more like home” was a way of enhancing the learner
experience. Further, these considerations may link to deeper contextual needs of L2 learners that take into account the emotional, psychological, and mental associations of place and comfort as they intersect with linguistic and rhetorical knowledge.

There is some overlap in this regard with the way writing gets manufactured in new technological contexts with their unique user interfaces and experiences, where research on language acquisition is connected to alternate digital spaces. Research has broached the threshold of various types of digital spaces as learning environments (Aydin & Yildiz, 2014; Chao & Lo, 2011; Fernsten, 2008)—wikis, use of avatars, bulletin boards, and document sharing platforms—but researchers might still delve into the greater personalization of these digital spaces and the potential scenarios within these environments which might be used to aid the development of linguistic and rhetorical knowledge and awareness.

**Genre knowledge and awareness**

Beyond these four components that each provide a way in to understanding how rhetorical knowledge gets enacted, developing situated and flexible genre knowledge is another important component to being considered advanced or gaining membership within specific social and rhetorical ecologies. Developing genre knowledge is also advanced because it encompasses all four knowledges of Tardy’s framework. Acquiring an awareness of the rhetorical components behind language but beyond syntactic conventions is another important way that a writer could be measured as advanced. However, acquisition of these underlying rhetorical components can be more challenging than acquiring linguistic patterns because one must develop greater meta-awareness as a result of greater text length, increased text complexity, and the potential fluidity of the genre (DePalma, 2015; Yang, Lu, & Weigle, 2015).

Discerning rhetorical structures as opposed to syntactic structures for language learners can be more challenging at times because one might develop and internalize the syntactic structure of a simple SVO sentence or even a compound sentence fairly quickly by looking at a series of sentence examples. One might fit a dozen or more of such examples on a single page and ascertain the pattern moderately quickly, but to observe (much less produce) larger rhetorical structures even the size of a single paragraph requires significantly more reading time and involves significantly more potential variation among samples—not to begin thinking of larger, more complex written products like an essay, lab report, professional cover letter, dissertation, or manuscript. Advanced writers may need to discern the relationships between sentences, chunks, paragraphs, and larger document sections. And further, advanced writers do all of this without distancing those who are already enculturated to the genre conventions of the specific context.

Multiple models have been introduced for enculturating writers to these challenges associated with genre acquisition, knowledge, and awareness (Bazerman, Bonini, & Figueiredo, 2009). More specifically, aside from substantially theorizing approaches to genre in her book, *Building Genre Knowledge*, Tardy (2009) also shares
an example of self-enculturation, where one student “…[sought] out rhetorical
information. By following a peer as he applied for jobs, asking non-experts for
advice, and searching Web sites, John learned not only what a text looks like, but
also how it operates within specific groups for specific purposes” (p. 94). This
example, as well as others included in Tardy’s monograph, provides tangible strat-
egies for helping writers acquire both knowledge and awareness of localized genre
conventions. Further, in an alternate review of genre-based writing instruction
(GBWI), Johns (2011) describes “a pedagogical and theoretical divide” between
genre acquisition and genre awareness, sharing that

Genre acquisition requires the direct teaching and student learning of specified text
types which are considered by practitioners to be common exemplars of genres.
Genre awareness, on the other hand, … refers to examining the relationships among
texts, their rhetorical purposes, and the broader contexts in which texts from a genre
may appear. (p. 57)

This divide between acquisition of knowledge and development of awareness
aligns with the differences between formal knowledge and rhetorical knowledge.
Approaches that emphasize acquiring specific formal genre knowledge are based
on more static perspectives of the rhetorical situation, often indicative of fossilized
or reified genre conventions, whereas flexible perspectives emphasize the need for
adaptability to rhetorical ecologies that see conventions as transforming and
adapting over time.

From an advanced SLA perspective, both of these approaches may be a way to
develop forms of advanced knowledge, by depicting what is commonly written and
also what level of risk the writer may chance when bending or breaking conventions
in the particular ecology. Even static knowledge about genre and conventions will
help writers position themselves as members of a context; however, more flexible
rhetorically based perspectives are likely to make for more adaptable writers capable
discerning the ideology behind the conventions. These advanced writers would
be able to adapt textual features, underlying rhetorical structures, and inferred ideo-
logical knowledge to the needs of the specific writing context (Coe, Lingard, &
Teslenko, 2002; Johns, 2011, p. 65). Adding to Tardy and Johns, Hyland (2002) indi-
cates the important consideration of interdiscursivity and genre mixing, which can
often make even more salient the need for greater flexible awareness as opposed to
tacit formal knowledge of common genres or schematic structures used in a
particular context. More recently, scholars (Marcela, Castro, & Chala, 2013; Yasuda,
2011) have considered developing genre flexibility across languages as a means to
develop awareness. And while at the time Hyland mentioned the longevity of some
genre structures, he also indicated the gap in regard to new technological genres
which have since taken genre mixing to unprecedented levels of, what can seem
like, almost constant flux. As a final component of our exploration of advanced rhet-
oric and writing, developing genre knowledge can be challenging as it encompasses
all four knowledges. Still, advanced writers and language users can gain signifi-
cantly by investing the extra effort necessary to develop this knowledge.
Conclusion

This chapter has been an exploration of rhetoric and writing as two important considerations of advanced SLA. We have elaborated on how these two considerations influence the development of socially situated language acquisition by first providing brief histories of both rhetoric and writing. Then we used a four-knowledges framework to establish the need for broader definitions of advanced in SLA, and we scaffolded rhetorically based models to highlight the unique contributions of rhetoric and writing to a socially situated understanding of advanced SLA. These models shift the focus of acquisition from the target-oriented complexity, accuracy, and fluency constructs to an efficacy-oriented multiple knowledges framework for advanced language learning.

Future directions of efficacy-oriented, rhetorically based models of SLA

The reorienting or shifting of emphasis that this chapter explicates reveals a need to adjust the focus of research as well. A copious amount of research in second language studies examines error and corrective feedback when learning and writing in a second language, but shifting focus toward research in efficacy could produce significant findings beyond what language is produced. Further, some branches of SLA research could benefit from a shift away from a primary emphasis on micro-level concerns of morphemes, lexicons, and syntax to a macro view of acquisition that includes larger rhetorical components and ecologies as well as situational and cultural influences on perceptions of language efficacy, especially on the part of the readers and receivers. And because much study of rhetoric has been situated in English-speaking countries and because much study of writing tends toward university and professional environments, there is a need for deeper, more widespread inquiry into advanced genre, writing, and rhetorical practices in languages other than English and a broader variety of ecologies as well. As a final note, a greater level of interdisciplinary research among language, writing, and rhetoric scholars needs to occur to provide a more robust depiction of SLA at a rhetorically advanced level in a variety of contexts.

NOTE

1 For those interested in a larger overview of the history of western rhetoric, Foss, Foss, and Trapp (1985) and Leki and Silva (2004) have given brief histories of rhetoric and some of its key epistemological anchors. Teachers might also enjoy the overview provided in chapter 4 of Érika Lindemann’s A Rhetoric for Writing Teachers (2001).
REFERENCES


Introduction

It is reasonable to question the importance of sociolinguistic variation in the context of second language speakers, particularly those at lower proficiency levels who are focused primarily on communicating basic messages to meet immediate needs. To be sure, a speaker who is unable to construct an utterance or to find the lexical items to express a given meaning faces far greater consequences than one who is unable to use linguistic resources to show they are forming an interpersonal connection with another speaker. Furthermore, research suggests that we afford certain liberties to non-native speakers. For example, Hanulíková, van Alphen, van Goch, and Weber (2012) found that Dutch listeners reacted more strongly (i.e. had larger P600 values) to sentences that contained syntactic errors than those that did not when those utterances were produced by native speakers of Dutch compared to when the same utterances (with and without errors) were uttered by a Turkish-accented speaker. Nevertheless, many second language learners have both integrative and instrumental goals for language acquisition and there is no doubt that failure to understand the social norms for communication may lead to difficulties in social interactions as well as commercial and academic ones. Beginning with early work, such as that by Canale and Swain (1980) who analyzed the skills that classroom learners must acquire in order to fully participate in second language communication, and moving to more recent work that examines language use in context (e.g. Hashimoto, 1994; Howard, 2005; Kanwit, Geeslin, & Fafulas, 2015; Marriott, 1993; Mougeon, Nadasdi, & Rehner, 2010; Olson Flanigan & Inal, 1996; Preston, 2000; Regan, 2010; Tarone & Liu, 1995), this body of research takes as a foundation that knowledge of the grammatical rules of a language is essential but not sufficient to communicate effectively in the
target language. Along with this grammatical knowledge, learners of a second language must also understand how to demonstrate familiarity, affection, politeness, solidarity, and distance. Failure to do so will inhibit one’s ability to form alliances, make friendships, or assert authority, to name only a few examples. Certainly, when the discussion of sociolinguistic variation focuses on advanced learners, as it does in the current volume, there are clear benefits for speakers who accurately convey and interpret linguistic, social, and geographic norms.

The current chapter focuses on linguistic variation in second languages, especially for learners with higher levels of proficiency. The first step in understanding second language variation is having a clear picture of what native speakers do. Although brief, we will include an overview of some of the ways that native speakers vary their language across interactional settings. Taking that information as a starting point, we may then consider how learners come to possess these same abilities. The current chapter will visit morphosyntactic and phonetic variation with linguistic, social, and geographic correlates and explore both interpretive and productive abilities. Once we understand what highly advanced learners are (not) able to do, we may ask several different questions about why that may be, what factors might improve this state of affairs, and what the consequences of these challenges may be across interactions. Finally, the chapter concludes with a discussion of what future research must consider in order to move the field forward in each of these areas.

What abilities must a sociolinguistically competent speaker possess?

When we think of using language in situationally appropriate ways it is often the more overt expressions of formality that come to mind. For example, competent language users are able to select the appropriate form of address for each interlocutor they encounter and to use situationally congruent lexical items. In essence, one knows when to say ‘hey grab me that thingamajig’ and when to say ‘excuse me sir, would you please pass the salt shaker?’ Polite vocabulary is something we may even explicitly teach, both to children and to L2 learners. However, native speakers vary language across interactional settings at other levels of the grammar as well, and this often occurs below the level of consciousness. A speaker of Spanish in Venezuela who never omits or aspirates a syllable-final/s/or a speaker in the Southern United States who never produces ‘-in’ (vs. ‘-ing’) in words like ‘swimmin’ or ‘fishin’ might sound excessively formal or even just odd (Loudermilk, 2015; Ruiz-Sánchez, 2004). These examples illustrate the importance of modifying the realization of the sounds that make up a language according to contexts of use. Likewise, syntax and morphology also vary according to characteristics of the speaker and the discourse setting. In English we know that formal language requires us to avoid stranded prepositions, but a classmate who approaches and asks “to whom are you speaking?” would sound quite out of the ordinary.
Likewise, verbal morphology can vary between regularized forms, acceptable in spoken language, and the more prescriptive forms expected in academic writing. From this very brief sampling, we easily see that the language a speaker produces provides information about the situation in which the interaction takes place as well as the social characteristics of the speaker and hearer(s).

Our knowledge of language variation is not limited to the language we produce. In fact, sociolinguistic competence involves sophisticated interpretive abilities as well. For example, native speakers can perceive information related to social characteristics of a speaker, such as age, ethnicity, region of origin, sexual orientation, and social class, just by listening to recorded speech (Campbell-Kibler, 2010; Levon, 2014). Well-known examples of this ability are seen in studies such as Staum Casasanto (2009), which examined the relationship between ethnicity and t/d deletion and showed that hearers expected greater rates of deletion from African American speakers and also perceived greater rates when they believed the speaker was African American, and in the so-called army brat studies of dialect identification, which explore how well speakers are able to identify various regional dialects (e.g. Clopper & Pisoni, 2004). At a minimum, native speakers are fairly good at knowing which speakers share their own geographic background and which speakers do not. Taken together, this body of research shows that native speakers glean detailed non-linguistic information during the process of perception and interpretation and, in fact, it has been suggested that instructed second language learners should begin the process of acquiring sociolinguistic competence by focusing on these interpretive abilities (Geeslin & Long, 2014).

Interestingly, native speakers can also use social information to guide perception. We see this in cases like the research by Staum Casasanto (2009) described previously for ethnic identity, where perception was influenced by the picture of the speaker leading to greater rates of perception of t/d deletion with African American photos. This effect is not limited to ethnicity, however, and has also been widely attested for geographic differences. For example, Niedzielski (1999) showed that labeling the participant response sheet as either “Michigander” or “Canadian” influenced vowel perception of the Detroit-based participants. Likewise, Drager and colleagues (e.g. Hay & Drager, 2010) have shown that activation of a particular country of origin (i.e. Australia or New Zealand), even through the use of stuffed animals that are present but not mentioned during the session, can influence the perception of vowels in recorded speech. A final example in Spanish can be seen in recent work by Leal, Geeslin, and Escalona Torres (2016), who showed that listeners in Mexico were influenced by the dialect of the speaker in the recorded instructions that preceded the self-paced reading task. It is important to note that for native speakers, these influences on perception are not necessarily at the level of consciousness. The growing body of research on the role of hearer expectations continues to demonstrate the complex relationship between social and linguistic factors. Given all that we know about how language varies sociolinguistically and how social information influences language use and interpretation, it is clear that highly proficient L2 learners must develop sophisticated abilities that are exhibited across a range of linguistic behaviors, including production and perception.
What are advanced learners able to do?

In the brief overview of native speaker sociolinguistic competence, we have seen that the ability to produce variable structures according to social norms and also to interpret linguistic and extra-linguistic information in the language to which one has access is implicated. To be sure, we would not expect beginning learners to exhibit the same variable patterns of production, interpretation, and processing as native speakers. However, highly proficient learners often develop native-like tendencies or at least show development in that direction on some or all of these measures (e.g. Adamson & Regan, 1991; Geeslin & Fafulas, 2012). In the current section we explore in greater depth the sociolinguistic competence of highly proficient second language learners.²

We know that native speakers converge on the frequency of use and the constraints that govern the use of a given variant within a particular speech community. Thus, for example, forms of subject pronouns are related to the person and number of the verbal subject, to the previous mention of that same referent, and to the verbal tense, mood, and aspect (e.g. Carvalho, Orozco, & Shin, 2015). Likewise, we know that while these constraints are relatively constant across varieties of Spanish, speakers with different countries of origin show varying rates of overt subject pronoun use (e.g. Erker & Otheguy, 2016). In sum, we are able to characterize native speaker patterns by examining the frequency of use of a given form and the linguistic, social, and geographic constraints on that use. Likewise, we are able to measure the degree to which variable structures are used in similar ways by non-native speakers by using this same metric. There is a wealth of research now available on linguistic variation and highly proficient learners in many languages, including English, Japanese, French, Spanish, and Chinese (e.g. Durham, 2014; Fafulas, 2013; Hashimoto, 1994; Kanwit, 2014; Li, 2010; Linford, 2016; Long, 2016; Mougeon, Nadasdi, & Rehner, 2010; Olson Flanigan & Inal, 1996; Regan, Howard, & Lemée, 2009; Rehner, 2004). The structures examined are also quite diverse. In French, for example, research has covered l-deletion (Howard, Lemée, & Regan, 2006), interrogatives (Donaldson, 2016), future-time reference (Gudmestad & Edmonds, 2016), the use of nous versus on (Sax, 2003), ne-deletion (Dewaele & Regan, 2002; Donaldson, 2017), and schwa-deletion (Uritescu, Mougeon, Rehner, & Nadasdi, 2004), to name only a few. Generally, learners with extensive experience (e.g. a year abroad) were able to move toward native-like patterns, but nevertheless, in many cases continued to use informal variants at lower rates than native-speaking counterparts or in slightly different contexts. In work on Spanish, researchers have studied the production of the copula contrast, subjunctive, subject forms, present progressive aspect, future-time reference, differential object marking, /d/-deletion, and geographic variants, such as use of the interdental fricative /θ/ and /s/-weakening, often with highly proficient learners included in the participant pool (Fafulas, 2013; Geeslin, 2003; Geeslin & Gudmestad, 2008, 2010; Gudmestad, 2012, Chapter 18 of this volume; Kanwit, 2017; Killam, 2011; Knouse, 2012; Linford, 2016; Ringer-Hilfinger, 2012;
Solon, Linford, & Geeslin, in press). In general, we see that the higher the level of proficiency, the more subtle the differences between native and non-native speakers.

One way to measure the subtle differences between native and advanced non-native speakers is by looking at the frequency of use of a form and the linguistic predictors of those patterns of use (for brief overviews see Gudmestad, 2014, for Spanish, and Regan et al., 2009, for French). It is sometimes the case that advanced learners differ from native speakers on only one measure rather than across all measures. For example, Geeslin (2003) found that advanced learners were able to approximate native rates of use of the copula estar ‘to be’ (relative to ser ‘to be’) on a preference task, where the rate was about 45% for both groups and the groups were not significantly different. Nevertheless, the predictive model for those rates of selection identified several linguistic factors that predicted preferences for both groups, but differences also arose. For example the native speaker model included the factor ‘predicate type’ as a significant predictor of selection of estar whereas the non-native speaker predictive model included the factors ‘frame of reference’ and ‘experience with the referent,’ but not ‘predicate type.’

It is also the case that sometimes the frequency of use or selection of a form is significantly different for native and non-native speakers but the predictors of these rates are fairly similar. For example, Gudmestad (2012) showed that advanced learners of Spanish had the same linguistic predictors of use of the subjunctive as native speakers do. Likewise, Geeslin and Fafulas (2012) found different rates of use of the present progressive in Spanish (vs. the present), which was used more frequently by non-native speakers than by native speakers, but the linguistic predictors of this use were very similar. For both groups, the linguistic factors aspect, clause type, object form, object type, object number, and object position were all significant. For the non-native group the factors ‘adverb’ and ‘clause type’ were also significant even though these did not enter into the predictive model for the native speakers. Nevertheless, the pattern across the categories of these two linguistic variables (i.e. the direction of the effect) was the same for both groups. In other words, despite the differences in frequency of use of these forms, the patterns of use were the same. Fafulas (2013) found something similar, again for the present progressive form, except that the factor ‘polarity’ proved to be significant for the learners but not for native speakers in Mexico or Spain. As a final example of this type of difference, Solon et al. (in press) examined patterns of /d/-deletion in Spanish in sociolinguistic interviews with advanced non-native and native speakers. This variant is associated with less formal speech but also with a host of linguistic factors. In word-internal intervocalic position, native speakers produced /d/ as a stop only 3.3% of the time and advanced non-natives did so 7.3% of the time. The greater difference between groups, however, was seen in rates of deletion in these same contexts (as opposed to approximant productions). The native speakers deleted /d/ 44.5% of the time whereas the non-natives did so only 18% of the time. The predictors of this deletion were the preceding vowel, grammatical category of the word, stress, and the following vowel. The same predictors were significant for the non-natives except that frequency was also
significant and, moreover, it was the most significant predictor in the model. These examples are not unusual and, in fact, these linguistic predictors of patterns of use and preference are often the best indicators of how natives and non-natives differ.

In addition to frequency of use of a form and the linguistic factors related to those patterns of use, several studies of advanced learners show individual characteristics that predict patterns of use. We take as an example Li (2014), who studied subject expression in Chinese and showed that subject expression was linked to the linguistic factors person and number of the verb and switch reference for native and advanced non-native learners. However, the patterns of expression for non-native speakers in that study were also linked to length of residence in China, gender, proficiency, and discourse context. Likewise, in Li’s 2010 study of de-deletion in Chinese her analysis of interview data produced by 12 native speakers and 20 non-native speakers (eight high-intermediate and 12 advanced) found that frequencies of use differed slightly, but that there were both linguistic and social predictors of these differences. The particle de is used differentially across functions, such that both native and non-natives used de 100% of the time in conditional clauses but in contexts that are more variable, such as adjective + noun contexts, native speakers used de only 28% of the time but the learners were still using de at a rate of 55%. One factor linked to these differences was first language (including Korean, Russian, English, and Japanese), as was speaker gender, degree of formality, and interaction with native speakers. In fact, several studies show a link between individual experiences with the target and patterns of use. For example, Sax (2003) showed that use of several different structures in French by 30 American learners, including ne-deletion, /l/-deletion, and the nous versus on alternation, was linked to proficiency in French and experience abroad. Fafulas (2013) showed that experience abroad was a predictive factor of use of the present progressive, and George (2014) showed that social networks were important in understanding use of the geographically variable interdental fricative in Spanish.

In recent research, Kanwit (2017) analyzed data elicited through a personal prompt response task from native speakers and five levels of English-speaking learners of Spanish. He showed that the most advanced group of learners was quite similar to the native speakers in both the frequency of use of the four forms produced to express future-time events (i.e. the morphological future, periphrastic future, present indicative, and lexical future forms) and the linguistic predictors of that use. Interestingly, there were also common social factors that held across the advanced non-native and the native group. For example, native speakers from Mexico and learners with experience in Mexico were less likely to use the morphological future over the periphrastic future than Spanish native speakers and learners who had studied in Spain. To be sure, this type of finding moves research beyond the role of experience abroad as a function of quantity of input or proficiency level, and toward identifying the groups with whom a speaker identifies and interacts. This is not unlike the earlier work of Eisenstein (1982), who showed that second language learners of English adopted native dialect norms and were influenced by the target-speaking group with whom the learners identified.
Looking at social factors beyond those related to language experience, Adamson and Regan (1991) showed that gender of the learner was linked to production of the variant ‘in’ (vs. ‘ing’) in English. Perhaps the most interesting fact about this last case is that while male native speakers use this variant more than female ones, the non-native speaking males were not only more frequent than their female counterparts, they overshot the male native speaker norms as well. Collectively, these cases show that the subtle differences between natives and non-natives can be linked to rates of use as well as linguistic and extra-linguistic factors that condition that use.

The importance of individual characteristics in explaining the use of variable structures has led to a handful of studies that question whether it is appropriate to analyze second language learners as a group, potentially obscuring the differences between individuals. Looking at sociolinguistically variable forms of subject expression, Geeslin, Linford, Fafulas, Long, and Díaz-Campos (2013) examined native speakers and learners at five levels of proficiency on a contextualized preference task. They found that the rates of selection for the most advanced group and for the native speakers were not statistically different (66% vs. 67.5%). These two groups also had the same significant linguistic factors in the predictive models identified by the analysis. In addition to these aggregate analyses of group patterns, the authors of that study looked at the range and standard deviation in rates of use of overt and null forms, finding that the standard deviation was near 20% except for the most advanced group (13.8%) and the native speakers (14.1%). Additionally, when the factor ‘individual participant’ was entered into the analysis, this factor was included in the predictive models for all groups (including native speakers) but it lost statistical power (i.e. decreased in significance) as proficiency increased. The authors took these two findings together as evidence that the role of individual differences may diminish as proficiency increases. Although more research in the area is needed, this finding may depend on the linguistic structure examined. For example, Geeslin and Gudmestad (2012) conducted a similar analysis of the role of individual differences in the rates of selection and factors predicting the selection of forms of future-time reference. They focused their analysis on the two highest levels of a group of 151 English-speaking learners at five proficiency levels of Spanish and on 22 native speakers and showed that the highest level of learners actually varied less in rates of selection than the native speakers. They found the linguistic constraint ‘temporal distance of the event’ to be the most important predictor and used distribution across the categories of this variable as a second measure of differences between groups and individuals. In general they found that learners, even at the highest levels, still differed from the group norms, sometimes in both frequency of selection and the predictors of selection, but more often in only one of these measures. The upshot of this analysis is that even as proficiency increases we see a decrease in individual variability but this variability remains prevalent.

As demonstrated in the preceding paragraph, it may be the case that different patterns exist for different variable structures, and evidence is starting to suggest this more clearly. Certainly one could argue that the structures examined in Geeslin
et al. (2013) and Geeslin and Gudmestad (2012) contribute to the differing findings across studies. Sax (2003) also found that for English-speaking learners of French, the degree of stylistic variation exhibited depended on the linguistic structure. Specifically, she found that ne-deletion showed more stylistic variation, followed by /l/-deletion and then by the nous/on variation. Likewise, Kanwit et al. (2015) analyzed data from two separate groups of study-abroad learners, one in Mexico and one in Spain. They used a contextualized preference task to compare learner preferences on three structures (the copula contrast, the present perfect vs. preterit contrast, and the simple present vs. progressive contrast) at the start and end of their stay abroad, against data from native speakers on the same task. Their analysis showed that learners moved toward native norms for all three structures in both study-abroad sites, but the extent and rate of change varied across linguistic structures. Linford (2016) showed similar findings for English-speaking students of Spanish in the Dominican Republic and Spain, after a semester abroad, when the patterns and rates of use of subjects, objects, and past tense forms in Spanish were examined. There is no doubt that the learners in these last two studies do not reach the same level of proficiency as those in the other studies cited previously. Nevertheless, taken together, we see that it is worthwhile to consider the role of the variable structure itself in the differences between natives and non-natives on the one hand and between individual and group norms on the other.

Throughout this chapter, we have seen that sociolinguistic competence extends beyond the production of or preference for a form and includes the ability to perceive and interpret social information and to evaluate that information according to norms that are appropriate for particular groups or individuals. For example, the ability to detect and properly interpret markers of informality is an important component of understanding the interaction. There are relatively few examples of this type of research in the second language context, but the examples we describe here illustrate promising directions for future study. Kanwit and Geeslin (2014) examined the interpretation of verb forms in the subjunctive and indicative moods following adverbial conjunctions by 16 native speakers of Spanish and 97 English-speaking learners of Spanish at three levels of proficiency. Their 24-item interpretation task, which manipulated the verb form (subjunctive or indicative) as well as the adverbial conjunction and the regularity of the verbal morphology, asked participants to indicate whether an event was habitual (usually associated with the indicative mood form) or whether it had not yet occurred (generally associated with the subjunctive mood form). In comparing their results for the most advanced group to those for the native speaker group, we see that the advanced learners actually show less variability and more prescriptive interpretations. Moreover, the native speaker patterns were less constraint-governed and more lexically specific than those of the non-natives. The upshot is that even advanced learners have not converged entirely on native norms of interpretation of this socially variable structure.

Turning to the evaluation that accompanies socially variable forms, it is also important to know whether advanced non-natives evaluate social information in the same way as natives. Research on production, such as Kanwit (2017),
suggests that affiliation or experience with a particular region does, in fact, influence patterns of future-time expression. Geeslin and Schmidt (in press) sought to examine how learners evaluate geographically indexed speech by native speakers from several different regions of origin. They employed a matched-guise design and asked participants to rate speakers on scales of individual qualities, such as intelligence, employability, likeability, and friendliness. In fact, the learners in that study did evaluate speakers from different countries of origin differently but this was not influenced by experience with that variety in predictable ways. Even the level-four learners in that study (arguably less advanced than those in Kanwit, 2017) showed that experience did not consistently improve (or decrease) ratings on any of the individual characteristic scales. To be sure, there are several possible explanations, including links to social networks (e.g. Ringer-Hilfinger, 2012) or to intelligibility (e.g. Schoonmaker-Gates, 2012), but again, this preliminary research suggests that further work in this area will yield important information about sociolinguistic competence among highly proficient second language learners.

We are also beginning to see the importance of examining the manner in which sociolinguistically variable structures are processed by native and non-native speakers. Geeslin and Leal (2016) used a self-paced reading task to explore the processing of the geographically variable direct object pronoun in Spanish. It is widely attested that in central Spain the dative form le is used variably in accusative contexts where lo occurs in other regional varieties, especially with masculine nouns (e.g. Klein-Andreu, 2000). The authors designed a task that manipulated grammatical gender of the noun and the pronominal form (le vs. lo) to determine whether speakers reacted differently to the two forms based on language experience. Their analysis showed that native speakers of Spanish read the le form slightly more slowly than the feminine form la with feminine nouns (as predicted in the literature) but did not read le more slowly than lo with masculine nouns, indicating that they accepted this geographically variable form and the grammatical gender bias associated with it. In contrast, even the most advanced learners read le more slowly with masculine nouns than lo, showing differences from the native speakers in their processing patterns despite extensive experience in some cases in regions where le is prevalent in these contexts.

This review is by no means exhaustive but rather a selection of empirical studies showing the range of abilities subsumed under the construct of sociolinguistic competence as well as the varying degrees to which highly proficient learners of second languages have been shown to achieve such competence. In every case we note that learners show evidence of approximating but generally not reaching native norms for production, interpretation, or perception (see Donaldson, 2017, for an example of an exception). To be sure, nearly every study looks at aggregate data and we know that, even amongst highly proficient learners, there is tremendous variability in the ultimate attainment. To explore this variability, and more concretely how to overcome these differential outcomes in learning, we turn now to some of the factors that have been said to contribute to successful acquisition of the norms for linguistic variation in a second language.
What additional factors play a role in the acquisition of sociolinguistic competence?

Throughout this chapter we have seen that highly proficient learners of a second language can converge on native speaker norms, although the degree of convergence appears to depend on facts related to the individual, the learning context, and the variable structure itself, to name only a few potential sources of difference. These differences might lead us to ask which factors matter the most for successful acquisition of variable structures. In this section we explore a handful of factors that have been shown to contribute to positive outcomes in the acquisition of sociolinguistic variation.

One factor that has been linked to attainment in general and to attainment of sociolinguistic variation more specifically is access to an experience in the target language and culture. While the length and nature of the stay in the target environment can vary from one study to another, the central research focus across settings is increased access in quantity and range of contexts to language input (e.g. Díaz-Campos, 2004; Kinginger, 2008; Lafford, 2006). For sociolinguistic variation in particular, this increased range of interactional contexts is important for learners to acquire the full range of social meaning associated with variable structures (e.g. Regan et al., 2009). Beginning with early research, such as Ryan and Lafford (1992), which confirmed the universal stages of acquisition established for the copula contrast in Spanish (e.g. VanPatten, 1987), differences across learning contexts are attested in the acquisition of variable structures that are linked to the input associated with a stay abroad. Specifically, Ryan and Lafford showed that the general stages of acquisition held across learning contexts but the order of acquisition of adjectives of condition with estar and locatives with ser was reversed, perhaps owing to increased exposure to expressions, such as está cerrado ‘it is closed,’ which are related to daily scheduling challenges that students abroad often face. Of interest for the current discussion is that adjectives of condition are variable and their acquisition appears to have been supported through study abroad. Since that early work, we have seen that learners can and do move toward native and regional norms over the course of study abroad. In Spanish alone this has been documented for future-time forms (Kanwit & Solon, 2013), subject forms (Linford, 2016), progressive aspect (Fafulas, 2013; Kanwit et al., 2015), the copula contrast (Kanwit et al., 2015), past time expression (Geeslin, Fafulas, & Kanwit, 2013), question intonation (Henriksen, Geeslin, & Willis, 2010), and the interdental fricative /θ/ (George, 2014; Ringer-Hilfinger, 2012; Willis, Geeslin, & Henriksen, 2009). In general, this body of work supports the assertion that even highly proficient learners benefit from experience in the target environment and that exposure to a wider range of contexts does lead non-native speakers to develop toward more native-like patterns of variation.

Despite the consensus that study abroad is helpful in the development of sociolinguistic variation, however, we also have considerable evidence that this experience alone does not guarantee native-like use of variable structures, particularly in
the case of geographically indexed variants. For example, in a large-scale study of 130 learners at five levels of proficiency, Geeslin and Gudmestad (2008) found that even at the highest levels in which many individual learners had experience in a study-abroad context only nine learners produced the interdental fricative /θ/ even a single time and only five demonstrated even a single instance of /s/-weakening. What is especially interesting about this result is that use of a variant was not contingent on experience abroad (e.g. only 50% of the learners of Levels 3 and 4 who produced the interdental fricative /θ/ had been abroad), and conversely, many of the learners who did not use these variants had been abroad (e.g. 75% at Level 4 and 92% at Level 5). In short, the relationship between study abroad and use of regional variants is neither causal nor categorical across all learners.

Given the disparity between production of regional variants and exposure to them, researchers have begun to develop methods to explore a second factor, the nature and extent of contact, both abroad and over the life span of the acquisition process. The idea that social relations influence acquisition is certainly not a new one (c.f. Schuman, 1978) and we know that identity is connected to one’s investment in the acquisition process (Norton & McKinney, 2011). In the study-abroad context more specifically, Lybeck (2002) showed that for sojourners in Norway second language pronunciation was mediated by the degree of acceptance and involvement they felt in the target community. Likewise, Isabelli-García (2006) showed a connection between performance on a simulated Oral Proficiency Interview on the one hand and social networks during the stay abroad in Argentina on the other.

In addition to increasingly sophisticated and detailed questionnaires involving the degree of social connections (e.g. Freed, Dewey, Segalowitz, & Halter, 2004), recent research has paired in-depth analysis of the acquisition of regional variants and social interactions during a stay in the target environment. One excellent example is George (2014), who examined production of two variants associated with central Spain, the interdental /θ/ and uvular [χ] fricatives, by English-speaking learners at three times during a one-semester stay in central Spain. She analyzed three production tasks (a word list, an interview, and a passage read aloud) as well as a language contact profile and a social networks questionnaire, which elicited information about how often and with whom each individual used Spanish, and their attitudes, anxiety, and motivation during the stay abroad. This analysis showed that amount of contact, level of anxiety, and attitude toward Castilian Spanish were all related to increased use of the regional variants over the course of the semester. The general upshot is that production of variable structures is intricately linked to the identity and experience of the language learner.⁶

A third and final factor that has been explored is the role of instruction in the development of sociolinguistic competence. It is certainly the case that the role of instruction in the acquisition of variable structures is not widely studied (see Lord, 2010 for an example of related research that does not focus on sociolinguistic variation). A notable area of interest, however, is the growing body of research on the role of instruction in the acquisition of various speech acts. This area of pragmatics is related to other work on variable structures
because speech acts can be completed according to local target norms through strategies that tend to conform across speech communities, but which vary according to speakers and the interactional context. For example, Cohen and Shively (2007) studied the acquisition of requests and apologies over the course of a stay abroad (some learners were in Spain and some were in France). One group of learners in each location received explicit instruction and a journaling project assignment, while the other half did not. Rather than documenting the importance of instruction, however, their results showed that both groups improved over time and there was no significant advantage for the instructed group. In a similar analysis, Shively (2011) analyzed the development of requests in service encounters during a semester-long stay in Spain. One key difference in that study was that Shively did find that some of the strategies taught were incorporated by the learners over the course of the stay abroad. Outside the study-abroad context, Hasler-Barker (2013) examined the effect of instruction on compliment and compliment responses in a second language classroom context. Her designed included an input-only group, an implicit metapragmatic instruction group, and an explicit metapragmatic instruction group. Her results did show an advantage for the groups that had received compliment instruction of either type. We see from this handful of studies on the role of instruction in the second language development of pragmatics, and likely all variable linguistic strategies for using language in context, that the effectiveness of instruction varies across structures, learning contexts, and methods of instruction. As with study abroad, there is some evidence that instruction can be helpful, but no clear evidence that it is uniformly required for success nor that it is categorically successful for all learners.

Future directions for research on advanced proficiency and sociolinguistic competence

The study of the acquisition of variable structures and sociolinguistic variation has become increasingly robust in recent years and, even when we limit our discussion to high-proficiency learners, we find examples of empirical studies spanning multiple language pairs, linguistic structures, and learning contexts. Nevertheless, there are several directions that offer the opportunity to expand our knowledge in these areas and to refine our understanding of second language acquisition more broadly. This final section addresses three broad directions and provides a brief discussion of the research initiatives that fall within each. This discussion is not exhaustive and its primary goal is to stimulate further interest in the topic.

In looking at high-proficiency learners in particular, we see that the rates of production and constraints on production of variable forms tend to converge on native norms, particularly in the area of morphosyntactic variation (e.g. Geeslin & Fafulas, 2012). Nevertheless, we are starting to see that tasks that
measure non-production constructs reveal important differences between native speakers and highly proficient learners. For example, Kanwit and Geeslin (2014) showed that on an interpretation task examining subjunctive and indicative verb forms paired with a range of adverbial conjunctions, the most advanced learner group was more prescriptive in the interpretation of subjunctive mood forms than the native speakers. Likewise, the most advanced group in Geeslin and Schmidt (in press) showed differences in the social evaluation of various dialectal groups on a matched-guise task. Finally, Geeslin and Leal (2016) found differences in reading time between native speakers and learners across contexts containing the variable pronominal form le in Spanish. Through these examples, we see that a key direction for future research is in the area of task design and the measurement of language interpretation and processing as these developments allow us to explore the aspects of language acquisition that are not readily visible in production data.

A second area that requires additional study in the future is that of the role of the individual in the attainment of advanced proficiency. It is a long-standing generalization that most adult learners will not achieve high levels of proficiency in a second language (Birdsong, 1999) and that even within groups of the most advanced learners, there will be differences at the level of the individual (e.g. Zafar & Meenakshi, 2012). In the area of sociolinguistic variation where the characteristics of the individual are known to affect the language a speaker produces as well as the language other speakers direct toward that individual, the analysis of aggregate data is a reasonable starting point, but cannot be the lone source of our understanding of the nature of acquisition of variation. Nevertheless, relatively few studies (see earlier discussion for exceptions) explore the degree to which individuals within a particular proficiency group vary. In the face of increasing research documenting the importance of the individual (e.g. Gurzynski-Wiess, 2013), this too represents an important area of development.

Finally, the review of existing research on high-proficiency learners and the acquisition of linguistic variation has revealed that not all variable structures are acquired at the same rate and even advanced learners do not always converge on native norms. What remains understudied in this regard is what differences between grammatical structures might explain this difference. It is likely that one key difference has to do with the social value of each variable structure (for discussion see Kanwit et al., 2015). It is reasonable to hypothesize that structures that carry social stigma pose additional acquisitional challenges in part because learners tend to have less access to informal language contexts (e.g. Tarone & Swain, 1995) and because the social meaning must be acquired in addition to the linguistic one. We have some evidence that socially stigmatized forms are processed differently in native language (e.g. Squires, 2013), but we know very little of how the social value of a form might influence acquisition. This is an area where research on advanced learners may prove particularly useful precisely because lower-proficiency learners are unlikely to have sufficient opportunity to acquire these forms.
NOTES

1 See Tsai (2013), Coperías Aguilar (2002), and Bachman (1990) for communicative and sociolinguistic competence, as well as classic work by Dell Hymes (Hymes, 1972).

2 Several different means for classifying participants as advanced, including particular experiences (e.g. study abroad) or independent measures (e.g. the Oral Proficiency Interview), exist in the literature. With no means for corroborating these assessments independently, the discussion that follows includes studies in which the authors believed some participants to be highly proficient. We understand that actual proficiency levels across studies do vary.

3 Our examples come from work in Spanish because that body of work tends to include native speaker data on the same tasks and within a single analysis. However, inferences from work on other languages, such as French, are also possible through comparisons to existing sociolinguistic research on native speakers. For example, Regan, Howard, and Lemée (2009) compare learner use of *ne*-deletion to established figures for native speakers in Ashby (1981).

4 Bayley and Langman (2004) also explore this topic but with non-native (Type I) variation.

5 For more on use of first language matched-guise tasks to measure social evaluation, see Campbell-Kibler (2010), and for more on dialect identification in native languages, see Clopper and Pisoni (2004).

6 To be sure, this finding is not exclusive to variable structures. In an exploration of a number of Spanish phones, Díaz-Campos (2004) showed that improvement by English-speaking learners was better predicted by social contact than by access to study abroad and, thus, several of the at-home learners actually outperformed the study-abroad ones as a result of close social connections to native speakers of Spanish in the United States.

REFERENCES


M. DuFon (Eds.), Language learners in study abroad contexts (pp. 231–258). Clevedon, UK: Multilingual Matters.


Staum Casasanto, L. (2009). How is sociolinguistic knowledge represented? In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of
Index

Page references to Figures are followed by the letter ‘f’ in italics, while references to Tables are followed by the letter ‘t’. References to Notes contain the letter ‘n’, followed by the number of the note. ‘L1’ stands for ‘first language’ and L2’ for ‘second language’.

Abrahamsson, N. 52, 247, 252–253
Abramson, A. S. 325–326
abroad experience
  individual differences in advanced proficiency 159, 161, 162, 166
  mood distinction 346–349
ABX segment discrimination 250
accents 241, 246, 250, 251
Acevedo, A. 224
Achugar, M. 16, 19
acoustic analysis 291
acoustically enhanced input 295
ACTFL see American Council of Teaching Foreign Languages (ACTFL)
actor network theory (ANT) 540
adjectival subordinate clauses 347
advanced ideational resources 16–17
advanced knowledge, phonology 242–245
  feature competition model 244–245
  functional load 243–244
  and target-like phonology 247
advanced proficiency
  and agreement processing 41–43
  connected speech see connected speech corrective feedback and advanced proficiency learners 96, 98–100
Critical Period (CP) research see Critical Period (CP) research
defining L2 proficiency/advanced proficiency 13, 97
individual differences see individual differences in advanced proficiency
  and L2 parsing 31–41
lexical development 401–418
mood distinction 345–346
peer interaction and advanced proficiency learners 100–103, 105
pre-task (strategic) planning with high-proficiency learners 221–224
psycholinguistic approaches see psycholinguistic approaches and referential processing 38–39
scope interpretation see scope interpretation at advanced proficiency
second language acquisition (SLA) see second language acquisition (SLA)
semantics, interface with discourse at 452–456
semantics of the nominal domain at 444–446
SFL-informed research agenda 24–25
societal demands for 201–203

Edited by Paul A. Malovrh and Alessandro G. Benati.
© 2018 John Wiley & Sons, Inc. Published 2018 by John Wiley & Sons, Inc.
sociocultural theory, L2 education to promote 120–128
speech attainment see speech attainment, advanced level
task repetition effects 229–230
term ‘advanced,’ notions about 309–310
advancedness
concept 157–158, 174
construct of L2 advancedness 10–15
as a discourse phenomenon 143–146
in listening 167
in speaking 174
adverb processing 242
adverbal subordinate clauses 347
age of acquisition (AoA) 247, 289
age of onset (AO), CP 51, 52
see also Critical Period (CP) research early results 53, 54
means and standard deviations 60
morphology and syntax 54, 55, 59–61
Ahern, A. 353
Ahmadian, M. J. 225, 228–229
Akiba, S 450–451
Alameen, G. 304, 307, 308, 314
Alarcón, I. V. 187
Alavi, S. M. 100
Alcón, E. S. 472
Alderson, J. C. 345
Aldosari, A. 102
Aleman Banon, J. 395, 396
Al-Gahtani, S. 519, 520, 523n1
Allen, I. E. 204
alveolar flap 308
Amenos-Pons., J. 353
American Council of Teaching Foreign Languages (ACTFL) 97, 186, 279, 309, 350
benchmarks 191
Oral Proficiency Interview (OPI) see Oral Proficiency Interview (OPI)
oral proficiency scale 203
Proficiency Guidelines 162
Reading Proficiency Test (RPT) 162
tests 161
Amoy Chinese 271
ANCOVA (analysis of covariance) 64
Anderson, B. 80
Anderson, J. 271
Anderson, R. C. 488
AO see age of onset (AO), CP
AoA (age of acquisition) 247, 289
applied linguistics 528
aptitude–treatment interaction (ATI) studies 68, 107
Arche, M. J. 425, 426
Archibald, J. 252, 254, 257, 258, 276, 277
argumentative texts 20
Armenian language/speaking 159
Arnaud, P. J. 405
articles
definite and indefinite 140
English language 140
nominal domain, semantics of at advanced proficiency
definiteness and specificity 445
genericity 445–446
zero article 140
articulatory effort 324
Articulatory Phonetics 329
Asencion-Delaney, Y. 354
aspect
acquisition and development across bilingual continuum 368–369
Competing Systems Hypothesis (CSH) 369–373
Concept-Based Instruction (CBI) 369–373
guided induction 373–374
imperfective, as the true aspectual marker 365
imprecise and incomplete descriptions of 369
instructional settings and pedagogical rules 369–372
iterativity and habituality 365–367
knowledge of 370–372
advanced 369
developing advanced conceptualization of knowledge 370
problems with advanced conceptualizations (CBI and CSH) 370–372
language data and metalinguistic awareness 373–374
representation of 362–363
grammatical 363–364
temporal vs. non-temporal meanings 367–368
and tense 361–380
aspiration 324
Atanassova, G. 99
| ATI (aptitude–treatment interaction) studies | 68, 107 |
| attention control | 249 |
| noticing, situating in the instructional design | 207–208 |
| role, in phonology | 248–249 |
| Australian Department of Defence | 201 |
| AXB task | 249 |
| Ayres-Bennett, W. | 344 |

| Bachman, L. F. | 464 |
| Baker, M. C. | 77 |
| Baker, W. | 286, 291 |
| Bakhtin, M. | 12, 144 |
| balanced bilinguals | 179–180 |
| Ballantyne, J. C. | 45n3 |
| Baralt, M. | 206 |
| bare assertions | 19 |
| Bartning, L. | 252 |
| Bastos, M.-T. | 160, 167 |
| Battistella, E. | 269 |
| Bayesian Information Criterion (BIC) | 165 |
| Bayley, R. | 560n4 |
| Bazergui, N. | 286 |
| Beckett, G. | 24 |
| Bell, H. | 404 |
| Berman, R. A. | 146 |
| Best, C. T. | 333 |
| Bialystok, E. | 73 |
| BIC (Bayesian Information Criterion) | 165 |
| Bilingual Language Profile (BLP) | 186 |
| bilingualism |  |
| allocation of resources | 86 |
| balanced bilinguals | 179–180 |
| bilingual continuum, acquisition and development across | 368–369 |
| concept | 180–181 |
| continuum of | 192 |
| early bilinguals | 289, 292 |
| emergent bilinguals | 180 |
| heritage bilinguals | see heritage bilinguals |
| incomplete bilinguals | 180 |
| late bilinguals | 289 |
| late Chinese-Spanish bilinguals | 296 |
| semilingual bilinguals | 180 |
| simultaneous bilinguals | 81 |
| speech performance | 288 |
| terminology | 179 |
| binding and co-reference | 454–455 |
| Birdsong, D. | 253, 443 |
| Blackwell, S. E. | 423 |
| Blake, R. J. | 160, 167, 175, 191 |
| Blanchet, J. | 314, 315 |
| Blom, E. | 389 |
| Boersma, P. | 245 |
| Bofman, T. | 138 |
| Bohn, O. S. | 331 |
| Bohnacker, U. | 432 |
| Bonferroni correction | 55 |
| Bongaerts, T. | 247 |
| Bonilla, C. L. | 347, 348 |
| book reviews | 21 |
| Bottleneck Hypothesis | 384, 421, 446 |
| Bowden, H. | 181, 191–192 |
| Bowles, M. | 187, 193 |
| brain | 287, 292, 392 |
| adaptive response | 332–333 |
| brain-based modular subsystems | 75 |
| event-related brain potentials (ERP) | 394 |
| mammalian wiring | 326 |
| mid/brain representations | 78 |
| Brecht, R. D. | 202 |
| Brisard, F. | 364, 367 |
| British Council | 201 |
| British National Corpus wordlist | 230 |
| Britton, J. | 534 |
| Broselow, E. | 243 |
| Brown, G. | 314 |
| Brown, J. D. | 313 |
| Brown, N. A. | 158, 175 |
| Bruhn de Garavito, C. | 61 |
| Bruhn de Garavito, J. | 391 |
| Bui, G. | 221, 222 |
| Bulté, B. | 142 |
| Bundgaard-Nielsen, R. | 285 |
| Bygate, M. | 142, 228, 229 |
| Byrnes, H. | 1, 17, 157–158, 174, 176, 203 |
| Cabrelli Amaro, J. | 85–86 |
| CAF (complexity, accuracy and fluency) |  |
| advanced-level grammatical abilities, observing through lens of | 141–143 |
| constantly changing subsystems, failure to recognize | 143 |
| pre-task planning | 220, 222–224 |
| rhetoric, advanced | 530 |
| task condition effects | 220 |
| task repetition | 229 |
| CAH (Contrastive Analysis Hypothesis) | 270, 271 |
CALF measures 220, 225, 228, 232  
see also CAF (complexity, accuracy and fluency)  
effects of pre-task planning on 222–224  
Callies, M. 434  
Canale, M. 547  
canonical word order 242  
Cardoso, W. 259  
Carlisle, R.S. 272, 273, 274  
“Carolina Project, The” 211–213  
Carpenter, H. 99  
Carrasco-Ortiz, H. 394  
Carreira, M. M. 160  
Carrell, P. 509  
Carruthers, J. 344  
Casillas, G. 80  
CCSARP (cultural study of speech act realization patterns) 470, 479  
CEFR (Common European Framework of Reference for Languages) 137, 222, 278, 279, 353, 427  
Celce-Murcia, M. 143  
Central Intelligence Agency (CIA) 201  
Cerezo, L. 200, 203, 204, 206, 210, 214  
Chamot, A. U. 221  
Chang, A. 287  
Chang, A, C.-S. 406  
Chang, P. 23  
Cheng, A. C. 350, 353, 355  
Chinese language/speaking  
Amoy Chinese 271  
compared to English 122  
Critical Period (CP) research 53  
English as a foreign language (EFL) Chinese students 99  
expository writing 123, 124  
generative approaches to second language (L2) acquisition (GenSLA) 85  
individual differences in advanced proficiency 158, 161  
parsing processes 32  
prosodic word adjunction, lacking 85  
suprasegmental attainment/learning 286  
tense 122  
theme 124  
writing proficiency in L2 Chinese 123  
Choi, H. 101  
Chomón Zamora, C. 187  
Chomsky, Avram Noam 15, 37, 77, 265, 328  
Chung, E. 449  
Clahsen, H. 40, 242  
Claremont Colleges Library (CCL) 465  
Clark, H. H. 510  
clause final filled pauses 228  
cleft constructions 433  
clitic left dislocation (CLLD) 428–430  
clitics and topicalization 455  
cloze-tests 318  
cluster analysis 169–173, 175  
Akaike’s Information Criterion (AIC) 165  
Bayesian Information Criterion (BIC) 165  
hierarchical clustering 164  
on-heritage speakers 170, 171  
onset cluster lexical decision task 250  
TwoStep clustering algorithm 164, 165, 173  
video watching 169–170, 174, 175  
Coffin, C. 21  
cognition, speech attainment, advanced level 296  
Cognition Hypothesis 141, 220, 233  
Cohen, A. D. 558  
coherence, contextual generic 21  
collegiate foreign language learners, advanced reading proficiency in see reading  
Collentine, J. 140, 354, 357n3  
collocations and lexis 61–64, 65f, 66  
Colombi, C. 16, 18, 19  
Comajoan, Ll. 375  
Common European Framework of Reference for Languages (CEFR) 137, 222, 278, 279, 353, 427, 465  
communicative language teaching 219  
Competing Systems Hypothesis (CSH) 86  
comprehensibility 245, 246, 248, 251, 259  
compression 307  
computer-assisted instruction (CAI) 211  
computer-assisted language learning (CALL) 160, 204, 210  
computer-assisted visualizations 336  
computerized formats 126  
computer-mediated communication (CMC) 210  
computer-mediated contexts (CMC) 95  
see also synchronous computer-mediated communication (SCMC)
Concept-Based Instruction (CBI) 369, 370–372, 373
conceptual refiguration 17
Conceptualizer stage, task condition effects 222, 223, 228, 229
concordancers 411
congruent semiosis 148
connected speech 304–322
see also connected speech processes (CSPs) and advanced L2 learners 304–306
assimilation 308
effects of training and individual differences 313–316
as ‘everyday speech’ 312
flapping 308
further research 316–319
glottalization 308
individual differences
and effects of training 313–316
research suggestions 317–318
motivation and challenges 310–313
palatalization 308
reduced forms of speech 306
reductions 306
research suggestions
individual learner differences 317–318
instruction and training 316–317
methods and tools 318–319
sandhi variations 306
term ‘advanced,’ notions about 309–310
terminology 308
training of L2 learners 310–316
weak forms 306
connected speech processes (CSPs) 306–309, 315
deletion CSPs 307
further research 316–318
insertion CSPs 307
linking CSPs 307
manipulation for humorous effect 319n1
modification CSPs 307, 308
multiple processes, CSPs 307
reduction CSPs 307, 308
Connor, U. 527–528
consonants 290
consonant clusters 243
intervocalic consonant (VCV) 308
word-final voiceless 308
contexts
computer-mediated 95
culture 11
differentiation among 12–13
instructed learning 136–137
learning, defining 204–205
rhetoric 540–541
secondaries 13
speaking, and markedness 274–275
strata 11
and text, complementary of 11–12
Contrastive Analysis Hypothesis (CAH) 270, 271
Conversation Analysis (CA) 517–518
Conway, L 450
Cook, M. 510
Cooper, M. M 535–536
Coppieters, R. 253, 364, 443
Corder, S. P. 265
corrective feedback 96, 98–100
Correia, S. 256
Coughlin, C. 43
Council of Europe 319n2
CP see Critical Period (CP) research
Crain, S. 450
Cramer Scalz, T. R. 32
Critical Period (CP) research
and advanced L2 proficiency 51–71
age effects and advanced L2 proficiency 51, 52–53
age of onset (AO) see age of onset (AO), CP
collocations and lexis 61–64, 65, 66
constitution of innate language faculty 87n2
core overextension 62
core preferences 62
Critical Period Hypothesis (CPH) 54, 61, 64, 66, 72, 79, 292, 336
early results 53–54
early-childhood learners compared with NSs 55–57
fossilization literature 68
further research 68–69
grammaticality judgment tests (GJTs) 53, 54, 55
items resistant to learning, at advanced levels 67–68
Kent–Rosanoff word-association task 61, 63
L2 learning at advanced proficiency levels 66–67
late-childhood L2 learners 57
length of residence 64
maturational constraints 52, 68
morphology and syntax 54–57, 58, 59–61
multi-word unit completion 62
multi-word unit correction 62
multi-word unit transformation 62
multi-word unity supply 62
native speakers (NSs) 52, 53, 55–57, 61–64
non-native speakers (NNSs) 54, 61, 63
peak sensitivity, critical period of 51
plural nouns 53
re-analyses 54–55, 57
relationship with advanced L2 proficiency 51–53
reliability of tests 63
salience 53, 54
sensitive periods 51, 52
summary of findings 66–67
terminology 87n1
third-person singular present 53
unmarked word order 59–60
vocabulary 61
word discrimination 62
Crivos, M. 386
Crossley, S. A. 145
cross-linguistic L1/L2 distance 289
cross-sectional studies
pragmatic competence 505, 507, 521
suprasegmental attainment/learning 286
systemic functional linguistics (SFL) 21
cultural study of speech act realization patterns (CCSARP) 470, 479
culture, context of 11
Cumming, A. 15
Cunnings, I. 37, 39
curriculum
foreign language, instructed SLA in 203–211
proficiency-based curricular change, example 211–213
SFL-inspired 22–24
DA (Dynamic Assessment) 115, 120, 125–127, 128
Danelund, L. 406
D’Angelo, F. 534
Dao, P. 102, 105
Darcy, L. 249
Davidson, D. E. 159, 167, 173, 187
De Jong, N. H. 229
decomposition 23
Defense Language Proficiency Test (DLPT) 201
DeKeyser, R. M. 54, 139, 434
Dekydtspotter, L. 80, 385, 451–452
DELE (Diploma of Spanish as a Foreign Language) 183, 186
deletion 245t
dental stops 248
depth of processing 208–210, 213
Derakhshan, A. 477
Derwing, T. 245, 247, 248, 252, 284, 286
descriptive statistics 167, 168t
Devrim, D. Y. 23
Dewaele, J.-M. 252, 298
dialectical materialism 114
Diaubalick, T. 367
Diaz-Campos, M. 560n6
dictation tests 318
Dimitrakopoulou, M. 80
Dippold, D. 519
disambiguating verbs 32
discourse
advancedness as a discourse phenomenon 143–146
interface with semantics at advanced proficiency 452–456
binding and co-reference 454–455
clitics and topicalization 455
interface phenomena in SLA 456
overt vs. null subject interpretation 453–454
knowledge 489–490
secondary contexts 13, 16
shift to 143
syntax-discourse interface 420–430
information structure at 419–420
native-like attainment at 430, 431t
discourse completion task 468
discourse semiotics 148
distance 13
Doiz, A. 365, 366
Domain by Age model 389
Dominguez, L. 425, 426
Donaldson, B. 430
Dörnyei, Z. 160–161
Doughty, C. 205
Doval Suarez, S. M. 433
Drager, K. 549
Dupoux, E. 255, 256, 257
Dussias, P. E. 32, 35, 40–41
Dutch language/speaking, parsing processes 38–39
Dynamic Assessment (DA) 115, 120, 125–127, 128
Index

Eckman, F. R. 273–274, 277
Edbauer, J. 534, 536
Edmonds, A. 354
Eefting, W. 326, 330
Egyptian Arabic 271
Eimas, P. D. 326
EIT (Elicited Imitation Task) 187, 192–193
electroencephalogram (EEG) studies 41, 45n2
Elicited Imitation Task (EIT) 187, 192–193
Ellis, N. C. 189
Ellis, R. 205, 223–226, 314–315
English as a foreign language (EFL) 99, 223, 413, 472, 508
Test of English as a Foreign Language (TOEFL) 97, 137, 311, 465
English language/speaking articles 140
compared to Chinese 122
consonant clusters 243t
Cree speakers of 251
exemplar sentences 121
expository writing 123
Hebrew speakers 61
Japanese speakers 17
Vietnamese speakers 20
epenthesis repair strategies 245t
Ernestus, M. 316
Escalona Torres, J. 549
Eslami, Z. R. 474, 476, 477
Eslami-Rasekh, Z. R 476
Estremera, R. 183
event-related brain potentials (ERP) 394
experiential grammatical metaphor 14, 16–17, 21
expository writing 123, 124
eye-tracking 31, 38, 40, 41, 279
face-to-face (FTF) interaction 103, 104
Fafulas, S. 551, 552
faithfulness 273
Faudree, M. 271
FBI 201
feature competition model, phonology 244–245
feature interpretability, inflectional morphology 387–388
feature reassembly 444
Feature Reassembly Hypothesis 85, 367
feedback corrective 96, 98–100
explicit 99
immediate 99
implicit 97, 98, 99
metalinguistic 96, 99
360-degree feedback 202
Felser, C. 32–33, 37, 39, 40
Feuerstein, R. 119, 123
Filiaci, F. 452, 453
filler-gap constructions 33, 37
Finer, D. 243
Finnish language/speaking 256
Fiorentino, R. 395
first language (L1) beyond L1 transfer 82–86
transfer, role of 32, 78, 80–82
Flege, J. 248, 251, 282, 285, 289, 290, 292, 294, 326, 330, 331
fluency 252, 309
see also CAF (complexity, accuracy and fluency)
lexical development 410–411
Foote, R. 43
Fordyce, K. 478
Foreign Service Institute, School of Language Studies 200
formal disciplines, theory of 118
formal education 183
formal knowledge 531
formulaic language, advanced lexical development 404–405
Formulator stage, task condition effects 222, 228
Forsberg Lundell, F. 405
fossilization 68, 83, 136–137, 138
Foster, P. 230, 231, 232
Fourakis, M. 332
Fowler, C. A. 330
Franceschina, F. 387, 391
Frawley, W. 115
Freed, B. F. 286
French language/speaking connected speech 308
Hexagonal French 354
individual differences in advanced proficiency 158, 161
mood distinction 343–345
Frenck-Mestre, C. 41, 394
frequency distributions 165–167
Frota, S. N. 207
Fukuta, J. 228, 229
Fukuya, Y. J. 474, 475
functional load, phonology 243–244
functional magnetic resonance imaging (fMRI) 279
functional morphemes 382, 383, 394
functional theory of language 10–11
further research
advanced proficiency and sociolinguistic competence 558–560
connected speech 316–319
Critical Period (CP) research 68–69
globalization demands, meeting 213–214
grammatical development, advanced-level (in instructed SLA) 149–150
heritage bilinguals, prior language experience 192–193
inflectional morphology 396–397
interaction-driven L2 learning 105–108
lexical development 412–413
markedness 279
phonology 258–259
pragmatic competence 521–522
pragmatics, advance-level (in ISLA) 478–480
sociocultural theory 127–128
speech attainment, advanced level 297–298
task condition effects 232–234
voice onset time (VOT) 335–336

Gabriele, A. 395
Gal’perin, P. Y. 122, 127
Garcia, P. 510–511
García Mayo, M. D. P. 100
‘garden-path’ sentences (ambiguous constructions) 31, 32, 33
Gass, S. M. 99, 190, 206, 229, 276, 334
Geeslin, K. L. 140, 141, 350, 351, 357n6, 357n7, 549, 551, 553, 554, 555, 557, 559
generative approaches to second language (L2) acquisition (GenSLA)
and advanced L2 proficiency 72–93
extraneous factors 73
Feature Reassembly Hypothesis 85
generative linguistic theory, main tenets 75–78
linking easy-to-observe and hard-to-observe properties 77
and logical problem of language acquisition 77
mind/brain representations 78
Missing Surface Inflection Hypothesis 84
Null Subject Parameter (NSP) 76, 81, 82
Overt Pronoun Constraint (OPC) 76–77
parameters of grammatical variation between languages 75
parametric line of thinking 77
Poverty of the Stimulus (PoS) 75, 77, 78, 80, 83, 87
Pro-drop Parameter 77
Prosodic Transfer Hypothesis 85
role of L1 transfer 80–82
Separation Hypothesis 84
ultimate attainment 78–86
Universal Grammar (UG), access to 79–80
genre-based writing instruction (GBWI) 541
genres
complementarity of context and text 11
complementarity of system and instance 12
genre analysis 143
genre-based tasks 24
grammatical metaphor, development 17
ideological 144
knowledge and awareness 541–542
orientation 144–146
primary and secondary 144
shift to genre 143
within a functional grammar 147–148
George, A. 552, 557
Georgetown University, Washington DC, German Program 23, 200
German language/speaking
Critical Period (CP) research 67
markedness 274–275
parsing processes 34
relativization 139
Gertken, L. 253
Gilbers, S. 290
Glass, W. R. 80
globalization demands, meeting 199–218
example of a proficiency-based curricular change 211–213
further research 213–214
instructed SLA 200, 203–211
societal demands for advanced proficiency and academia’s response to 201–203
GM see grammatical metaphor (GM)
Goad, H. 85, 86, 331
Gökgöz-Kurt, B. 315
Goldschneider, J. M. 54
Gonzalez Alvarez, E. 433
González López, V. G. 331
Goodman, K. S 484–485
Google Scholar 183
Goro, T. 450–451
Gradual Learning Algorithm 328
grammatical metaphor (GM) 13
anaphoric reconstruction 17
development of 17
experiential 14, 16–17, 21
functional theory of language 148
ideational 22
interpersonal 14, 18–20
logical 14, 16–17
subjective interpersonal 19
types 14
grammaticality judgment tests (GJTs) 252
Critical Period (CP) research 53, 54, 55
mood distinction 346, 348
grammatics 134
Granena, G. 55, 57, 59, 60, 64, 65t, 296
Greenberg, J. 266, 278
Gregg, K. 258
Grice, P. 513
Grosjean, F. 180
Grüter, T. 450, 451
Gu, Y. 223
Gualmini, A. 450, 451
Gudmestad, A. 351–354, 356, 357n6, 357n7, 357n9, 551, 553, 554, 557
guided induction 373–374
Guijarro-Fuentes, P. 353, 367
Guion, S. 248, 291
Guiraud Index 229
Gundel, J. 274
Gurtin, S. 331

habituality 365–367
Hakuta, K. 73
Halliday, M. A. K. 10, 11f, 134, 144, 145
Han, Z.-H. 68
Hancin-Bhatt, B. 244
Hanulíková, A. 547
Hanzawa, K. 298
Harrington, M. 32
Harrison, R. 103
Harsch, C. 127–128
Hasan, R. 145
Hasler-Barker, M. 558
Haspelmath, M. 77
Hatch, E. 205
Hawkins, R. 80, 286
Hebbian learning 332
Hedgcock, J. 98
Hellmuth, S. 256, 257
Hempel, C. G. 277
Henrichsen, L. E. 310, 311
Index

Henshaw, F. 191
heritage bilinguals
  advanced proficient heritage speakers 187–188
  further research 192–193
heritage language learners (HLLs) 160, 179–182, 187, 188, 190–192
heritage speaker benchmark 251
HL speakers as additional language (Ln) learners 191–192
pedagogical interventions 188–191
prior language experience of 179–198
individual differences 181–183
proficiency measurements 183, 184–185, 186–187
heritage language (HL) 181
Heritage Language Journal 183
heritage language learners (HLLs) 160, 179–182, 187, 188, 190–192
heritage speaker status 159–162, 166, 170, 176
  heritage speaker benchmark 251
heritage status 159–160, 161
Hertel, T. J. 424, 425
heteroglossic statements 19, 20
Hewings, A. 21
high variability input 295
High Variability Phonetic Training (HVPT) 295
Higher Education Bureau Examination 223
Hilferty, A. 313
Hilles, S. 82
Hindi retroflex 248
Ho, V. L. 20
homographs, interlingual 253
homophones, interlingual 253–254
Hopp, H. 42, 43, 392, 393, 395
Housen, A. 142
Howard, M. 346, 347, 348
Hsieh, H.-C. 211
Hsu, H. 225, 226
Hulstijn, J. H. 182, 226, 345
Hulstijn, W. 226
Hungarian language/speaking 256, 257
HVPT (High Variability Phonetic Training) 295
Hwang, S. 444
Hyams, N. 77
Hyland, K. 537
Hyltenstam, K. 52, 247, 252
Hymes, D. 463
ideational distance 13
ideational metafunction 10
IELTS (International English Language Testing System) 21, 241
ILR (Interlanguage Round Table) 279
Imai, S. 285
imperfective, as aspectual marker 365
implicit feedback 97
IMRD (Introduction, Methods, Results, and Discussion) format, lab reports 539
incongruency, notion of 13
incremental processing, L2 parsing 31–36
  complex sentences, L1/L2 differences 35, 36
independent construction 23
indirect meaning, comprehension of 506–512
individual differences in advanced proficiency
  abroad experience 159, 161, 162, 166
  connected speech
    research suggestions 317–318
    training effects 313–316
  heritage speaker status 159–162, 166, 170, 176
  L2 speech attainment 292, 294–296
  literature review 157–158
  motivation 160–173
  prior language experience 181–183
  research method
    data analysis 164–165
    participants and materials 161–164
  research results
    cluster analysis 169–173
    continuous survey variables and score profiles 167, 168f
    descriptive statistics 167, 168f
    frequency distributions 165–167
    use of authentic L2 resources outside of class 160, 161
Infante, P. 123
inflectional morphemes 189, 396
inflectional morphology 381–400
see also morphology
  dissociations between representation and processing 392–396
  feature interpretability 387–388
  function/place in language faculty 382–388
  further research 396–397
  grammatical and semantic features expressed by 382–384
in advanced SLA 419–441

cleft constructions 433
combining experimental and corpus evidence 435–436
functional approaches 430, 432–435
grammatical variation 434–435
information structure at syntax-discourse interface 419–420
and left periphery 427–428
and prenominal subject distribution 421–423
raising and alternations 434
subject-verb inversion 423–427
transfer of principles in advanced stages of acquisition 432–433
instructed second language acquisition (ISLA) 159
advanced-level grammatical development see grammatical development,
advanced-level
context of instructed learning 136–137
and Critical Period (CP) research 69
depth of processing in, controlling for 208–210, 213
in a foreign language curriculum 203–211
genre orientation, implications for advanced instructed language learning 144–146
and globalization demands see globalization demands, meeting
higher forms of consciousness, development through schooling 117–120
incorporating models in processing-based pedagogy 206–207
noticing, situating in the instructional design 207–208
phonology 247–248
pragmatics, advanced-level 463–482
further research 478–480
pragmatic competence and ISLA 464–465
pragmatic features and measures 468, 469t, 470–472
pragmatic instructions, effectiveness of 472, 473t, 474–478
research method 465–466, 467–468t results 468–477
redefining learning context 204–205
vs. second language acquisition 205
technology control 210–211
theory/practice relationship, institutionalizing 205–206
intelligibility 245–248
interaction and advanced learners 97–104
computer-mediated contexts (CMC) 95
corrective feedback and advanced proficiency learners 96, 98–100
face-to-face (FTF) 103, 104
further research 105–108
interaction-driven L2 learning 94–113
Interactions Hypothesis 94, 95
interlocutor proficiency, impact on LREs 101, 102
key elements of approach 95–96
Language Related Episodes (LREs) 100–102, 104, 105
low-proficiency learners 99
mixed-proficiency 102
negotiation 96
pair dynamics 101–102
pedagogical implications of interaction-driven L2 learning 104–105
peer interaction and advanced proficiency learners 100–103, 105
task-based language teaching (TBLT) 95, 100, 104
teacher–learner, as mediation for L2 proficiency 125–127
and technology 103–104
intercultural competence 128
Interface Hypothesis 452–454
in advanced SLA 420–430
generative approaches to second language (L2) acquisition (GenSLA) 86
interlanguage (IL) 95, 96, 265–266, 419
Interlanguage Round Table (ILR) 279
internalization 116–117
International English Language Testing System (IELTS) 21, 241
interpersonal distance 13
interpersonal metafunction 10
Interpretability Hypothesis 367, 395
intervocalic consonant (VCV) 308
intransitive verbs 32
Introduction to Instructed Second Language Acquisition (Loewen) 200
inverse-scope reading 448
inversion 243
Ionin, T. 386, 445
Isaacs, T. 252
Isabelli, C. A. 347, 348
Isabelli-Garcia, C. 557
ISLA see instructed second language acquisition (ISLA)
island effects 37–38, 45n1
iterativity 365–367
Ivanov, I. 429
Iverson, G. K. 273–274, 332
Iverson, M. 77, 80, 350
Iwashita, N. 101, 102
Jackson, C. N. 33–34, 35
Jaeggli, O. 77
Japanese language/speaking adult learners of English 17
past tense feature 85–86
Jenkins, J. 319n3
Jeon, E. H. 464
Jesney, K. 251
Jessner, U. 374
Jiang, N. 393, 396
Johns, A. M. 541
Johnson, M. 122, 224
joint construction 23
Joyce, P. 311
Judy, T. 76, 82
Juffs, A. 32
Kaan, E. 45n3
Kagan, O. 160, 182
Kaivanpanah, S. 100
Kang, H. 191
Kanno, K. 80
Kanwit, M. 350, 351, 352, 354, 359
Kao, Y.-T. 123–124
Kaplan, R. B. 123
Karpov, Y. V. 118–119
Kaya, T. 464
Keating, G. D. 43
Keating, P. A. 326, 393
Kempchinsky, P. 428–429, 455
Kennedy, S. 314, 315
Kent–Rosanoff word-association task 61, 63
Kijak, A. M. 257
Kim, E. 272, 455
Kim, H.-Y. 315
Kim, Y. 101, 102
Kinneavy, J. L. 534
Kissling, E. 331
Kitade, K. 104
Klein, W. 253, 362, 373, 374
K-Means clustering 164
knowledge advanced 242–245
aspectual 374
advanced 369
developing advanced conceptualization of knowledge 370
problems with advanced conceptualizations (CBI and CSH) 370–372
discourse 489–490
formal 531
functional morphology, semantics reflexes of 385–387
genre 541–542
grammatical, applying during on-line processing 35–38
linguistic, contributions to reading operations 487–490
metalinguistic 74, 83, 101, 186
morphological 488
orthographic 487
passive 310
phonological 242–245, 487
feature competition model 244–245
functional load 243–244
target-like 247
procedural 119
process 531
rhetorical 531–542
vs. skill 242
subject matter 531
syntactic 489
theoretical 119
vocabulary 488–489
Ko, H. 445
Kochetov, A. 331
Kohler, K. J. 306, 307
Kondo-Brown, K. 160
Korean language/speaking 274, 286, 449
Kormos, J. 228
Kostin, I. 311
Kouwenhoven, H. 316
Kowal, M. 98
Kozulin, A. 119, 120
Kuhl, P. K. 326, 329, 333
Kupisch, T. 86, 245, 251

Lafford, B. 357n3
Lahmann, C. 292, 294
Lai, W. 122, 123
Lakoff, G. 122
Lambert, C. 228, 229
Lancaster, Z. 20
Langacker, R. 366, 376n5
Langman, J. 560n4

language
aspiring languages 326
categorization of types for number marking 268t
changing views of proficiency of in L2 research 484
data 373–374
dominance 182
evaluative superstructure of 269
formulaic 404–405
functional theory of 10–11
interlanguage (IL) 95, 96, 265–266
metafunctions 10
meta-language 23
natural functional theory of 148–149
phonetic inventory 269
proficiency in, and reading ability 484–485
role in communication 11
SFL-based assumptions about 10–15
switching 253–255
systemic model 15
target language (TL) 96
usage-based theories of language development 12
voicing languages 326
language dominance 182
Language Flagship Program 200–201
Language Hiring Bonus Program 201
Language Proficiency Flagship (NSEP) 161
Language Related Episodes (LREs) 100–101, 102, 104, 105
Language Testing International (LTI) 161, 162
language-focused learning, lexical development 409–410
Lantolf, J. P. 115, 120, 121, 369, 370, 371
Lapkin, S. 101
Lardiere, D. 258, 384, 385, 396, 444
Laskin, L. 432

Lasnik, H. 78
Laufer, B. 61, 402–404
law of economy 306–307
Leal, T. 549, 555, 559
Leal Méndez, T. 187
learner corpus research (LCR) 435, 436
learner language 265
Lee, E. J. 100
Leeser, M. J. 101
left periphery 427–428
Lekie, M. D. 187
Lemmet, N. 395
length of residence (LoR) 282, 291, 330
Lenneberg, E. H. 332
Leontjev, D. 126
Leow, R. P. 200, 203, 204, 206–207, 208, 210, 212, 214
Levelt, W. J. M. 222, 227, 233
Levi, T. 126
Levis, J. M. 304, 307, 308
Lewis, K. 190
lexical deficits 64
lexical density 16, 24n2
lexical development, advanced 401–418
corpus-driven studies 403
defining 401–406
fluency 410–411
formulaic language 404–405
further research 412–413
helping learners 406–411
language-focused learning 409–410
meaning-focused inputs 407–408
meaning-focused outputs 408–409
next steps 411–412
processing 405–406
sophistication 404
vocabulary learning program, four strands of 407
vocabulary size 401–403
lexical profiling studies 402
lexical sophistication 404
lexical verbs 34
lexis, and collocations 61–64, 65t, 66
Li, S. 98–99, 508
Li, X. 552
Liardet, C. L. 17, 18
Liddicoat, A. 510
Lidz, J. 78
Lieberman, M. 450, 451
Limited Attentional Capacity Model 141, 220, 233
Lin, Y-H. 98
Lindqvist, C. 405
Linford, B. 554
*Linguistic and Language Behavior Abstracts*
database 183
linguistic universals 266
Lisker, L. 325–326
Listening Proficiency Test (LPT) 162
literacy/literacies 13
Liu, Y. 508
lived experience 128
LoCastro, V. 505
Loewen, S. 200, 204, 207
Logan, J. S. 295
logical grammatical metaphor 14
Log‐likelihood 165
Long, M. H. 57, 59, 60, 64, 65t, 68, 94, 106, 189, 205, 207, 210, 296
longitudinal studies
pragmatic competence 507
systemic functional linguistics (SFL) 16, 17, 20
voice onset time (VOT) 335
Lozano, C. 422, 424, 426
LTI (Language Testing International) 161, 162
Lucy, P. 510
Luria, A. R. 116
Lybeck, K. 557
Lynch, T. 228
McAllister, R. 290
McCandliss, B. D. 295
McCarthy, C. 388, 389
Macdonald, D. 102, 105
McDonough, K. 102, 105
MacKay, I. R. 251, 289, 292
Mackey, A. 95, 96, 99
McLaughlin, B. 207
McLean, J. 228
McLeod, B. 207
McManus, K. 346–349
McNamara, D. 145
magnetoencephalography (MEG) studies 254
Major, R. 271, 272, 287
Malovrh, P. A. 202, 211–213
Mandarin 267, 268, 271, 284, 387
*see also* Chinese language/speaking
Cantonese learners 287
Manzon-Omundson, S. 512
Marinis, T. 36, 37
markedness
and advanced development 264–281
background 265–269
as evaluative measure in L2 phonology 269–278
as evaluative measure of advanced development 278–279
as evaluative measure of IL grammar 273–274
as evaluative measure of phonological environments/speaking contexts 274–275
faithfulness 273
further research 279
interlanguage 265–266
L2 phonology 265, 269–278
as learning difficulty in L2 phonology 270–273
typological 264, 266–269, 275–278
Markedness Differential Hypothesis (MDH) 269, 270, 271, 273
Martin, J. R. 12
Martinez-Flor, A. 475
Martins, C. 367
mass/count distinction and plural interpretation 444–445
Master, P. 140
Mastropavlou, M. 387–388
Matsuzawa, T. 313, 318
Matthiessen, C. M. I. 11f
McClelland, J. L. 487
McDonough, K. 101
MDH (Markedness Differential Hypothesis) 269–271, 273
meaning-making
congruent vs. incongruent 12–15
metaphorical style of 15
Meara, P. 404, 413
Mediated Development (MD) 115, 121–124
Mediated Learning Experience 119, 120
Medina-Rivera, A. 355
Meisel, J. M. 67, 88n3, 242
Menendez-Benito, P. 376n4
Menke, M. 512
Mennen, S. 247
Mercado, L. 224
“Merge Files” method 162
meta-language 23
metalinguistic awareness 373–374, 376n7
metalinguistic explanations 98–99
metalinguistic feedback 96, 99
metalinguistic knowledge 74, 83, 101, 186
metaphors
  faded 17
  fresh 17
  grammatical metaphor (GM) see grammatical metaphor (GM)
  time, spatial metaphors for 122
methodological principles (MPs) 205
MFH (Multiple Feature Hypothesis) 327
Michaelis, L. 362
Michigan State University Flagship team 162
Miller, D. 396
Miller, J. D. 326
Miller, R. 119
Mind in Society 114
Minimal Sonority Distance (MSD) 243
Minimalism 369, 371, 382
Minn, D. 228
Miralpeix, I. 413
Missing Surface Inflection Hypothesis (MSIH) 84, 385, 388, 396
Mitchell, R. 346–349
MLA (Modern Language Association) 200, 201
Modern Language Association (MLA) 200, 201
Moeller, A. J. 175
Moffett, J. 534
Mohan, B. 24
Mojica-Diaz, C. C. 350, 353, 355
monoglossic statements 19
monolinguals 288–289
Montalbetti, M. 76
mood distinction
  academic year abroad 346–349
  advanced-level 343–360
  and advanced-level learners 345–346
  established proficiency metric 353–355
  in French and Spanish 343–345
  graduate-level NNSs 349–353
  near-native speakers 355
Mooi, E. 165
Mora, J. C. 286
Moranski, K. 376n8
Morcillo Gómez, J. 206
Moreno, N. 204, 210–213
Morgan-Short, K. 208
morphemes 53, 446, 488
  free 382
  functional 382, 383, 394
  grammatical 54
  inflectional 189, 396
  number 386
Morphological Congruency Hypothesis 393, 394
Morphological Underspecification Hypothesis (MUH) 388, 396
morphology
  inflectional see inflectional morphology
  morphological knowledge 488
  morphological regularity 53
  semantic reflexes of functional
    morphology knowledge 385–387
    and semantics 444–446
  grammatical and semantic features expressed 382–384
  inflectional morphology 382–387
  in nominal domain 446
  semantic reflexes of functional
    morphology knowledge 385–387
    and syntax 54–57, 58t, 59–61
  morphology-before-syntax 84
  syntax-before-morphology 84, 384–385
morphosyntax 140, 252, 443
motivation
  connected speech 310–313
  individual differences in advanced
    proficiency 160–173
  speech attainment, advanced level 296
Motor Theory of Speech Perception 329
Moyer, A. 253
MSD (Minimal Sonority Distance) 243
MSIH (Missing Surface Inflection Hypothesis) 84, 385, 388
MUH ( Morphological Underspecification Hypothesis) 388
multicompetence 12
multilingual turn 137
multilingualism 173, 253
multiple analyses of variance (MANOVAs) 213
Multiple Feature Hypothesis (MFH) 327
Munro, M. 245, 247, 251, 252, 284, 289, 292
Muradás-Taylor, B. 256
Nagle, C. 331
Nagy, N. 331
Nagy, W. E. 488, 489
Nakayama, M. 254
Index

Nathan, G. S. 331
Nation, P. 403, 404, 407, 413
National Council for Accreditation of Teacher Education (NCATE) 203
National Heritage Language Resource Center 186
National Security Education Program (NSEP), Language Proficiency Flagship 161
Native Language Magnet (NLM) 329, 333
native speakers (NSs)
see also near-native speakers; non-native speakers (NNSs)
advanced skill 245
Critical Period (CP) research 52, 53, 55–57, 61–64
vs. late L2 learners 44
mood distinction 344, 349
psycholinguistic approaches 33, 34, 37
native-likeness 242, 251, 258, 282, 291
Nativist Hypothesis 336
naturalistic learning 135
near-native speakers
see also non-native speakers (NNSs)
mood distinction 355
phonology 252–253
near-native-likeness 289
negative evidence 96
negative polarity item (NPI) 450
negotiation, for meaning 96
Negueruela, E. 115, 121–123, 370, 371
neighbourhood density 298n1
Nelson, C. L. 531
Nemser, W. 265
Newmeyer, F. J. 77
Nguyen, T. M. H. 413
Nguyen, T. T. M. 474
Ni, L. 450
Niedzielski, N. 549
Nishida, C. 347, 348
NLM (Native Language Magnet) 329, 333
NNSs see non-native speakers (NNSs)
nominal domain, semantics of at advanced proficiency
see also semantics
articles
definiteness and specificity 445
genericity 445–446
mass/count distinction and plural interpretation 444–445
and morphology 446
nominal inflectional morphology 390–392
nominal subordinate clauses 347
nominalizations 16, 17
non-native speakers (NNSs)
Critical Period (CP) research 54, 61, 63
graduate-level 349–353
heritage speaker benchmark 251
mood distinction 346
morphology and syntax 54
Norris, J. M. 142, 158, 174, 176, 314
Norton, B. 251, 298
Nota, A. 290
noticing hypothesis 207, 209
Noun Phrase Accessibility 139
noun phrases (NPs) 34–36, 39–40, 148
see also parsing processes (L2);
psycholinguistic approaches
ambiguous 32, 33
plausible ambiguous 33
NPI (negative polarity item) 450
NPs see noun phrases (NPs)
NSEP see National Security Education Program (NSEP)
NSs see native speakers (NSs)
Null Subject Parameter (NSP) 76, 81, 82
oblique–subject alternations 434
obuchenie (teaching/learning) 118, 120
O’Malley, J. M. 221
on-line planning
see also pre-task planning with high-proficiency learners
defining 224–225
hybrid 225–226
operationalization and effects 225–226
theory 224–225
unpressured 225
on-line processing, applying grammatical knowledge during 35–38
onset cluster lexical decision task 250
Ontogeny Phylogeny Model (OPM) 272, 287, 329
OPM (Ontogeny Phylogeny Model) 272, 287, 329
Oppenheim, P. 277
Optimality Theory (OT) 272, 328
Oral Proficiency Interview (OPI)
( ACTFL) 97, 183, 186, 201, 309
see also American Council of Teaching Foreign Languages (ACTFL)
Oral Proficiency Interview Computerized (OPIC) 161, 186
Orosz, A. 406
orthographic knowledge 487
OT (Optimality Theory) 272, 328
Otálora, M. 391
Overt Pronoun Constraint (OPC) 76–77
Oxford, R. 221

Palmer, A. S. 464
Pang, F. 222
Papadopoulou, D. 40
parametric models 77, 243, 382
Park, H. 249

parsing processes (L2)
see also psycholinguistic approaches
‘garden-path’ sentences (ambiguous constructions) 31–33
incremental processing 31–36
island effects 37–38, 45n1
on-line processing, applying grammatical knowledge during 35–38
lower proficient L2 learners 35
misanalyses 31, 33–34
noun phrases (NPs) 32–36
pronominals 31, 38
real-time parsing 35–36
referential processing, and advanced proficiency 38–39
relative (RC) attachment ambiguities 39–41
structural attachment preferences 31
subcategorization 32
syntactically complex constructions 31
Pascual y Cabo, D. 183
past simple 34
Pater, J. 258, 331
Pawley, A. 15

pedagogy
and aspect 369–372
interaction-driven L2 learning, pedagogical implications 104–105
prior language experience of heritage bilinguals 188–191
processing-based 206–207
Reading to Learn (SFL-based) 23
systemic functional linguistics (SFL) 22–24
Pederson, E. 248
peer interaction and advanced proficiency learners 100–103, 105
Pekarek Doehler, S. 519

Peperkamp, S. 256
perception, advanced 245
Perceptual Assimilation Model 329, 333
Perceptual Difficulty Hypothesis 139
Perez-Leroux, A. T. 80
Perfetti, C. A. 229
Philp, J. 99
Phinney, M. 82
phonemes 244, 328, 333, 335
phonetic identity, loss of 305
phonetics
vs. phonology 246
target-like 247
phonological knowledge 242–245, 487
feature competition model 244–245
functional load 243–244
target-like 247
phonology
‘advanced,’ locus of 242–246
advanced knowledge 242–245
feature competition model 244–245
functional load 243–244
and target-like phonology 247
advanced perception 245–246
advanced production 245
advanced skill 245–246
advanced-level second language 241–263
attention, role of 248–249
comprehensibility 245, 246, 248, 251, 259
encapsulation of 252
factors influencing outcomes 247–251
further research 258–259
global accent 251
heritage speaker benchmark 251
instruction, effects of 247–248
intelligibility 245–248
language switching 253–255
markedness and L2 phonology 265, 269–278
and morphosyntactic ability 252
native-likeness 242, 251
near-nativeness 252–253
perception 245–246
vs. phonetics 246
phonological knowledge see phonological knowledge
phonological salience 53
processing 250
production 245
prosody, advanced 255–258
solidarity aspects 250–251
target-like 247, 285
translation equivalents 254
working memory 249
Pica, T. 100
Pienemann, M. 139, 242
Pinker, S. 304
Plonsky, L. 44
Pochon-Berger, E. 519
Poehner, M. E. 115, 120, 121, 125, 127–128, 371
Polinsky, M. 159, 182, 188
Polio, C. 140, 202
Polish language/speaking 256, 257
postverbal subjects 426
Poverty of the Stimulus (PoS)
generative approaches to second language acquisition (SLA) 75, 77, 78, 80, 83, 87
inflectional morphology 386
overcoming 447–452
disjunction, scope of 450–451
inverse-scope reading 448
PoS argument in L2 acquisition 447–448
quantifiers, scope of 448–450
scope ambiguity and discontinuity 451–452
scope and the PoS problem 452
surface-scope reading 448
wh-expressions, scope of 448–450
semantics, advanced-level 443, 447–452, 456
PPP (present-practice-produce) 209
pragmatic competence
see also pragmatics, advance-level (in ISLA)
advanced L2 505–526
further research 521–522
indirect comprehension, characteristics of advanced competence in 508–512
indirect meaning, comprehension of 506–512
review methods 507–508
speech acts, production of 513–521
advanced competence characteristics 515–520
linguistic strategies of requests 515–517
review methods 514–515
sequential organization of requests 517–520
pragmatics
advanced-level see pragmatics, advance-level (in ISLA)
defining 505
pragmatic competence see pragmatic competence
pragmatic plausibility 32
pragmatics-inspired approaches to language development 15
pragmatics, advanced-level (in ISLA) 463–482
see also pragmatic competence
further research 478–480
mixed findings 475–476
pragmatic competence 464–465
pragmatic instructions, effectiveness of 472, 473t, 474–478
explicit and implicit instruction as a continuum 476–477
explicit vs. control 472, 474
explicit vs. implicit 474–475
implicit vs. control 474
research method 464–466, 467–468
Prague School of Linguistics (1930s) 266
praxis 120
Predicate Proximity 40
prenominals 31, 421–423
present perfect 34, 141
pre-task planning with high-proficiency learners
see also task condition effects
Articulator stage 223
behavior of advanced learners 221–222
effects on CALF 222–224
principle of least effort 306–307
Principles and Parameters Theory 328
prior language experience
of heritage bilinguals see heritage bilinguals
individual differences in 181–183
procedural knowledge 119
process knowledge 531
processing speed 249
processing-based pedagogy, incorporating ISLA models in 206–207
production, advanced 245
Productive Levels Test 403
prenominals 31, 38
pronunciation 214, 247, 253, 285, 290, 313–317, 324, 329, 331, 333, 557
accuracy 296
attainment 292
ease of 273
errors 287
forms 298
patterns 265, 272
pedagogic 259
pronunciation (cont’d)
performance 285, 287, 296
teaching of 336
term ‘advanced,’ notions about 309
Pronunciation in Second Language Learning and Teaching (PSLLT) 259
ProQuest 507
Prosodic Transfer Hypothesis 85, 385
prosody, advanced 255–258, 286
linguistic assumptions 258
prosodic word adjunction, Chinese language lacking 85
representation 257–258
stress deafness 255–257
psycholinguistic approaches 30–50
‘advanced late-learners’ 30, 44
agreement processing and advanced proficiency 41–43
electroencephalogram (EEG) studies 41, 45n2
eye-tracking 31, 38, 40, 41
on-line grammaticality detection 43
native speakers 33, 34, 37
parsing (L2) and advanced proficiency see parsing processes (L2)
real-time sentence processing studies 44
time-sensitive measures 31
types of L2 sentence processing studies 30–31
“psycholinguistic guessing game” model 485
psychology
general principles of 116
higher and lower psychological functions 116
scientific 114
psychophysical salience 139
Pullum, G. K. 77
Q_Lex test 413
Qian, D. D. 402
quantifiers, scope of 448–450
Quesada, M. L. 423
‘quotative-echoic’ interpretation 354, 357n10
raising constructions 434
Range software 404
Rankin, T 432
Rau, D. 287
Ravenhorst-Kalovski, G. C. 402
reading
advanced L2 reading proficiency in practice 490
analysis of advanced-level textbooks 495, 496–499, 500
analysis of proficiency guidelines 491, 492–494
changing views of ability 484–485
collegete foreign language learners, advanced reading proficiency 483–504
component skills of 485–486
inverse-scope 448
and language proficiency 484–485
linguistic knowledge, contributions to 485–486
operations, contributions of linguistic knowledge to 487–490
surface-scope 448
Reading Proficiency Test (RPT) 162
recasts (implicit forms of feedback) 97, 98, 99
Reduced Ability to Generate Expectations (RAGE) 31
Reeder, J. T. 331
register, context of 11
register analysis 145
Reinhardt, J. 175
relative clause (RC) attachment ambiguities 31, 39–41
relative clause (RC) islands 37
relativization 139
representation of aspect 362–364
inflectional morphology 392–396
prosody, advanced 257–258
resource-directing factors 220
resource-dispersing factors 220
resources, authentic 160, 161
response time 253, 254
rhetoric, advanced 527–546
audience considerations 538–540
CAF constructs 530
communicative purpose 537–538
context, concerning 540–541
contrastive rhetoric 123
future directions of efficacy-oriented, rhetorically-based SLA models 543
genre knowledge and awareness 541–542
knowledge 531, 532–542
North American approaches 146
overview 527–528
progressively scaffolded models of rhetorical knowledge 532–542
rhetorically based framework 531–532
rhetorically situating the writer 536–537
writing 528–529
Index 585

Rizzi, L. 77
Roberts, L. 32–33, 37, 38–39
Robinson, P. 141, 207, 220, 234
Rodriguez-Louro, C. 422
Rodriguez-Ordonez, I. 390
Roever, C. 519, 520, 523n1
Romance languages 364, 428, 446
Roscoe, R. 145
Rose, K. R. 516, 517
Rosen, C. 432
Rossman, T. 207
Rothman, J. 74, 76, 77, 80, 82, 245, 251, 369, 396, 428–429, 455
Russian language/speaking
advanced proficient heritage speakers 187
Czech speakers 251
heritage language learners (HLLs) 182
individual differences in advanced proficiency 158, 161
inflectional morphology 53
verbs 187
Ryshina-Pankova, M. 13, 17, 21
Sagarra, N. 40–41, 189
Saito, K. 286, 287, 291, 292, 294, 298
Salaberry, M. R. 367, 368, 375
salience 53, 54, 139–140
salient associations 139
SALT (suggestive accelerative learning and teaching) method 477
Santa Barbara Corpus of Spoken American English 506
Sanz, C. 142, 181, 189, 191–192
Sarstedt, M. 165
Sauro, S. 104
Savignon, S. J 405
Sawallis, T. 244
Sax, K. 554
scaffolding, textual 23
SCH (Structural Conformity Hypothesis) 273, 275, 276, 277
Schleppegrell, M. J. 18
Schmid, M. S. 290, 292
Schmidt, A. M. 330
Schmidt, R. 207, 209, 555, 559
Schmitt, N. 405–406
Scholz, B. C. 77
School of Language Studies, Foreign Service Institute 200
Schoonen, R. 345
Schumacher, N. 141
Schwartz, B. D. 78, 389, 447–448, 449, 450
SCMC (synchronous computer-mediated communication) 103–104, 106
scope interpretation at advanced proficiency disjunction, scope of 450–451
inverse-scope reading 448
Poverty of the Stimulus (PoS), overcoming 447–452
PoS argument in L2 acquisition 447–448
scope and the PoS problem 452
quantifiers, scope of 448–450
scope ambiguity and discontinuity 451–452
surface-scope reading 448
wh-expressions, scope of 448–450
Scovel, T. 250
SCT see sociocultural theory (SCT)
SDRH (Similarity Differential Rate Hypothesis) 269, 270, 272
Seaman, J. 204
Sebastián-Galles, N. 256
second language acquisition (SLA) advanced
functional approaches to word order and information structure in 430, 432–435
Interface Hypothesis 420–430
voice onset time see voice onset time (VOT)
age of first exposure to a new language and eventual attainment 51
Critical Period (CP) research see Critical Period (CP) research
and globalization demands 199, 200
information structure and word order see information structure (IS)
instructed see instructed second language acquisition (ISLA)
v.s. instructed second language acquisition 205
and markedness 264
and sociocultural theory 115
speech attainment see speech attainment, advanced level
Segalowitz, N. 286
segmental attainment/learning 250, 283–285, 297
product of advanced-level L2 speech attainment 288–291, 293t
Seidenberg, M. S. 487
selective fossilization 68
Selinker, L. 265, 276
Index

semantic priming 298
Semantic Subset Principle (SSP) 450
semanitics
advanced-level 442–460
and functional grammar 147
interface with discourse at advanced proficiency 452–456
binding and co-reference 454–455
clitics and topicalization 455
interface phenomena in SLA 456
overt vs. null subject interpretation 453–454
language switching 253
and morphology 444–446
grammatical and semantic features expressed by 382–384
inflential 382–387
in nominal domain 446
semantic reflexes of functional morphology knowledge 385–387
of the nominal domain at advanced proficiency 444–446
articles 444–446
mass/count distinction and plural interpretation 444–445
scope interpretation at advanced proficiency see scope interpretation at advanced proficiency
semantic complexity 53
semiosis 148
sentence processing studies, types 30–31
see also psycholinguistic approaches
Separation Hypothesis 84, 397n2
Sepehrinia, S. 100
Serafini, E. J. 142, 189
serial reaction time 297–298
Serratrice, L. 420
SFH (Single Feature Hypothesis) 327–328
SFL see systemic functional linguistics (SFL)
Shallow Structure Hypothesis (SSH) 30–31
Shaw, P. M. 145
Shi, D. 23
Shively, R. L. 512, 558
Shum, M. S. K. 23
Shum, S. 231
signs, internalization and mediation through 115–117
Silva-Corvalan, C. 368, 374
Similarity Differential Rate Hypothesis (SDRH) 269, 270, 272
Simões, A. 315
simple past 141
Simple Span Task 249
Single Feature Hypothesis (SFH) 327–328
situation, context of 11
Siyanova, A. 405–406
Skehan, P. 141, 220, 222, 230, 231, 233
Skinner, B. 328
SLA see second language acquisition (SLA)
Slavic languages 284, 286
SLM (Speech Learning Model) 248, 289, 329, 333
Slobin, D. 138
Smith, C. 363
smysl (personal meaning) 121
Snape, N. 86, 444, 445
social media use 175
socially interactive input 295–296
sociocultural theory (SCT) 114–130
content of L2 instruction 121–124
Dynamic Assessment (DA) 115, 125–127
further research 127–128
higher forms of consciousness, development of 117–120
internalization and mediation through signs 115–117
L2 education to promote advanced proficiency 120–128
Mediated Development (MD) 115, 121–124
principles 115–120
schooling, developing higher forms of consciousness through 117–120
Systemic-Theoretical Instruction (STI) 115, 121–124
Teaching and Learning Cycle 23, 24
sociolinguistic competence, further research 558–560
sociopragmatics 127
Solon, M. 551
Song, H. S. 449, 450
Sorace, A. 67, 419, 420, 452, 453
Sotillo, S. 104
sound–meaning correspondences (meaning-attending) 249
Spadaro, K. 61, 63, 64
Spanish language/speaking
advanced proficient heritage speakers 187
Basque verbs 390
Catalan-Spanish 181
connected speech 308
Critical Period (CP) research 59
embeddedness of grammatical features 140
exemplar sentences 121
generative approaches to second language (L2) acquisition (GenSLA) 76, 81
heritage language learners (HLLs) 182
importance of Spanish 201
individual differences in advanced proficiency 158, 161
lacking of prosodic word adjunction 86
and markedness 272
mood distinction 343–345
past tense feature 85–86
Spanish-English L2 learners 32, 43
verbs 187
Spanish Learner Language Oral Corpora (SPLLOC) 389
spatial orientations, languages following 122
Speech Act Theory 513
speech acts, production of 513–521
advanced competence
characteristics 515–520
linguistic strategies of requests 515–517
review methods 514–515
sequential organization of requests 517–520
speech attainment, advanced-level
see also connected speech; second language acquisition (SLA)
age 292, 294
cognition 296
further research 297–298
individual differences 292, 294–296
input quality and quantity 294–296
acoustically enhanced input 295
high variability input 295
socially interactive input 295–296
motivation 296
perception and production 287–288
process of 283–298
product of 288–292
segmental attainment/learning 284–285, 288–291, 293t
suprasegmental attainment/learning 286, 291–292, 293t
Speech Learning Model (SLM) 248, 289, 329, 333
speech recognition technology 319
speech styles 275
Spenader, A. J. 175
Springer, P. 432
Sprouse, R. 78, 80, 385, 447–448
SPSS 22.0 162, 164
Stæhr, L. S. 402
Stafford, C. 181, 191–192
Stampe, D. 245
standardized testing 158
Staum Casasanto, L. 549
Steinkrauss, R. 292
STI (Systemic-Theoretical Instruction) 115, 121–124
Stochastic Optimality Theory 328
stop burst 324
stop release 324
stress deafness 250, 258
cue-dependent 257
previous research 255–257
re-construction of 257
stress placement 246
stress sequence repetition 250
structural attachment preferences 31
Structural Conformity Hypothesis (SCH) 273, 275, 276, 277
subject interpretation, overt vs. null 453–454
subject matter knowledge 531
subjective-indicative alternation 343, 344
subject-verb inversion, and information structure 423–427
subordinate clauses 347
Subset Principle 81
suggestive accelerative learning and teaching (SALT) method 477
suprasegmental attainment/learning 250, 283, 286, 297
product of advanced-level L2 speech attainment 291–292, 293t
surface-scope reading 448
surveys, continuous variables 167, 168t
Swain, M. 95, 98, 101, 102, 128, 547
Swales, J. M. 143
Swan, M. 374
Swender, E. B. 182, 199
Syder, F. H. 15
syllables
duration ratio between stressed/unstressed 286
English structure 271
open/CV 305
stressing of 246, 247
structure 308
unstressed 307
voice contrasts in codas 271
tvowel-initial 305
synchronous computer-mediated communication (SCMC) 103–104, 106
synoptic semiosis 148
syntax
  see also syntax-discourse interface
  and morphology 54–57, 58t, 59–61
  morphology-before-syntax 84
  syntax-before-morphology 84, 384–385
  syntactic complexity 16, 31, 223, 224, 344
  syntactic knowledge 489
  syntactic processing, problems with 35
  syntax–semantics interface 443
syntax-discourse interface 420–430
  information structure at 419–420
  native-like attainment at 430, 431t
systemic functional linguistics (SFL)
  advanced ideational resources 16–17
  advanced interpersonal resources 18–20
  and advanced L2 proficiency 9–29
  advanced textual resources 20–22
  complementarity of context and text 11–12
  complementarity of system and instance 12
  congruent vs. incongruent meaning-making 12–15
  and construct of L2 advancedness 10–15
  dialogic balance and control 20
  discourse participants, constructing relationship between 18–20
  Engagement 19, 20, 22
  fostering advanced L2 proficiency 22–24
  and genres 146
  grammatical metaphor (GM) 13, 14, 16–20
  L2 advanced proficiency, SFL-informed research agenda 24–25
  longitudinal studies 16, 17, 20
  nominalizations 16, 17
  organizing textual meanings through thematic choices 20–22
  SFL-based assumptions about language and language development 10–15
  SFL-based descriptions 16–22
  SFL-inspired curriculum, pedagogy and assessment 22–24
  specifying L2 advancedness 16–22
  stance 20
  systemic model of language 15
  theme 20–21, 22
Systemic-Theoretical Instruction (STI) 115, 121–124
Taguchi N. 508
Tai, C. P. 23
Takahashi, S. 478
Tardy, C. M. 144, 531, 541
target language (TL) 96, 264
Tarone, E. 287
task condition effects
  see also task repetition
  on advanced-level foreign language performance 219–237
  further research 232–234
  and high-proficiency learners 226–227
  on-line planning with advanced learners 224–227
  native speaker performance 230–232
  pre-task planning with high-proficiency learners
    behavior of advanced learners 221–222
    effects on CALF 222–224
task repetition
  see also task condition effects
  defining 227
  effects, with advanced learners 220, 227–230
  following a time interval 228–229
  immediate 228
  justifications for 227–228
  multiple repetitions with time intervals 229, 230
  operationalizations of 228
  speech comprehension 227
  speech monitoring system 227
Task-Based Language Teaching (TBLT) 95, 100, 104, 107, 205, 210, 221
tasks
  characteristics 220
  independent performance of 119–120
  task condition effects see task condition effects
  task repetition see task repetition
  Task-Based Language Teaching (TBLT) see Task-Based Language Teaching (TBLT)
task-essentialness 210
Tavakoli, M. 225, 228, 229, 231, 232
Taylor, B. 256, 257
TBLT see Task-Based Language Teaching (TBLT)
Teaching and Learning Cycle, sociocultural theory 23, 24
technology
  control for judicious uses 210–211
  interaction and advanced learners 103–104
Temple, L. 344
Teng, F. 222
tenor, context of 11, 13
Test of English as a Foreign Language (TOEFL) 97, 137, 311, 465
Test of Spoken English (TSE) 97
texts
  advanced resources 20–22
  argumentative 20
  and context 11–12
  organizing meanings in, through thematic choices 20–22
textual configuration 17
textual metafunctions 10
textual modeling 23
textual scaffolding 23
theme
  Chinese language 124
  systemic functional linguistics (SFL) 20–21, 22
theoretical knowledge 119
Thomas, J. 191
Thomas, M. 182, 183
Thomson, R. I. 248, 252
Thorndike, E. L. 118
Thorne, S. L. 175
360-degree feedback 202
time, spatial metaphors for 122
TOEFL (Test of English as a Foreign Language) 97, 137, 311, 465
Tokowicz, N. 45n2
Tomlin, R. S. 207
Toomela, A. 116–117
topicalization and clitics 455
Torres, J. 183, 190
Toth, P. 376n8
Towell, R. 252, 286
Townsend, D. 489
Toyoda, E. 103
translation equivalents 254
Tremblay, A. 43, 258
Trenkic, D. 445
Trofimovich, P. 286, 291, 314
Trubetzkoy, N. S. 329
TSE (Test of Spoken English) 97
Tsimpli, I. M. 80, 387–388, 452
TwoStep clustering algorithm 164, 165, 173
typological markedness 264, 266–269
  concept 268
  as an explanatory principle 275–278
generalizations 276
Underwood, M. 312
Underwood, P. R. 314
United States
  see also Georgetown University, Washington DC
Census Bureau 201
Department of Labor 201
Department of State 278–279
Foreign Language (FL) teachers 201
qualified FL users 202
Universal Grammar (UG) 75, 87, 349, 369
  access to 78, 79–80, 443
  beyond UG access 82–86
Valdés, G. 159
Valls-Ferrer, M. 286
van Compernolle, R. A. 127
van den Branden, K. 142
van der Slik, F. 247
van der Zwaard, R. 103
van Hell, J. G. 33–34, 45n2
van Mulken, M. 316
van Vuuren, S. 432
vanPatten, B. 200, 201, 203, 206–208
variable structures
  abilities of advanced learners 550–555
  abilities required by a sociolinguistically competent speaker 548–549
  acquisition of sociolinguistic competence, additional factors 556–558
  future research on advanced proficiency/sociolinguistic competence 558–560
  and sociolinguistic variation 547–565
VCV (intervocalic consonant) 308
Vendelin, I. 256
Vendler, Z. 363
verb end 243
verb final languages 34
verb separation 243
verbal inflectional morphology 388–390
verb-second (V2) syntax 432
Verhoeven, L. 146
Verkuyl, H. 362
video watching 169–170, 174, 175
Villa, V. 207
vocabulary
  Critical Period (CP) research 61
  knowledge 488–489
  learning program, four strands of 407
  lexical development, advanced 401–403
  size 61, 401–403
Index

Vocabulary Size Test 403
voice contrasts in onsets 267
voice onset time (VOT)
acquisition of in word-initial stops 291
in advanced SLA 323–339
aspirating languages 326
attributes 324–327
ceiling effects 330
common sense 323
concept 324
contrast lead 326
further research 335–336
gap between expected and actual values 332–334
gestures 325
how and when of acquisition 328–329
laryngeal contrasts 323
long-lag 324, 326
negative 326
positive 326
research contributions 330–331
segmental learning 285
short-lag 326
target of acquisition 327–328
two-way contrast between voiced and voiceless obstruents 325
utterance-initial 326
voicing 325
voicing languages 326
vowel acquisition 290
vowel identification 285
Vygotsky, L. S. 114–121, 127, 128

Wagner, D. E. 376n8
Waldman, T. 61
Wallace, W. 314
Walley, A. 285
Wang, Z. 225, 226, 228
Watanabe, Y. 102
Webb, S. 403, 413
Webb, S. A. 406
Webb, W. B. 409
Wexler, K. 445

Weyers, J. R. 203

wh-expressions
dependencies 37
questions 33, 34
scope of 448–450
unambiguous extractions 35
White, B. 256, 257
White, L. 85, 86, 276, 385, 391, 396, 455
Wigglesworth, G. 223
Wijnen, F. 45n3
Wildner-Basset, M. 477
within-task planning 224
Wong, S. W. L. 315
Wong, W. 208
word boundaries 309
word families 284, 285
word order, Critical Period (CP)
research 59–60
word-boundary palatalization 315
word-final position, voice contrasts 274
word-final voiceless consonants 308
word-initial stops 291
word-level prominence 291
working memory 43, 189, 206, 249
writing 528–529
written contextualized task (WCT) 351
Wu, S. M. 20
Wu, X. 287

Yanez-Prieto, C. M. 370, 371
Yang, C.-L. 249
Yashima, T. 298
Yasuda, S. 17, 23
Yavas, M. 271
Yoon, J. 455
Yuan, F. 223, 224, 225
Yule, G. 102, 105

Zhang, Y. 474
Ziegler, N. 104

znachenie (conventionalized meaning) 121
Zone of Proximal Development (ZPD) 115, 119, 120
Zyzik, E. 140, 191, 202
WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.