

Second Language Acquisition Theory

The legacy of
Professor Michael H. Long

EDITED BY

**Alessandro G. Benati
and John W. Schwieter**

**Bilingual Processing
and Acquisition**

14

John Benjamins Publishing Company

Second Language Acquisition Theory

Bilingual Processing and Acquisition (BPA)

ISSN 2352-0531

Psycholinguistic and neurocognitive approaches to bilingualism/multilingualism and language acquisition continue to gain momentum and uncover valuable findings explaining how multiple languages are represented in and processed by the human mind. With these intensified scholarly efforts come thought-provoking inquiries, pioneering findings, and new research directions. The *Bilingual Processing and Acquisition* book series seeks to provide a unified home, unlike any other, for this enterprise by providing a single forum and home for the highest-quality monographs and collective volumes related to language processing issues among multilinguals and learners of non-native languages. These volumes are authoritative works in their areas and should not only interest researchers and scholars investigating psycholinguistic and neurocognitive approaches to bilingualism/multilingualism and language acquisition but also appeal to professional practitioners and advanced undergraduate and graduate students.

For an overview of all books published in this series, please see benjamins.com/catalog/bpa

Executive Editor

John W. Schwieter
Wilfrid Laurier University

Associate Editor

Aline Ferreira
University of California, Santa Barbara

Editorial Advisory Board

Jeanette Altarriba
*University at Albany, State
University of New York*

Panos Athanasopoulos
Lancaster University

Laura Bosch
Universitat de Barcelona

Marc Brysbaert
Ghent University

Kees de Bot
University of Groningen

Yanping Dong
Zhejiang University

Mira Goral
*Lehman College, The City
University of New York*

Roberto R. Heredia
Texas A&M International University

Arturo E. Hernandez
University of Houston

Ludmila Isurin
Ohio State University

Janet G. van Hell
Pennsylvania State University

Walter J.B. van Heuven
University of Nottingham

Iring Koch
RWTH Aachen University

Li Wei
UCL IOE

Gerrit Jan Kootstra
*Radboud University Nijmegen &
Windesheim University of Applied
Sciences*

Gary Libben
Brock University

Silvina Montrul
*University of Illinois at Urbana-
Champaign*

Kara Morgan-Short
University of Illinois at Chicago

Greg Poarch
University of Groningen

Leah Roberts
University of York

Norman Segalowitz
Concordia University

Antonella Sorace
University of Edinburgh

Volume 14

Second Language Acquisition Theory. The legacy of professor Michael H. Long
Edited by Alessandro G. Benati and John W. Schwieter

Second Language Acquisition Theory

The legacy of professor Michael H. Long

Edited by

Alessandro G. Benati

University of Hong Kong

John W. Schwieter

Wilfrid Laurer University

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/bpa.14

**Cataloging-in-Publication Data available from Library of Congress:
LCCN 2022025788 (PRINT) / 2022025789 (E-BOOK)**

ISBN 978 90 272 1256 6 (HB)

ISBN 978 90 272 5748 2 (E-BOOK)

© 2022 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · <https://benjamins.com>

Table of contents

In memory of Professor Michael H. Long (1945–2021)	VII
Acknowledgements	IX
Preface	XI
About the editor and contributors	XIX
CHAPTER 1	
The legacy of Professor Michael H. Long and his influence in second language acquisition	1
<i>Alessandro G. Benati and John W. Schwieter</i>	
CHAPTER 2	
Taking the Long view: A bibliometric analysis	9
<i>Susan Gass, Luke Plonsky and Lizz Huntley</i>	
CHAPTER 3	
Stabilization: A dynamic account	29
<i>Manfred Pienemann, Frank Lanze, Howard Nicholas and Anke Lenzing</i>	
CHAPTER 4	
Enhanced incidental learning	77
<i>Ilaria Borro</i>	
CHAPTER 5	
The uncertainty principle in second language acquisition	101
<i>Stefano Rastelli</i>	
CHAPTER 6	
Tandem interaction enhancement: Manipulating NS-NNS semi-spontaneous conversation to promote focus on form	113
<i>Elena Nuzzo</i>	
CHAPTER 7	
Immediate versus delayed oral negative feedback: A comparison of psycholinguistic advantages	127
<i>Gisela Granena and Yucel Yilmaz</i>	

CHAPTER 8	
A calculus for L1 transfer	143
<i>William O'Grady</i>	
CHAPTER 9	
Age of acquisition in second language thinking	177
<i>Emanuel Bylund and Panos Athanasopoulos</i>	
CHAPTER 10	
Age effects in naturalistic and instructed second language acquisition: Two sides of one coin	197
<i>Robert M. DeKeyser</i>	
CHAPTER 11	
Performance on second language speaking tasks: Supports and impediments	211
<i>Peter Skehan</i>	
CHAPTER 12	
A task-based needs analysis framework for TBLT: Theory, purpose, and application	235
<i>Ellen Grote and Rhonda Oliver</i>	
CHAPTER 13	
Can focus on form have an effect on language development?	257
<i>Alessandro G. Benati</i>	
Index	273

In memory of Professor Michael H. Long (1945–2021)

Professor Michael H. Long (1945–2021) played a huge role in the development of second language acquisition as a field of enquiry. He was a true academic, relentless in his publications right up until his untimely death. He was a believer in the power of dissemination to further the debate in our field. He was truly committed to the promotion of research in second language acquisition.

We are all indebted to Mike. On a personal level we remember his generosity of time for others, his dedication to his students and his friendship. It is terribly sad to lose him so early, but its legacy will be with us for the years to come. We will all miss him greatly.

Acknowledgements

We would like to thank the contributors for their excellent contributions highlighting some of the major areas of research in second language acquisition which were significantly shaped by the work of Professor Michael H. Long. We also are very grateful to Catherine J. S. Doughty for writing the preface.

We are very appreciative to John Benjamins Publishing for their professional partnership in putting together this volume in honour of Long's legacy. We thank the peer reviewers of the chapters for their time and expertise and for helping to improve the overall quality of the volume.

It was a pleasure to bring together the contributing scholars and everyone involved in the review and production process to collectively pay homage to Mike's legacy in the field of second language acquisition.

All royalties generated from the sales of this volume will go towards a selected charity deeply admired by Mike.

Preface

As so many good friends and colleagues throughout the world have noted, Mike was a giant in our field and inspirational to us all. Mike's own life adventures and experiences inspired and shaped his thinking about how adults learn second languages. After completing his undergraduate law degree, but by then a firm anarchist uninterested in practicing law, Mike decided to join Voluntary Service Overseas to "do something more meaningful." His assignment was changed at the last minute from Afghanistan to Peru, and Mike often wondered how differently his life might have turned out had he gone to Kabul. In any case, his VSO experience in Peru did shape him significantly. Shortly after arriving, there was an earthquake, and the volunteers were given leave to travel around a bit before their training could get organized and begin. That was the start of Mike's own Spanish "learning by doing" over years on and off in South and Central America. In addition, as Mike told it, during his volunteer service, he discovered that he did not have skills that were of much use to the VSO. Other volunteers had technical and engineering skills, which he greatly admired, but recognized his own lack of proclivity or one might even say resistance (always claiming that he had "a hard time figuring out where to put the quarter in the computer.") The VSO program manager gently suggested that Mike could perhaps get a qualification to teach English and come back. This is where Mike began to place an emphasis on expertise (see the amusing comments that Stefano Rastelli quotes in his chapter in this volume from Mike on his admiration of engineering skills).

Taking up this suggestion, Mike returned to England, earned a post-graduate certificate in English language teaching and did his teaching practice in Barcelona (which he called "Mecca"). This was the start of his lifelong love of Catalonia and his many relationships with "the Catalans," as colleagues and students (and the continuation of his scholarship on the Spanish Civil War). He then earned an MA in Applied Linguistics, but rather than return to Peru, he went to teach in Mexico City. It was there, in his new department, that he began his never-ending fight against the "I'm a book, you're a book" traditional approach to language teaching and to train teachers in a new way. It's a testament to Mike's influence (and was a wonderful surprise) that several of his colleagues from that department virtually attended his life celebration on May 1, 2021. They reminisced that Mike had changed their professional lives, marking perhaps the first among countless many such personal

influences over many decades on colleagues and students throughout the world. As for influences on him, Mike continued his Spanish acquisition to very high level. Decades later, when he taught in a summer institute in San Sebastian, Mike was greeted afterward by a student saying she hadn't heard such amusing Mexican Spanish in very many years. Also significantly, Mike roomed with a scientist in Mexico City who studied cycads, exposing him to the scientific method, and who taught him how to listen to modern jazz. Ever since, these were both important systems in Mike's daily work and life.

From South and Central America, Mike headed to North America, by then estranged from England. After a brief time at Concordia University in Montreal, where he made other life-long friends among colleagues, spurred by the need for expertise and a commitment to the scientific method, Mike felt the urge for more training and was admitted to the doctoral program at UCLA. This brings us to the time when we might say that the nascent cognitive science field of Second Language Acquisition began, with UCLA having the first such research-based program and Mike at the forefront. He was among the PhD students who conceptualized the Second Language Acquisition Forum, first held only at UCLA, then alternating between UCLA and USC, and then moving around to various universities with programs focusing on second language acquisition (SLA). SLRF remains a student-run conference that, by design, is free of dominance by commercial publishers, who were (and still are), according to Mike, responsible for many of the problems and lack of progress in language teaching. Mike argued strongly against forming a SLRF organization to keep costs down and focus the emphasis on SLA theory and research. Mike supported his own masters and doctoral students as they hosted SLRF at the Universities of Hawaii and Maryland. I truly hope that this conference will continue as originally conceptualized as a venue where graduate students and more established scholars share their latest SLA research findings.

Throughout his career as a professor at the Universities of Pennsylvania, Hawaii, and Maryland, Mike started new masters and doctoral programs in SLA from the ground up. Working with graduate students was extremely satisfying to him, and he insisted on their rigorous training leading to expertise. He also held himself to a very high standard of scholarship. Mike kept up with the scientific literature on SLA in general, as well as on the more theoretical topic of age effects and the critical period for SLA. By "keeping up," I mean that he read *everything* that was published, as well as yet-to-be published graduate student work, manuscripts he reviewed for journals (including *Studies in Second Language Acquisition* as a continuous editorial board member), and papers sent to him by researchers from around the world. Mike astounded everyone by remembering in detail, including names, dates, and places of publication, much of the past empirical work in areas of interest to him. His students (and colleagues!) benefitted tremendously from his wealth of knowledge and ability

to integrate new research findings into the exponentially expanding body of evidence on SLA. In addition to the students at his own universities, he was generous with his time with international students. Many of his students have become prominent in the field of SLA in their own right.

In his own research program, Mike was keenly interested in formulating evidence-based SLA principles and applying them in language teaching, which led him to invent Task-Based Language Teaching. TBLT has 10 methodological principles, all of which were derived from the empirical findings of instructed SLA research. In talks to teachers over the years, most recently to my colleagues at the Foreign Service Institute, which he saw as an ideal context for TBLT, Mike would point out that the popular language teaching approaches appearing in textbooks have little or no empirical evidence underpinning them. Not so with TBLT, which has attracted tremendous research interest (due to its practicality, as noted by Peter Skehan in this volume) and amassed considerable empirical support. And, as Peter Skehan put so well in his chapter: “No one else has argued more influentially that claims should not be data free and should not be based simply on calls to personal experience.” It was Mike’s academic life’s work over several decades to integrate the findings of ISLA into his formulation of TBLT as they became available (codified in Long, 2015).

One aspect of TBLT was still puzzling Mike, given the still insufficient evidence from studies. That is the question of how to sequence pedagogic tasks to efficiently and effectively lead learners to the ability to do target tasks in and through the second language. His money was still on Peter Robinson’s notion of sequencing according to the complexity of tasks rather than the complexity of language with a parallel notion stating that as task complexity increases, so does the language needed to accomplish them (Robinson, 2011). But Mike admitted that we need accumulate to more evidence. Hopefully ISLA researchers will continue this line of research. Meanwhile, Mike was confident enough in the complexity approach to suggest it as the best available guideline. I believe Mike would have responded very positively to the contribution from Peter Skehan in this volume to the discussion of how task conditions can be set up to improve speaking ability. Teachers interested in implementing TBLT in their classrooms constantly ask how to design tasks. In Mike’s own writing, he leaves the pedagogy up to the teachers themselves, as they are the experts in their own classrooms and are aware of the learning conditions each day. Peter provides a number of specific examples of scaffolding, which comes naturally to teachers. The importance of that chapter is that it explains to teachers exactly how, and perhaps more importantly, why and when to scaffold (e.g., task planning and task repetition) and what the outcomes are expected to be. It also touches upon another of Peter Robinson’s ideas that was appealing to Mike: task conditions engage learning processes in different ways, sometimes pushing learners

to notice and eventually acquire new language features and at other times allowing them to practice the newly acquired language (Robinson, 2011). The proposal that Alessandro G. Benati is pursuing in his structured input research described in this volume, with its four conditions of providing sufficient input first, checking the developmental schedule, taking note of how learners process input, and manipulating the processing with structured input, may offer another approach to task sequencing, but this remains to be tested.

Another chapter that teachers will find extremely useful and practical is the contribution from Ellen Grote and Mike's long-time Australian friend and colleague, Rhonda Oliver. While the reprise of Task-based Needs Analysis is very useful and practical-steps oriented, what teachers are seeking are examples of how to develop pedagogic tasks. This chapter delivers many rich examples, particularly enjoyable to read among which are those involving speakers of Aboriginal English learning Standard Australian English in Structured Workplace programs. The detail offered in these examples is highly informative in showing materials developers how to meet the needs identified in the TBLA.

That second language acquisition entails a set of processes that take place during interaction was fundamental to Mike from the very beginning of his scholarly career. As reflected in the contributions to this volume, over time, Mike systematically investigated input, interaction, interlanguage development, negotiation of meaning, effective instructional interventions, so-called fossilization, and age effects on language learning. He used the findings to advocate for improving the efficiency of instruction in general, and of course, for TBLT. For this to work in practice, Mike argued that language teachers need to be trained in and should draw upon SLA principles, enhancing their unique ability as the local, on-the-ground experts in how to apply them in their own contexts with their own students and under the moment-to-moment learning conditions that arise. Those who know Mike know that he had little tolerance for the "old guard" in language teaching methodology, whom he called arm-chair applied linguists, nor for solely using intuitions in language teaching. He championed theory and evidence in language teaching, a field that has come to be known as instructed SLA (ISLA). And he continued to eschew the commercialization of "one-size-fits all" language teaching, which is based on no theory or research at all.

Mike used his gift at interpreting cumulative research findings to help teachers (of any language really, not just English) to gain some knowledge that is helpful in making the day-to-day pedagogical decisions that he felt strongly are solely in their domain. For example, he always took pains to describe Manfred Pienemann's interlanguage development and Processability Theory work, which has been formulated incrementally and withstood the test of empirical research over time. Mike would have been pleased to read the contribution by Manfred and his colleagues in this

book which succinctly describes the research advances in our understanding of the development schedule and takes on the “loose end” of the notion of fossilization within the wider theory of dynamic interlanguage development. Mike would have been impressed by the emphasis in this chapter on making and testing predictions. For this same reason, he would have been highly interested in William O’Grady’s processing explanation for transfer, which shares some features with Processability Theory. William claims that transfer takes place to minimize processing cost by carrying familiar operations in the L1 over to the L2, where feasible, and he formulates and tests hypotheses that follow on from this claim. Another example of the similarity with Processability Theory is William’s emergentist perspective, much valued by Mike (also espoused by Nick Ellis *inter alia*). While there is a strong theoretical foundation for processing costs driving developmental schedules overall and transfer in particular, in practice, it is still somewhat difficult for teachers to implement this knowledge. The kind of careful, hypothesis-driven work described in the chapters of this volume will hopefully someday lead to practical guidelines on what to attend to and not to attend to with respect to those clues to those crucial cues to interlanguage development, i.e., the systematic errors that learners make.

The idea that most excited Mike recently in his efforts to make language learning more effective and efficient through TBLT is his notion of enhanced incidental learning (EIL). This is because EIL is an unobtrusive manipulation of learning conditions that could solve that problem that implicit and incidental learning – ideal for TBLT’s learning by doing – take a long time, time that learners in classrooms often do not have. Stimulated by the findings of Ilaria Borro’s dissertation (he served on her committee), Mike said to me that he really hopes that researchers will take this idea and run with it. It is interesting and illustrative of Mike’s length and breadth of influence to note that Peter Robinson was one of Mike’s very first doctoral students, and Ilaria was his last.

A parallel, equally important research interest of Mike’s was in the critical period for second language acquisition, which he shared with his University of Maryland, Robert DeKeyser. I think it is fair to say that the two of them were among the leading experts in critical period research in SLA and certainly the most informed on the intricacies of the research methodology needed to answer the complex questions of the outcomes of learning processes under different conditions at different ages and with high or low aptitude for implicit or explicit learning. Robert, like Mike, devoted much of his research energy to questions in ISLA. Robert’s contribution to this volume, lays out clearly how different learning processes are differentially effective at various ages and in particular contexts. This was to address a seeming discrepancy that older children are better than younger children in some studies. Robert carefully shows, and I know that Mike certainly agreed, that “it is the interaction of age and context (amount of input, quality of input, amount of

interaction, extent of focus on form, and student aptitude) that determines the degree of success.”

In carrying out his work, Mike was driven by logical thinking gained during his law training and strong belief in empirical research and parsimonious theory building based on research findings. I, among others, was always astounded that Mike could keep up on all the research published in both the ISLA and critical periods research domains. For him, this was essential. Highly admiring of the work of Manny Bylund (and his colleagues Kenneth Hyltenstam and Niclas Abrahamsson), I think Mike would certainly have delighted in engaging with Bylund and Athanasopoulos in discussing their investigations of critical period effects on L2 thinking. I’m not sure how he would have reacted to raising this “bête noire” again, but I do know that he would have appreciated their approach in which they explicitly honor Mike’s “legacy of critical and progressive discussions around age effects” in various language domains.

Because he was a prominent theory builder in our field, Mike was involved in a number of infamous debates during his career. Most amusingly, one of them was with his good colleague, Robert DeKeyser. While as I have mentioned, the two of them were in near perfect agreement on critical period issues, they continued to delight their students with a lively debate over the roles and prevalence of implicit and explicit learning processes in SLA. Another prominent debate was the epistemological one of rationalism vs. post-modernist socio-cognitive approaches, the latter of which caused him considerable consternation. Mike was firmly in the rationalist camp because, as pointed out by Stefano Rastelli in this volume, as a keen student of the philosophy of science, Mike followed the Popperian approach of testing falsifiable hypotheses and lived by Occam’s razor invoking parsimony. I’m certain that Mike would have taken up Stefano’s invitation to discuss the uncertainty principle, laid out in his contribution to this volume. I clearly recall Mike handing me a copy of Stefano’s book (Rastelli, 2014), *Discontinuity in second language acquisition* (which he reviewed), saying, “this guy is a genius.”

Mike had big ideas, but he required and respected the empirical work that is needed to permit them entry into a theory. Negotiation for meaning was one of his early big ideas within which feedback on error has received considerable attention from ISLA researchers. It was a delight to read the personal communication Gisela Granena and Yucel Yilmaz had with Mike on the notion of delayed feedback, which at face value should not be effective but, which could be implemented in such a way that it might be. Gisela found her way to the University of Maryland via a Fulbright scholarship. Mike admired her intellect tremendously and was amazed by her work ethic and productivity, enhanced by her partner in work and life, Yucel. Their exchange with Mike shows clearly how he approached hypothesis testing (also that he

was not stubborn and dogmatic although at time he may have seemed to be!). I'm thinking now though that he might have pointed them in the direction of William O'Grady and Manfred Pienemann's work on cost effects of long-distance processing. Mike was a firm believer of the "psycholinguistically opportune moment" that he wrote about so long ago (Long, 1991). Along these lines, Mike would certainly have agreed with Elena Nuzzo that her promising approach to interactional enhancement during tandem learning described in this volume, spot on with respect to learning conditions, should be formulated systematically and is in need of more empirical investigation.

Mike was working and writing up to the last month. As I mentioned at the outset, Mike was a committed anarchist, and in the latter part of his career, he incorporated more and more of these principles into his ideas about language teaching (see, for instance Chapter 4 of Long 2015). In his last book, Mike collaborated with his life-long friend and colleague, Geoff Jordan (an honorary Catalan), on *How English language teaching is and how it could be*. In addition to putting their political views front and center, particularly with respect to their lively critique of publishers of textbooks and tests, an essential aim of the book is to make SLA theory and research accessible to teachers so that they can make evidence-based and principled decisions. In their own words (Jordan & Long, 2022, p. 3) describe who the book is for and what it is about:

We make many teaching recommendations, but this is not a book of classroom recipes. Rather, we aim to lay out what research has shown about how people learn English as a foreign or second language, and for that matter, any other foreign or second language. Understanding how languages are learned and how and when to intervene are fundamental to the knowledge base of qualified language teachers, just as understanding human anatomy and physiology, the respiratory and cardiovascular systems, and the treatment of diseases is critically important for physicians, in fact required of them. There is still nothing for language teachers like the body of knowledge available to medical practitioners, and much is still *not* known about language learning, but many of the research findings over the past 50 years are important and very useful in the classroom. It is incumbent upon all of us to keep abreast of what is known about how language learning can be done most efficiently and how best to teach and test, but it is equally important to be aware of what is not yet known.

I highly recommend this book to readers for Mike's cumulative and recent thinking on second language acquisition. At a very difficult time, he was deeply satisfied by this collaborative work with Geoff.

References

- Jordan, G. and Long, M. H. (2022). *English language teaching now and how it could be*. Cambridge: Cambridge Scholars Publishing.
- Long, M. H. (1991). The design and psycholinguistic motivation of research on foreign language learning. In Freed, B. (ed.), *Foreign language acquisition research and the classroom* (pp. 309–20). Lexington, Mass.: D. C. Heath.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Rastelli, S. (2014). *Discontinuity in second language acquisition: The switch between statistical and grammatical learning*. Multilingual Matters.
- Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance*. John Benjamins.

About the editor and contributors

About the editors

Alessandro G. Benati is Director of CAES at The University of Hong Kong. He is known for his research on how L2 learners process information and what internal strategies they use in language processing across romance and non-romance languages. Most recently his ground-breaking research on the effects of structured input has been driven by the use of new online measurements (e.g., eye tracking, self-paced reading, and reaction times).

John W. Schwieter is Professor of Spanish and Linguistics and Cross-Appointed in Psychology at Wilfrid Laurier University in Waterloo in Canada, where he is also Director of the Language Acquisition, Multilingualism, and Cognition Lab and Bilingualism Matters @ Laurier. His research spanning bilingualism and language acquisition has appeared in journals such as *Acta Psychologica*; *Applied Linguistics Review*; *Bilingualism: Language and Cognition*; *Cerebral Cortex*; *Critical Inquiry in Language Studies*; *Frontiers in Psychology*; *Intercultural Education*; *International Journal of Bilingualism*; *International Journal of Bilingual Education and Bilingualism*; *Language Learning*; *Linguistic Approaches to Bilingualism*; *Psychophysiology*; and *Neuropsychologia*.

About the contributors

CHAPTER 1

Alessandro G. Benati (see “about the editors”).

John W. Schwieter (see “about the editors”).

CHAPTER 2

Susan Gass is University Distinguished Professor Emerita in the Second Language Studies Program at Michigan State University. She has published widely in the field of second language acquisition. She has served as President of the American Association for Applied Linguistics (AAAL) and of the International Association of Applied Linguistics (AILA). She is the winner of numerous local, national, and international awards and is the current co-author of two recent books: *Second language acquisition: An introductory course* (5th ed.), with Jennifer Behney and Luke

Plonsky and *Second language research: Methodology and design* (3rd ed.) with Alison Mackey. She is the former editor of *Studies in Second Language Acquisition*.

Luke Plonsky is an Associate Professor of Applied Linguistics at Northern Arizona University in the United States where he teaches courses in SLA and research methods. His work in these and other areas has resulted in over 100 articles, book chapters, and books. Luke is Senior Associate Editor of *Studies in Second Language Acquisition*, Managing Editor of *Foreign Language Annals*, Co-Editor of De Gruyter Mouton's Series on Language Acquisition, and Co-Director of the IRIS Database (iris-database.org). In the midst of the Covid pandemic Luke published the highly acclaimed *Professional development in applied linguistics: A guide to success for graduate students and early career faculty* with John Benjamins. In addition to prior appointments at Georgetown and University College London, Luke has lectured in China, Japan, The Netherlands, Poland, Spain, and Puerto Rico.

Lizz Huntley is a Doctoral Candidate in the Second Language Studies program at Michigan State University, United States. Her research interests include bibliometrics, open science practices, multimodal input, and psycholinguistic approaches to second language variation. She is the co-author of two recent textbooks for second language learners of Arabic: '*Arabiyyaat al-Naas fii MaSr*, and '*Arabiyyaat al-Naas Part II* (with Munther Younes, Makda Weatherspoon, Jonathan Featherstone, and Hanada Al-Masri). Additionally, she has served as editorial assistant for the journal *Studies in Second Language Acquisition* for four years.

CHAPTER 3

Manfred Pienemann is Professor Emeritus of Linguistics at Paderborn University, Germany and co-chair of the linguistic engineering company. He was previously Professor of Applied Linguistics at the University of Newcastle (UK), and also at the Australian National University. He founded the Language Acquisition Research Centre at the University of Sydney and was one of the founding members of PacSLRF. He has been involved in second language acquisition research since the 1970s, when he worked on variational aspects of the ZISA study directed by Jürgen Meisel. His main contribution to the field of SLA has been the development of processability theory (PT), a cognitive approach to SLA that has been tested for a range of L2s and applied to the profiling of L2 development. His current research focuses on automating L2 profiling and on integrating formal aspects of dynamic systems theory into PT.

Frank Lanze is co-chair of the linguistic engineering company. He is a computer scientist with 12 years of industry experience as a software developer and software architect. His R&D activities focus on artificial intelligence, evolutionary algorithms

and computational linguistics. His current research deals with automating a L2 profiling system based on processability theory.

Howard Nicholas is an Adjunct Professor in the School of Education at La Trobe University in Australia. His research interests encompass delineating communicative repertoires and their development (Nicholas & Starks, 2014). In 1981, he worked on variational aspects of the longitudinal extension of the ZISA study directed by Jürgen Meisel that investigated adult second language acquisition of German. This complemented his research interests in child second language acquisition as part of a life-long endeavour to establish that second language acquisition encompasses more than adult language acquisition. In his work with children and in his work on additional language and literacy acquisition/development by refugee adults he has continuously engaged with the widest possible spectrum of educational and life experiences among second language learners, as for example in *The kaleidoscope of adult second language learning: Learner, teacher and researcher perspectives* (edited by G. Wigglesworth, 2003).

Anke Lenzing is Professor of English Language Education at Innsbruck University in Austria. Her research interests lie in psycholinguistic aspects of (instructed) second language acquisition. She has published on topics such as language processing in beginning second language learners, the role of formulaic sequences in instructed SLA, and relationships between L2 production and comprehension. She is the author of the books *The development of the grammatical system in early second language acquisition: The Multiple Constraints Hypothesis* (2013, John Benjamins) and *The production-comprehension interface in second language acquisition: An integrated encoding-decoding model* (2021, Bloomsbury). Her current research engages with psycholinguistic perspectives on L2 communicative interaction.

CHAPTER 4

Ilaria Borro recently completed her PhD at the University of Portsmouth in England and a thesis related to the incidental learning of second languages, which was supervised by Mike Long. For over 10 years, she has taught Italian as a second language in public and private academic institutions. At the same time, she has contributed to the scientific research of second language acquisition through experimental research, presentations at international conferences, and publications. She is an active member of GRAAL (Gruppo di Ricerca e Azione sull'Apprendimento delle Lingue), where she contributes to the organization of professional development events for language teachers and to the support of research activities whose objective is to disseminate language teaching pedagogy with well-founded scientific and empirical evidence.

CHAPTER 5

Stefano Rastelli is Associate Professor of Psycholinguistics and Second Language Acquisition Linguistics at the University of Pavia in Italy, where he directs the Laboratorio di Linguistica e Glottodidattica Sperimentale (LLEGS). He utilizes EEG, eye-tracking and behavioral techniques to investigate syntactic representations, the relationship between statistical and grammatical learning and the comprehension of formal texts (e.g., laws, contracts and regulations) by low-literacy readers, elders, and immigrants.

CHAPTER 6

Elena Nuzzo is an Associate Professor of Applied Linguistics at the Department of Foreign Languages Literatures and Cultures, Università Roma Tre in Italy. She teaches and researches in the fields of second language acquisition and teaching, with a focus on second language Italian. Her main research interests include applications of speech act theory to second language learning and cross-cultural communication, and task-based language teaching. Her research interests have recently expanded to include the role of external manipulation in low-structured learning environments, such as tandem and telecollaboration.

CHAPTER 7

Gisela Granena is an Associate Professor of English as a Second Language and Director of the technology-mediated second language teaching and learning program at the Universitat Oberta de Catalunya in Spain. Her research interests include the role of cognitive aptitudes in naturalistic and instructed contexts; aptitude-treatment interactions; task-based language teaching (TBLT); measures of implicit and explicit language knowledge; and the effects of early and late bilingualism on long-term second language achievement.

Yucel Yilmaz is an Associate Professor of Second Language Studies at Indiana University in the United States. He received an M.A. in TESOL from the University of Pennsylvania and a PhD in Multilingual/Multicultural Education from the Florida State University. His research focuses on second language interaction and corrective feedback, computer-mediated communication, task-based language teaching, individual differences in second language acquisition, and explicit and implicit learning processes

CHAPTER 8

William O'Grady is a Professor in the Department of Linguistics at the University of Hawai'i at Mānoa, United States. He is well known for his writings on syntactic theory and language acquisition, including second language learning. A

major theme in O'Grady's research is his commitment to emergentism, the idea that language is a complex system whose properties derive from the interaction of more basic factors and forces, especially processing pressures. He has explored the relevance of this idea to SLA in several publications and presentations, often with a focus on the phenomenon of transfer. He is the author of several books and numerous journal articles. His books include *How children learn language*, *Syntactic carpentry*, and *Jejueo: The language of Korea's Jeju Island*. He is the co-editor (with Brian MacWhinney) of *The handbook of language emergence* and the co-editor (with John Archibald) of a widely used textbook, *Contemporary linguistics analysis*, now in its ninth edition.

CHAPTER 9

Emanuel Bylund is professor of linguistics (Stellenbosch University in South Africa, and Stockholm University in Sweden) with a background in bilingualism research and Spanish linguistics. His current research concerns the role of age in language acquisition and language loss, and conceptual representation in multilingualism, and has appeared in journals such as *Applied Linguistics*, *Cognition*, *Journal of Experimental Psychology: General*, and *Studies in Second Language Acquisition*.

Panos Athanopoulos is a Professor in the Department of Linguistics and English Language at Lancaster University. His multi-disciplinary research program addresses two fundamental questions in the Humanities and Social Sciences. The first concerns the extent to which language representations interact with cognitive processing of the perceived world. The second probes the dynamic nature of these interactions, asking whether they can be restructured by learning a new language with contrasting representations from the first. His work on the representation of categories such as number, colour, gender, aspect, motion, space and time, has appeared in some of the most impactful journals in the fields of Linguistics and Psychology such as *Applied Linguistics*, *Cognition*, *Language Learning*, and *Psychological Science*, as well as in documentaries for the BBC, in *Science Magazine*, and in the TEDx Talks series.

CHAPTER 10

Robert M. DeKeyser is a Professor emeritus of Second Language Acquisition at the University of Maryland, College Park, United States. His research interests include language learning during a stay abroad, age effects on language learning, other individual differences in language learning, and the interaction between age, aptitude, and teaching contexts/methodologies. He has published in a wide variety of journals. In 2007 he published an edited book on practice activities in second language learning, in 2019 a co-edited book with Goretta Prieto Botana about research issues

in instructed language learning, and in 2021 an edited book on aptitude-treatment interaction. He was editor of *Language Learning* from 2005 to 2010 and associate editor of *Bilingualism: Language and Cognition* from 2010 to 2014.

CHAPTER 11

Peter Skehan is an Honorary Research Fellow at the Institute of Education, University College London, U.K. Previously he taught at King's College, London; the Chinese University of Hong Kong, and the University of Auckland. His interests are in second language acquisition, task-based learning and teaching, second language spoken performance, language testing and foreign language aptitude. Recently he has published in the *Modern Language Journal*, the *Journal of Instructed Second Language Acquisition*, the *Annual Review of Applied Linguistics*, and *System*, as well as in Ahmadian & Long's *Cambridge handbook of task-based teaching* (2022). He is currently working on methods of measuring performance on second language speaking tasks.

CHAPTER 12

Ellen Grote has a background in applied linguistics and has consulted for university research centres and various government departments in Australia. Her research areas include Aboriginal English, cross cultural communication in education, vocational education training, legal and justice contexts as well as Aboriginal Education more generally in schools and tertiary institutions. She has published journal articles and book chapters on Aboriginal English, second language acquisition research and task-based needs analysis for learners of an additional language and/or additional dialect.

Rhonda Oliver completed her PhD under the primary supervision of Mike Long in the mid 1990s. Since that time, she has gone on to undertake extensive research about and is widely published in the areas of second language and dialect acquisition, and task-based language learning especially in relation to child and adolescent language learners in schools and universities. Her more recent work includes studies within Australian Aboriginal education settings. She has also held numerous leadership roles in Western Australian universities including Dean and Director of Research and most recently Head of Education, at Curtin University.

CHAPTER 13

Alessandro G. Benati (see “about the editors”).

CHAPTER 1

The legacy of Professor Michael H. Long and his influence in second language acquisition

Alessandro G. Benati and John W. Schwieter
University of Hong Kong / Wilfrid Laurier University

This chapter will provide a brief overview of the influence of Professor Michael H. Long in the field of second language acquisition whose work has shaped our understanding of many key concepts in the field. The chapter continues with a short synopsis of the contributions in this volume motivated by his work.

What was the main contribution of Michael H. Long to second language acquisition (SLA)? We can't answer this question and summarize his many contributions in a nutshell. Michael Long had a huge influence in the development of many areas of SLA as a field of research, and we must talk about a multi-strands influence of his work on research, theory and language pedagogy. Over the last fifty years, he has coined key terms, developed theoretical accounts and introduced key concepts/ideas which served a springboard for more research, reflections and discussions in our field.

Focus on form is a term first introduced by Michael Long in a paper where he made a clear distinction between 'focus on form' and 'focus on forms' (Long, 1991). Focus on form involves a combination of focus on meaning and a focus on form. More specifically, focus on form is type of pedagogical intervention aiming at drawing L2 learners' attention to the linguistic properties of a language through exposure to meaningful input. Michael Long's definition sparked several branches of investigations on the role of different types of focus on form approaches in SLA. He originally noted (Long, 1983a) that classroom research on the effects of explicit instruction on SLA has principally focused on one main issue: whether focus on form per se makes a positive impact on the acquisition of a second language (L2). The question as to how we should provide focus on form and particularly whether there is one type of instruction which is more effective than others had been somewhat neglected. Michael Long's paper (1983a) on the question as to whether or not instruction makes a difference in SLA was the starting point for a long and fruitful debate in the field and for the start of a number of branches of empirical research

investigating the effects of instruction. This new wave of research systematically described how instruction is operationalized in the classroom, and addressed the question of whether there are particular types of focus on form more beneficial than others (Long, 2016).

Michael Long (2007) has also influenced the field of SLA in contributing to the development of one of the main contemporary theories in this field in the early 1980s (Long, 1983b): the so-called Interaction Hypothesis. According to this theory conversational interaction and negotiation of meaning can facilitate acquisition. Long (1996: 451) has initially suggested that ‘negotiation for meaning, and especially negotiation work that triggers interaction adjustments by the NS (Native Speakers) or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive way. Interaction plays a central role in SLA processes in two ways: (1) by modifying input, (2) by providing feedback related to the linking of meaning and form. Input modification (Long, 1985, 2020) occurs when the interlocutor perceives that the learner does not understand what is being said, and restates something by simplifying, exemplifying, or otherwise altering the original statement. Feedback occurs when the interlocutor uses particular devices to inform the learner about something he or she has said. One of the main assumptions of this theoretical framework is that native speakers and non-native speakers can work actively together to achieve mutual understanding. The Interaction Hypothesis had a huge influence in SLA and language instruction making a number of claims: (i) input plays a crucial role in SLA. Input is crucial but it is not sufficient. Interaction also plays a key role. Through interactions learners may be led to notice things they wouldn’t notice otherwise, and this noticing can affect acquisition; (ii) output is necessary for the development of language. Output modifications through negotiation of meaning help learners to notice important aspects of the target language; (iii) negative feedback obtained during negotiation of meaning might facilitate the acquisition of vocabulary, morphology, syntax, and pronunciation.

Michael Long’s contribution to our understanding of the role of age in second language acquisition has been crucial. His work on the differences of initial rate of acquisition among learners of different ages (Long, 1990) acted as springboard for future research on the nature and effects of age in language development.

Michael Long had also a tremendous influence on language pedagogy contributing to the development of Task Based Language Teaching (Long, 2015, 2016, 2017). Classroom-based research conducted in the 1980s and 1990s has suggested that a focus on the grammatical properties of a target language is a necessary ingredient in L2 development. This research emphasized the need for a ‘focus on form’ (Long, 1991; Long and Robinson, 1998) in L2 teaching; for the need of output

practice; and for the need to enhance the quality of the input received by L2 learners in the classroom.

In the 1980s and the 1990s SLA scholars and researchers also suggested that psycholinguistics factors and processing conditions are relevant for SLA. It was believed that instruction in which L2 learners are given communicative real-world tasks to complete in which they have opportunities for exposure to comprehensible and meaningful input, opportunities for interaction and negotiation of meaning, might ultimately engage natural acquisitional mechanisms, cause a change in L2 learner's interlanguage system and drive forward development. Task-Based Instruction referred to a type of language teaching which takes "tasks" as its key units for designing and implementing language instruction. Task-Based Instruction should enhance language acquisition by providing learners with opportunities to make the language input they receive more comprehensible; furnishing contexts in which learners need to produce output which others can understand; making the classroom closer to real-life language situations. Learners work to complete a task and have plenty of opportunities for interaction and negotiation of meaning as they have to understand each other and express their own meaning. Task-Based Instruction is an alternative model which is based on sound theoretical foundations and one which takes account of the need for authentic communication.

With his famous seminal paper "Inside the black box," Michael Long (1980) examined the importance of carry out language classroom research to make appropriate advances in language acquisition which have direct implications for language instruction. Long's definition and classification of classroom research is that 'the investigation of classroom language learning may be defined as research on L2 learning and teaching, all or part of whose data are derived from the observation or measurement of the classroom performance of teachers and students' (Long, 1980: 3). He compared the classroom as a 'black box' and made a call for more research to systematically describe instruction, practices and behaviors. The mysterious 'black box' in Long's title suggests how little we understand about language classroom processes in the late 1970s, and the need to carry out appropriate empirical research to shed some lights on processes involved in the acquisition of an L2. Michael Long was committed to experimentation as the strongest research design available for instructed L2 research, emphasizing the importance of the combination of a product and process-oriented approach. He also emphasized the importance for a 'process and product' approach to L2 research (Long, 1984).

Michael Long (2005) developed an important rationale and methodology for needs analysis and needs analysis research. He strongly argued about the importance of carefully researching learners' needs as the main prerequisite for developing effective course designs.

One of the most insightful books written by Michael Long was *Problems in second language acquisition* (Long, 2007). Two of the many quotes from this work relates to his concerns of both the future of SLA as a field of enquiry and its implications for L2 language. In relation to the lack of an accepted theory of SLA and a solid empirical base for classroom instruction “teachers and learners have always been, and will always be, vulnerable to drastic pendulum swings of fashion, the coming and going of various unconventional and unlamented “Wonder Methods” being an obvious recent example. The sad truth is that after at least 2,000 years, most language teaching takes place on a wing and a prayer – sometimes successfully, but often a relative failure.” (Long, 2007: 157). Reflecting on SLA research as an important form of enquiry he stated:

whether or not these ideas alone would solve any of the problems discussed, I look forward to the day when SLA is more widely recognized as the serious and socially responsive discipline I believe it can be. Chapters like this one (unpleasant for writer and assuredly some readers alike) would no longer be needed. One could instead concentrate on the genuine controversies and excitement in SLA and L3A: the roles of nature and nurture; special and general nativism; child-adult differences and the possibility of maturational constraints; cross-linguistic influence; acquisition and socialization; cognitive and social factors; resilience; stabilization; fossilization, and other putative mechanisms and processes in interlanguage change; the feasibility of pedagogical intervention; and, most of all, the development of viable theories.

(Long, 2007, 168)

What are the impact and implications of Michael Long’s research and theory for L2 instruction? Through his *Interaction Hypothesis*, he has supported the view that (i) comprehensible input leads to language acquisition. Input becomes comprehensible when L2 learners negotiate meaning with other learners. (ii) Interaction among learners is a crucial and vital factor for acquisition. (iii) A number of strategies can be used in the language classroom which can help instructors to make adjustments in conversation to increase comprehensibility (e.g., confirmation checks; comprehension checks; clarification requests; self-repetitions). (iv) Focus on form is a key feature in L2 instruction particularly because it makes a particular form meaningful and salient in the classroom context. Traditional approaches (focus on forms) to teach grammar and vocabulary in isolation and without meaning do not promote language acquisition. Michael Long has emphasized the fact that traditional form-focused approaches do not work and should not be used alongside task-based teaching.

Michael Long has, with his many contributions, widely influenced SLA and he has indicated the ways for recent and future research in this field. Many questions remain unanswered, whereas others are in need of more complete answers. The field has increased in size and scope, yet it is still sufficiently focused on questions

of learning and teaching for many voices and perspectives to be acknowledged. The richness and complexity of the L2 learning process as a field of study suggest that there are many perspectives to apply and many more applications to be found.

Contributions in the present volume

The chapters in this volume represent and are motivated by the work of Michael Long. Following this introductory chapter, Gass, Plonsky, and Huntley in Chapter 2 offer a bibliometric analysis of Long's research and impact in the field of SLA. Then begin with an objective elaboration of his research areas categorized in three themes (input/interaction, task-based language teaching, and the role of age and aptitude in ultimate L2 attainment) followed by an analysis of his citation record. On this impressive scholarly backdrop, the authors discuss how Long's research has been an influential force in the development of the field of SLA. Chapter 3 by Pienemann, Lanze, Nicholas, and Lenzing echo Long's rejection of fossilization as an explanation or description in SLA. Like Long, the authors argue that a cognitive *stabilisation* mechanism is a viable alternative to *fossilization*. Pienemann et al. demonstrate that linguistic simplification can be simulated using mathematical formalisms from dynamic systems theory and elaborate an agent-based model that is fully operationalized and tested with longitudinal empirical data.

In Chapter 4, Borro addresses an approach to facilitating L2 learning through input known as enhanced incidental learning (EIL). The chapter compares EIL and input enhancement, a practice which aims to raise learner's noticing of formal aspects of the L2 by altering perceptual salience. The author draws on current trends in the literature to argue that because EIL modifies the conditions under which L2 input is processed (Long, 2017), it may give rise to implicit knowledge through detection rather than noticing. Rastelli in Chapter 5 presents three areas in SLA which he has discussed with Long in recent years: The Discontinuity Hypothesis, the need to study 'intra-language' (in addition to 'interlanguage'), and the uncertainty principle. Some of the themes discussed include the re-assembly of L2 competence, the twofold nature of L2 development, and the benefits of studying interlanguage in SLA.

Chapter 6 by Nuzzo looks at how tandem interaction can promote focus on form. Tandem interaction refers to situations where two partners who are each learning each other's language assume the role of a native-speaker teacher. Nuzzo discuss how these semi-spontaneous conversations can help push learners to systematically use focus-on-form techniques in what she refers to as "enhanced interactions." In Chapter 7 by Granena and Yilmaz, the authors compare the psycholinguistic advantages associated with immediate and delayed oral feedback. Drawing on Long's

Interaction Hypothesis (1981), the chapter reviews key factors moderating the effectiveness of negative feedback and discuss arguments for the utility of delayed feedback based on personal interactions with Mike Long. The chapter concludes with a set of future research directions for research on feedback timing.

O'Grady in Chapter 8 discusses the phenomenon of translation and its role in approaches to SLA such as Focus on Form and in the Interaction Hypothesis. The author argues that transfers should be viewed as a processing-driven strategy in which learners transfer processes from their L1 to L2 except in cases where the process is more costly in the L2 than the L1. Drawing on data from a variety of languages, O'Grady situates transfer into a larger emergentist program devoted to processing-based explanations. Chapter 9 by Bylund and Athanasopoulos reviews the role that age of L2 acquisition plays for phenomena of thinking-for-speaking and linguistic relativity. Building on Long's work on age effects (e.g., Long, 1990), the chapter expands the scope to conceptual and cognitive restructuring and suggests innovative methods for future research.

In Chapter 10, DeKeyser addresses age effects in naturalistic and instructed SLA. He reviews recent classroom-based studies demonstrating that children are not necessarily better L2 learners. While this growing body of work appears to call into question the "younger-the-better" belief along with the critical period hypothesis, DeKeyser argues that these two bodies of literature are not contradictory, but rather two sides of the same coin: one cannot be understood without reference to the other. Following this, Chapter 11 by Skehan discusses the effects of task characteristics and task conditions on L2 speaking performance. In the frame of Level's distinction between the conceptualization and formulation stages, the chapter reviews task research findings relevant to L2 pedagogy and presents ways in which L2 learners can achieve a parallel mode of speech..

In Chapter 12 by Grote and Oliver, the authors review theoretical underpinnings of the Task-Based Needs Analysis model, a tool developed by Long for Task Based Language Teaching. The chapter offers contextualized examples demonstrating how the model contributes to task and syllabus design and 'real life' aspects of L2 learning. The volume concludes with Chapter 13, a contribution by co-editor Benati which offers an overview of focus on form. Focus on form is a type of L2 instructional method that explicitly focuses on the presentation and analysis of linguistic form/structure without being contextualized in authentic or communicative situations. Benati discusses whether focus on form can have an effect on L2 development, highlights some limitations in current research, and provides suggestions for future work.

We sincerely hope that the contributions in the present volume not only showcase the breath and impact of Long's crucial research in SLA, but also pay tribute to his legacy. He is the author of the Interaction Hypothesis, and he made key

contributions to our understanding of focus on form, the role of age in second language acquisition and task-based learning and teaching among other areas of second language acquisition theory and research. We really hope that this small contribution will help to cherish his memory and be a springboard for more contributions to celebrate his immense legacy.

References

- Long, M. (1980). Inside the “black box”: Methodological issues in classroom research on language learning. *Language Learning*, 30(1), 1–42. <https://doi.org/10.1111/j.1467-1770.1980.tb00149.x>
- Long, M. (1983a). Does second language instruction make a difference? *TESOL Quarterly*, 17, 359–382. <https://doi.org/10.2307/3586253>
- Long, M. H. (1983b). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126–141. <https://doi.org/10.1093/applin/4.2.126>
- Long, M. (1984). Process and product in ESL program evaluation. *TESOL Quarterly*, 18, 3, 409–25. <https://doi.org/10.2307/3586712>
- Long, M. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input and second language acquisition* (pp. 377–393). Newbury House.
- Long, M. (1990). Maturation constraints in language development. *Studies in Second Language Acquisition*, 12, 251–285. <https://doi.org/10.1017/S0272263100009165>
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). John Benjamins. <https://doi.org/10.1075/sibil.2.07lon>
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. R. Ritchie & T. K. Bhatia (Eds), *Handbook of second language acquisition* (pp.413–68). Academic Press.
- Long, M. (2005). *Second language needs analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299>
- Long, M. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Long, M. (2016). In defense of tasks and TBLT: Non-issues and real issues. *Annual Review of Applied Linguistics*, 36, 5–33. <https://doi.org/10.1017/S0267190515000057>
- Long, M. (2017). Instructed second language acquisition: Geopolitics, methodological issues, and some major research questions. *Instructed Second Language Acquisition*, 1(1), 7–44. <https://doi.org/10.1558/isla.33314>
- Long, M. (2020). Optimal input for language learning: Genuine, simplified, elaborated, or modified elaborated? *Language Teaching*, 53(2), 169–182. <https://doi.org/10.1017/S0261444819000466>
- Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge University Press.

Taking the Long view

A bibliometric analysis

Susan Gass¹, Luke Plonsky² and Lizz Huntley¹

¹Michigan State University / ²Northern Arizona University

By any definition, Michael Long was a giant in the field of second language acquisition (SLA). He was one of the very few who could truly wed language learning and language teaching which he did in a seamless fashion. His impact is far-reaching and envelops a number of different areas as we will show in this chapter. He was remarkable in his encyclopedic knowledge of the field (and beyond), including access to what had come before and what was current. It was common to hear him say “As so and so showed in 2003...” with his listeners often thinking “How did he remember that?” The first author of this paper was always in awe of his numerous contributions to the fields of language learning and language teaching and his ability to see the big picture. We started our careers at about the same time (Gass received her PhD in 1979 and Mike received his in 1980) and our interests in those early years paralleled each other’s. His work influenced the work of Gass and one would hope that there was bidirectionality in this influence. Our paths crossed personally and professionally and so his death is both a personal loss as well as a tremendous loss to the field (see Ortega’s, 2021 tribute). In this paper we will have the easy task of demonstrating his impact. We include an objective consideration of his areas of influence as well as an analysis of citations of Mike’s work. Finally, we discuss how his research areas have been an influential force in the development of the field of SLA.

1. Introduction: Long’s impact

1.1 Citations

In our attempt to understand Mike’s influence, we availed ourselves of three databases, Web of Science, Google Scholar, and Scopus (see Section 2.1 below for a discussion of the limitations of databases).

There are many ways that a scholar’s impact can be determined. An initial look at his Google Scholar page shows an astounding number of citations (as of

this writing, nearly 59,000).¹ Figures 1 (from Google Scholar, 1991-present) and Figure 2 (from Web of Science, 1981-present) show the number of citations over the years. We present both figures given that the databases are different. Figure 1 shows a more complete picture of actual citations. As can be seen, there is a steady increase in the number of citations from 1981 until 2016, with the number remaining remarkably high over the past 5 years. The bars in Figure 2 show the publications during the 1981-present timeframe and the line demonstrates his citations from 1981 to the present. The left y-axis represents the number of publications; the right y-axis represents the number of citations. With a few peaks and valleys, the line shows that his influence has steadily increased over the years and to this date remains high.

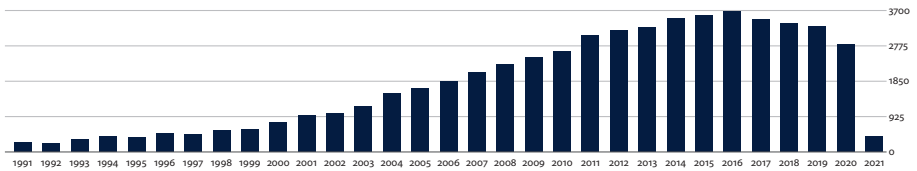


Figure 1. Publications from Google Scholar (1991-present)

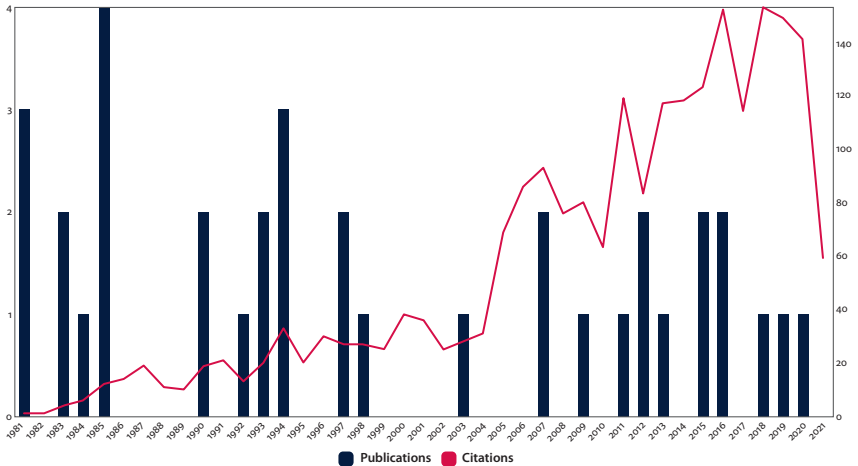


Figure 2. Publications from Web of Science (1981-present)

1. During the writing of this article, Mike Long's Google Scholar page was deleted. This forced us to use either Web of Science or Scopus. During the process, we worked extensively with Clarivate (owners of Web of Science) to make sure that Mike's papers were included in their database. We are grateful to them for their help, but in the end, the database for Mike's papers was insufficient for us to continue to use that database for our analysis. In the end, our bibliometric analysis in Section 3 is based on the Scopus database.

Not only are the citations high, but, as we will show below, Mike's influence extended across numerous areas of research. He is one of the most cited authors in multiple topics within the field. In Table 1, we provide citation counts (as of August 13, 2021) of his works (reflected on Google Scholar) that have received 1000 or more citations. These counts indicate quite compellingly that his work was influential in a range of domains including (1) ultimate attainment, (2) input/interaction, (3) instruction, (4) TBLT, as well as (5) general SLA-related research.

Table 1. Counts of Long's most-cited works in Google Scholar

Publication	Year	Topic	# citations
Age, rate and eventual attainment in second language acquisition (Krashen, Scarcella, & Long)	1979	Age-related acquisition	1109
Input, interaction, and second language acquisition (dissertation)	1980	Interaction	2333
Linguistic and conversational adjustments to non-native speakers	1983a	Interaction	1165
Native speaker/non-native speaker conversation and the negotiation of comprehensible input	1983b	Interaction	2553
Does second language instruction make a difference?: A review of research	1983c	Second language instruction	1415
Group work, interlanguage talk, and second language acquisition (Long & Porter)	1985	Interaction	2128
Maturational constraints on language development	1990a	Age-related acquisition	2128
An Introduction to second language acquisition Research (Larsen-Freeman & Long)	1991	General Book	5229
Three approaches to task-based syllabus design (Long & Crookes)	1992	TBLT	1575
The role of the linguistic environment in second language acquisition	1996	Interaction	6636
The Handbook of second language acquisition (Doughty & Long)	2003	General Book	3428
Problems in SLA	2007	General Book	1099

In sum, his writing covers a broad range of topics and the citations are numerous. Many of these works appeared when the field was developing, suggesting that Mike was involved in early research that has shaped the direction SLA research has taken as a whole.

1.2 Indices

But there are other indications of impact as well. Hirsch (2005) proposed the h-index which takes into account both the number of papers one publishes and the number of citations those publications receive. In Mike's case, his h-index (from Google Scholar) is 74 which means that he has had 74 papers that have been cited at least 74 times. But more important is the interpretation. According to Hirsch, "an h index of 20 after 20 years of scientific activity, characterizes a successful scientist" (p. 16571); "an h index of 40 after 20 years of scientific activity characterizes outstanding scientists" (p. 16571), and "an h index of 60 after 20 years, or 90 after 30 years, characterizes truly unique individuals" (p. 16571). Translating these suggested values into m-quotients which take into account the length of an individual's career, the first case (h-index of 20 after 20 years) yields an m-quotient of 1; the second case yields an m-quotient of 2 and the third scenario an h-index of 90 after 30 years yields an m-quotient of 3. Mike's m-quotient of nearly 2 puts him in the category of outstanding.

Using a different index, his i10-index of 116 shows that 116 of his works have been cited at least 10 times, another extraordinary count.

1.3 Other indications of impact

de Bot (2015) produced an admittedly somewhat subjective list of leaders in the field in his survey of applied linguistics covering a period of approximately 30 years. He distributed questionnaires to 68 researchers (56 were returned) in the field and conducted one-on-one interviews with 38 others. His goal was to determine 'leaders' in the field.² The following were the questions related to leaders:

- Who are the most important/influential leaders in the field? Why are they important?
- Who have you been most influenced by?
- What are the 5–10 most important articles for you over these 30 years?
- What are the 5–10 most important books for you over these 30 years?

Not surprisingly, of the 58 scholars whose name was mentioned 4 or more times on the course of interviews or in responses on the questionnaire, Mike was the fifth most common individual mentioned. de Bot also lists the h-index and number of

2. This is not the time nor place to discuss the limitations of de Bot's data collection. The interested reader is referred to his work where he discusses the biases and limitations.

citations to the top 25 scholars on his list. Mike with his h-index of 68 (note that this statistic came before the current h-index of 74 mentioned above) was the second highest of all scholars, demonstrating again Mike's influence on the field and the high regard his colleagues have of him.

There are two recent bibliometric analyses that similarly point to Mike Long's influence: Lei and Liu (2019) and Zhang (2020). Lei and Liu examined applied linguistic bibliographic data from 2005–2016. One of the questions they asked had to do with the publications and authors who have been the most highly cited (in their terms, suggesting greatest influence). They divided their time period in groups of 4 years (2005–2008, 2009–2012, and 2013–2016) and listed the top 20 most cited items in each period. Mike's 1996 article on interaction appeared on all three lists.

A second analysis that they performed was to determine the 20 most highly cited authors. The scope of their analysis was quite broad and went beyond SLA to include the whole of applied linguistics. Thus, to determine Mike's "ranking" within the field of SLA, we removed names to those who might be better thought of as linguists (Chomsky) or non-SLA applied linguists (Hyland). Mike was number four on that list.

The second bibliometric analysis we looked at was Zhang (2020), whose analysis consisted of publications between 1997 and 2018 with data coming from 16 journals. There are a number of important points in Zhang's article relevant to empirically establishing Mike Long's shaping of the field, one of which is Table 2 (modified from Zhang's article). He divided the research timeframe under investigation into two 10-year periods and counted citations to scholars in each period. We took the top 15 in each list and, as can be seen, Mike is one of only a handful of scholars to appear on both lists (#1 in the first 10-year period and #5 in the second).

Consistent with what we reported earlier, Mike's work spans the boundaries of intellectual SLA subdisciplines. Zhang used co-citation information to determine clusters in each of his two 10-year periods (Small, 1973). In the first 10-year period, four clusters emerged and Mike's work appears in two of them. The largest cluster in the first 10-year period is referred to as the interaction/communicative competence cluster; of the 21 articles that appear therein, Mike is author or co-author of four. These works include topics that cross subdisciplines of language teaching and interaction. In the second 10-year period, the interaction/communicative competence cluster is smaller, including 11 articles, but Mike's influence is even stronger with three of those articles having Mike as author. Again, both language teaching and interaction are part of Mike's influence in this cluster.

Table 2. The most influential scholars in the two time periods

1997–2007			2008–2018			
Author	Raw citation	Norm. citation	Author	Raw citation	Norm. citation*	
1	Long, M.	616	302	Ellis, R.	1822	313
2	Ellis, R.	554	271	Dörnyei, Z.	1537	264
3	Swain, M.	517	253	Swain, M.	1243	213
4	Krashen, S.	428	210	Ellis, N.	1135	195
5	Pica, T.	400	196	Long, M.	1046	180
6	Gass, S.	386	189	Bialystok, E.	960	165
7	Gardner, R.	358	175	DeKeyser, R.	816	140
8	Schmidt, R.	326	160	Robinson, P.	812	139
9	VanPatten, B.	320	157	Nation, I.S.P.	781	134
10	Oxford, R.	299	146	Flege, J. E.	768	132
11	Doughty, C.	272	133	Schmidt, R.	768	132
12	Flege, J. E.	251	123	Skehan, P.	752	129
13	Ellis, N.	250	122	Gass, S.	711	122
14	Lyster, R.	236	116	Mackey, A.	699	120
15	Dörnyei, Z.	222	109	Laufer, B.	685	118

* Normalized citation is in per thousand abstracts

1.4 Impact beyond publications

The development of disciplines is dependent on the strength of the scholars whose innovative work forms their basis, but there are other ways that a discipline develops. For example, disciplines need journals and associations devoted to their work. They also need intellectual and institutional homes (namely, university settings and learned societies) where work is fostered and thrives. Mike has been instrumental in this realm as well. Thomas (2013) surveyed Ph.D. programs specifically in SLA or Second Language Studies (SLS). She isolated 9 programs, of which two were founded during Long's tenure at those institutions: University of Hawai'i at Manoa (the oldest of the 9 programs) and University of Maryland at College Park (the youngest). As Thomas noted,

In the context of this article, it is the 1988 launching of the PhD program in SLA that is most significant, since that shift acknowledged the legitimacy of study of the general phenomena of second language acquisition at the doctoral level. Insofar as the field wants to identify the anniversary of its modern institutionalization, the 1988 opening of the University of Hawai'i PhD program in SLA provides a convenient date. (p. 514)

Ortega (2021) helps us understand Long's role in the establishment of this program by referring to him as "one of the founders of the oldest doctoral programs in SLA in the USA" (p. 1). Mike left the University of Hawai'i in 2003 when he founded another Ph.D. program at University of Maryland. Thomas (2013) points out the distinctiveness of that program: "It declares its strong cognitive science orientation, and highlights research on L2s other than English" (p. 522). This clearly shows Mike's underlying emphasis on the need to apply "perspectives from philosophy of science and cognitive science... to the evaluation of epistemologies for SLA and to the examination of theoretical progress and disciplinary change" (Ortega, 2021, p. 1) (see Long, 1990b, 2007). As Long (2007, p. vii) notes in the introduction of his book that the book itself is "a proposal concerning how the field might develop greater coherence and a clearer focus than it has now, and do so systematically, in part through guidance from work in the philosophy of science." Thus, one can clearly see how his intellectual orientation helped shape a new Ph.D. program and, in doing so, an entire generation of SLA scholars.

2. Bibliometric analysis

Long's influence can be observed throughout much of the fields of L2 teaching and L2 learning. However, for the purpose of our analysis, we wanted to identify – ideally using empirical means – the domains where he had been most active and where his influence was most clearly felt. As linguists, it is perhaps not surprising that we took a linguistic approach to doing so.

Our first step was to compile the titles of each of Long's books, articles, and book chapters, extracted from his CV. Our assumption here was that the language (i.e., words and phrases) in these titles would shed light on the topics Long was most interested in. This small but exhaustive corpus was then imported into LancsBox (Brezina, Weill-Tessier, & McEnery, 2020), which provided a list of every word in the corpus, ranked by frequency. Lancsbox also produced a ranked list of bi-grams that occurred two or more times.

Words of a general-SLA nature such as "language", "acquisition", and "teaching" (ranked 1, 7, and 12, respectively) and phrases such as "in SLA" and "review of" (ranked 11 and 17) were removed along with function words. The remaining items in these two lists were then closely examined. In doing so, three main themes emerged: Input/interaction, task-based language teaching (TBLT), and the role of age and aptitude in ultimate attainment.³ The words and phrases in the titles of Long's works that we determined to fall into each of these themes are provided in Table 3.

3. In actuality, a fourth theme emerged, needs analysis. However, we decided not to perform a citation analysis on this topic given its extreme breadth and relevance over many fields unrelated

Table 3. Themes in Long’s work and corresponding words/bi-grams from the titles of his publications

Topic 1: Input/interaction	Topic 2: TBLT	Topic 3: Ultimate Attainment: Age and aptitude
Input	Task-based (language)	Language aptitude
Feedback (on) form	TBLT	Sensitive periods
Foreigner talk	Group work	Advanced
Conversational	Task	(L2) Attainment
Simplified		Maturational constraints
Implicit negative		Age
Interaction		Constraints
Negative feedback		Ultimate L2
Recasts		Critical period

These topics were corroborated in Ortega’s tribute (2021) to Long. In addition to his work on interaction and TBLT, she notes “Mike continued his influential and prolific work, adding new knowledge across a variety of topics central to both SLA and his work, such as the critical period hypothesis, fossilization, aptitude for language learning and the advanced learner” (p. 1).

Our next step was to perform a bibliometric analysis. More specifically, we further investigated Long’s impact in these three areas by looking at the number of citations of his works within each one. What we intend to show is that not only does he have a diverse range of research areas, but that each of these research areas has had a large impact on others work in the field.

2.1 Data

The data for the present analysis were accessed and downloaded at the end of July, 2021.

We consulted Web of Science, Google Scholar, and Scopus to inform our discussion of Long’s work. Because these sources approach the curation and dissemination of publication data differently, the use of all three allows us to more fully capture the extent of his scholarly contributions. Notably, the Web of Science and Scopus databases are human-curated, whereas Google Scholar automatically indexes documents based on web crawler data. This means that Google Scholar data may contain duplicate entries or non-scholarly citation sources, but it can also provide insights into a researcher’s broader influence beyond traditional outlets such as journal articles and even beyond the Ivory Tower in the form of blogs, white papers, and technical reports. On the other hand, Google Scholar does not, to date, permit the bulk data extraction offered by Web of Science and Scopus

(Martín-Martín et al., 2018). In our comparison of Long’s citations covered in the human-curated databases, Web of Science identified 36 references whereas Scopus identified 41. These counts pale in comparison to what we had counted based on Long’s professional website (144 total: 19 books and monographs, 70 book chapters, 55 journal articles and reviews). Although Scopus and Web of Science undercount Long’s publications and citations thereof, the extent to which others have cited Long can still be represented through an analysis of citing references; that is, the number of times that Long’s work appears in any reference list of any article included in the database. This is because citations pulled from a reference list are included in a citation count regardless of whether or not the reference was independently indexed in the database.

With the above discussion in mind, we conducted a citation analysis of Long’s work from the Scopus database. We utilized Scopus because it contained a broader index of Long’s work. The first step was to determine the search terms needed to accurately identify articles in the three key areas previously identified: task-based language teaching, input/interaction, and age & aptitude. For all three domains, we began by searching for any occurrence of related terms in the title, abstract, or keywords of any article in the Scopus database. This search would then provide us with a list of all articles matching the search criteria, as well as access to the number and names of references cited by those articles (see Table 4).

Table 4. Summary of the scopus database search results

Domain	Search terms	Subject area	Article matches	Cited references
TBLT	“task-based language”	(n/a)	553	13,883
Input/ Interaction	“conversational interaction” or negotiation and L2 or SLA or “second language”	social sciences, arts & humanities, psychology	418	6,629
Age & aptitude	“ultimate attainment” and L2 or SLA or “second language”	social sciences, arts & humanities, psychology, neuroscience	97	4,091

Given that task-based language teaching has spawned a number of related sub-domains (task-based language assessment, task-based language learning, etc.) we searched for the broader term “task-based language.” Finally, to adequately capture the domains of input / interaction and age & aptitude in the field of SLA, we searched for combinations of terms that most reliably returned topically relevant journal articles while also delimiting the search to the most closely related subject areas such as “Social sciences” and “Arts and Humanities” as defined by Scopus.

2.2 Analysis

To analyze Long's influence in these three domains, we uploaded the Scopus search results (including article matches and cited references) to VOSviewer (van Eck & Waltman, 2021). VOSviewer is a freely available software tool for visualizing bibliometric networks.

For each of the three domains that we identified as most prominent within Long's work, we created a map based on the cited reference data. In order to visualize the impact of Long's most influential works, we selected "documents" rather than "authors" as the unit of analysis (e.g., Wang & Zhang, 2019). We did not set any limits on the maximum number of references any one document could contain. For the sake of visual clarity, however, we did limit the minimum number of times a reference needed to be cited across all articles in order to appear in the map. Finally, we restricted the output to the largest connected group of cited references to avoid mapping unrelated clusters. This resulted in a 28-item map for the domain of TBLT (citation count minimum: 30), a 79-item map for input / interaction (citation count minimum: 20), and a 28-item map for ultimate attainment (citation count minimum: 10).

In the following sections, we present temporal overlay maps. A visualization of bibliometric networks typically consists of two elements: nodes (the units being directly measured or quantified; in our case each node represents a single publication), and links (lines which connect individual nodes). The size of the nodes represents the relative weight of each unit, hence in the figures below a larger circle indicates a larger number of citations for that particular publication. The links in the citation networks depicted here indicate that one publication has cited the other. Finally, the coloring of the figures illustrates the year of publication. Darker shades represent older publications, while lighter shades represent newer publications. The range of publication years differs slightly for each figure depending on the set of nodes in each domain network; please see the key in the lower right-hand corner of each figure for exact years.

2.2.1 *Task-based language teaching and learning*

Figure 3 shows Long's (1992) article with Graham Crookes as having the largest number of citations. In addition, Long appears in three other highly-cited articles: Doughty and Long (2003, the 4th most cited paper), Long (2016, the 14th most cited), Serafini, Lake, and Long (2015, the 27th most cited paper). Taking all 50 articles that appeared in our search, we further determined the most highly cited scholars; Long, with a total of 559 citations to his four publications, topped the list.

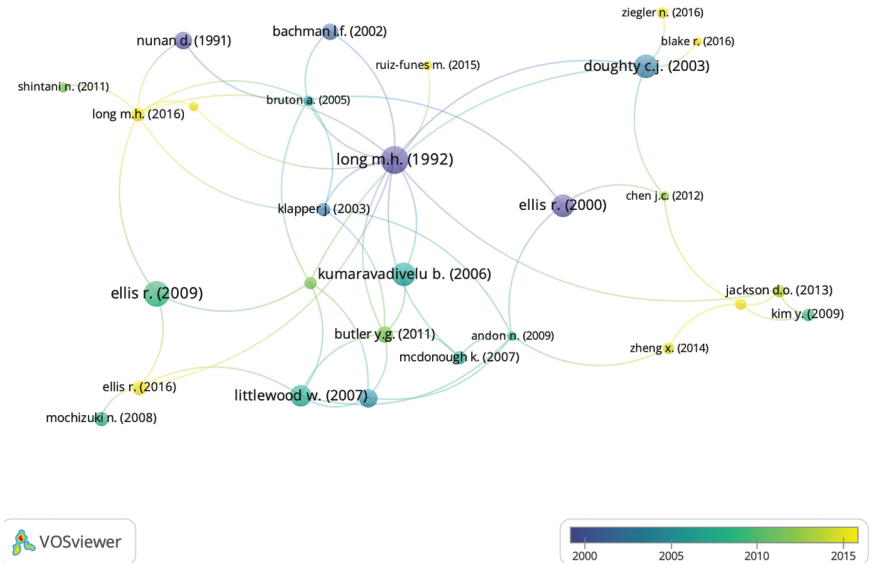


Figure 3. Bibliometric visualization of top SLA citations in the domain of TBLT

2.2.2 *Input/interaction*

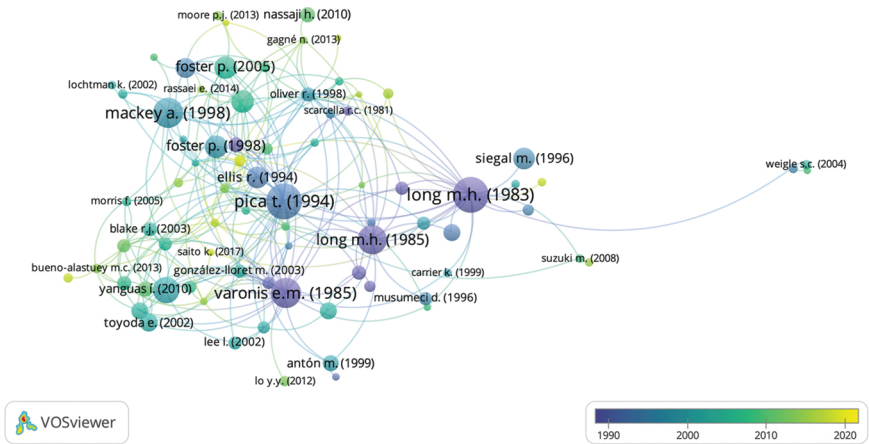


Figure 4. Bibliometric visualization of top SLA citations in the domain of input/interaction

Within the interaction research area, 79 papers met our search criteria (Figure 4), with three of Long's papers appearing on the list (1981a, 1983b, 1985). Another way of emphasizing his impact is to take the five most cited papers in this research

area (see Table 5). The total number of citations is 2053. Long's works represent 40% of those citations.

Table 5. Top five publications by citation count in the domain of input/interaction

Author	Publication	# of citations
Pica	1994	504
Long	1983b	499
Mackey	1998	364
Varonis & Gass	1985	355
Long	1985	331

2.2.3 *Ultimate attainment*

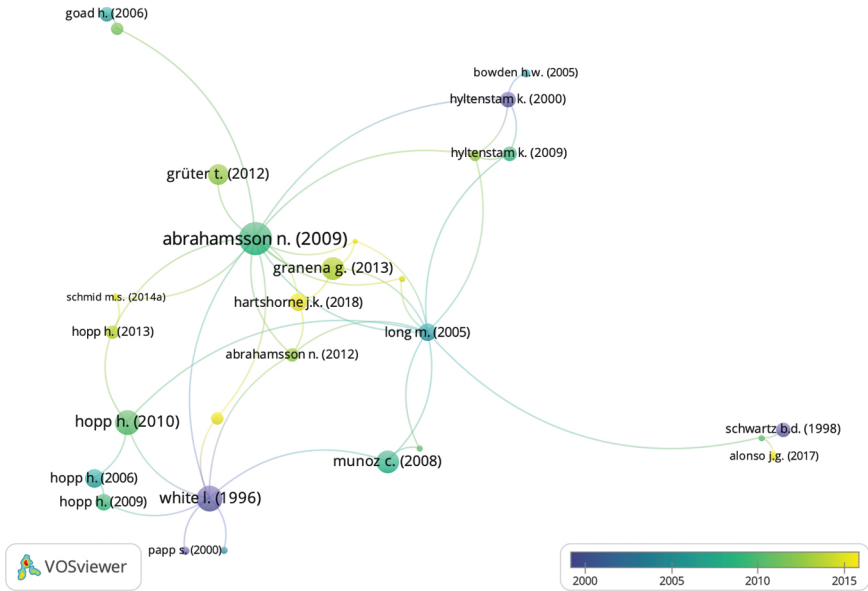


Figure 5. Bibliometric visualization of top SLA citations in the domain of ultimate attainment

As can be seen in Figure 5, similar to what we have found in the other two areas, Long's work is central to the research area which we have called ultimate attainment and which includes age and aptitude. There are two publications that appeared on the list of 26 highly-cited articles in this category, Long (2005), the 9th highest cited article and Granena and Long (2013), the 4th highest cited publication.

3. Discussion

3.1 Task based language teaching and learning

As noted above, three of Mike's papers are highly cited in the research area of TBLT. Additionally, it is important to note that Mike has been writing about classroom teaching since the mid-1970s (e.g., Long, 1975). Long and Crookes (1992) establish an important link between syllabus design and SLA. Taking "task" as the major unit of analysis for syllabus design, they argue that a task syllabus (when combined with a focus on form, Long, 1991) receives the greatest support (when compared to procedural and process syllabi) from an SLA perspective. The use of SLA research to justify language teaching approaches is a hallmark of Long's work. As Long and Crookes argue, "TBLT is distinguished by its compatibility with research findings on language learning, a principled approach to content selection, and an attempt to incorporate findings from classroom-centered research when making decisions concerning the design of materials and methodology" (pp. 45–46).

Doughty and Long (2003) carry the TBLT approach into distance learning programs, and particularly programs for the teaching of less commonly taught languages. They do this through an examination of 10 methodological principles (MP) necessary for TBLT, although they recognize that it is likely that, with further research, not all of these methodological principles will be well founded:

1. use task, not text, as the unit of analysis;
2. promote learning by doing;
3. elaborate input;
4. provide rich input;
5. encourage inductive ("Chunk") learning;
6. focus on form;
7. provide negative feedback;
8. respect developmental processes and "learner syllabuses";
9. promote co-operative/collaborative learning;
10. individualize instruction

Important for our purposes is the continued recognition of the need for underlying theoretical research. As Doughty and Long (2003) tell us:

... selections among the ever increasing range of technological options in distance learning need to be theoretically and empirically motivated, not simply market-driven, as is too often the case. As a basis for decision-making in the creation of optimal psycholinguistic environments for distance foreign language learning, theory and research in SLA, educational psychology, and other areas of cognitive science will clearly be foundational. The 10 MPs of TBLT reflect one attempt to integrate interpretations of such theory and research findings into a coherent design for the delivery of instruction. (pp. 67–68)

The third paper in the TBLT category is one published in 2016 in which he takes a serious look at criticisms of TBLT, some of which he argues are nonissues. He further argues that the methodological principles are “motivated by what ... SLA research has shown about how children and adults learn L2s successfully (Long, 2009, 2015, pp. 300–328) and independently by principles from the philosophy of education” (Long, 2015, pp. 63–83). He goes on to debunk many of the criticisms of the MPs and other criticisms of TBLT. True to character, Long concludes the article with true problems related to TBLT (e.g., establishing criteria for task complexity) and points the way to a deeper understanding of TBLT in the future: “Advances in theory and research, coupled with further field trials, will assuredly refine current models, and quite probably identify needed changes. The goal is for researchers and practitioners to move forward together systematically in what must be a collaborative endeavor” (pp. 28–29).

Finally, Serafini, Lake, and Long (2015), in a paper that focuses on a needs analysis, highlight the need for methodological rigor. They do this through an examination of TBLT practices. The paper, while only incidentally involving TBLT, provides a unique and detailed methodology for conducting a needs analysis with English for specific purposes learners. True to form, Long and his co-authors focus on rigorous methods that adhere to standards for reliability and validity.

A recent volume (Ahmadian & Long, 2021), with its list of well-regarded international scholars, is likely to provide a definitive statement of TBLT.

3.2 Input/interaction

In this section, we discuss three of Long’s papers that were on the list of most cited based on our bibliometric analysis. In addition, we discuss his 1996 paper. We feel justified to include this paper for two reasons: (1) Lei and Liu’s article, cited above, placed this paper on their list of most highly cited papers in the three time periods they examined and (2) because, as can be seen in Table 1, it is the most highly-cited paper.

In 1981, Long published a paper (1981a) that had been written while he was a Ph.D. student at UCLA. It is clearly a precursor to his dissertation work. This was an early exploration into differences between structures of native speaker (NS) NS-NS conversations and conversations involving non-native speakers (NNS). This study provided an analysis that suggested differences in discourse structures between NS-only conversations and those involving learners. He further hinted at the role of input available to learners and the order of appearance of linguistic features. This, coupled with his dissertation work (1980, 1981b) set the stage for interaction-based research.

Two years following the publication of his 1981a paper, Long published a paper in *Applied Linguistics* that showed the development of his thinking on this topic. In this paper he clearly demonstrates the importance of interactional modifications, as opposed to input modification when considering NNS comprehension. He provides the following example to exemplify this:

- NS: When did you finish?
 NNS: Um?
 NS: When did you finish?
 NNS: Ten clock
 NS: Ten o'clock?
 NNS: Yeah

Here the NS repeats the question following an indication of non-comprehension by the NNS. When the NNS responds with a somewhat accurate response (*ten clock*), the NS responds with a confirmation check to make sure that comprehension had taken place. Thus, the interactional structure differed from what one would expect in a conversation involving two NSs, but there was no input modification. As he notes in the article “[m]odified interaction is observed in cases where some kinds of input modification usually considered ‘basic’ (e.g., reduced syntactic complexity) are absent” (p. 131). He takes his data beyond a description of the conversational adjustments made by determining strategies and tactics that native speakers use to make input comprehensible. Modifications to input are clearly important, but Mike paved the way to an understanding of the types of modifications to discourse that occur.

Long and Porter (1985) take his research into the classroom and, consistent with Mike’s emphasis on theory, provides a psycholinguistic rationale for group work. The rationale itself stems from SLA research on conversations. The authors review the work on negotiation and “interlanguage talk” and argue that there needs to be a theoretical basis for pedagogical practices.

We would be remiss if we didn’t discuss Mike’s 1996 paper which, given the vagaries of using Scopus, did not appear on the list generated by the Scopus database, but which, as noted earlier (Table 1), is in actuality his most highly-cited publication. This article is a seminal one in the history of work on interaction, as is indicated by the very large number of citations. He draws on numerous theoretical constructs from other fields, such as negative and positive evidence, attention and awareness. He ties these constructs to focus on form about which he had previously written in relation to TBLT. In this article he proposes what he called *The Interaction Hypothesis* and famously claims that “*negotiation for meaning*, and especially negotiation work that triggers *interactional* adjustments by the NS or more

competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways” (pp. 451–452). This statement and the arguments presented in this article have been extraordinarily influential in all future work on input and interaction.

3.3 Ultimate attainment

Two of Long’s papers were revealed in Figure 5 to be central to research at the intersection of age, aptitude, and ultimate attainment. The earlier of these, Long (2005), is classic Mike in that it systematically addresses a set of complex theoretical issues related to research on the critical period hypothesis and its effects on L2 development. More specifically, Mike takes on a wide-reaching set of what he viewed as misconceptions in this domain. Also in classic form, he does so in a critical, systematic, and highly insightful manner, breaking down and refuting claims against the existence of a critical or sensitive period. The arguments are woven together to seamlessly blend logical, theoretical, and empirical evidence. In a number of instances, Mike also highlights methodological weaknesses (e.g., sampling, measurement, potentially bias among raters) among studies that claim to have found evidence against the critical period. His consideration of the theoretical issues in play as well as of the empirical studies that address them is meticulous throughout, down to level of individual test items and statistical results.

Complementary to the highly conceptual and methodological orientation of Long (2005) is Granena and Long (2013), an empirical study that also appeared as prominent in Figure 5. Employing a cross-sectional design, the authors examined ultimate attainment among Chinese learners of Spanish as a function of (a) age of onset, (b) length of residence in the target language community, and (c) aptitude. Critically, a variety of linguistic domains were investigated thus allowing Granena and Long to demonstrate differential effects of age and aptitude on L2 phonology, morphosyntax, and lexis. As we might expect, the effect of age was strongest for pronunciation. Contrary to much of the prior research and theorizing in this area, the findings also provide compelling support for the existence of multiple sensitive periods that correspond to different linguistic domains.

Conclusion

Mike was, simply put, a titan in the field of SLA. Given the robust health and status of SLA today, it is difficult to imagine the field's somewhat fledgling status as recently as three or four decades ago. To be sure, Mike was absolutely integral in its maturation toward SLA's current status as an independent and now-thriving discipline, from program building and theoretical development to conference organizing and, of course, writing and publishing. As the bibliometric analysis presented here shows, Mike was prolific in his scholarly efforts, leading to a vast body of work and intellectual influence in SLA both as a whole as well as in a range of unique domains. Evidence of his influence can be seen in part in the citation-based data described throughout this paper.

Those of us who had the privilege of knowing Mike knew that he had a one-of-a-kind personality. However, as we hope to have demonstrated in this chapter, Michael H. Long was also a one-of-a-kind scholar.

References

- Ahmadian, M. J., & Long, M. H. (Eds.). (2021). *The Cambridge handbook of task-based language teaching*. Cambridge University Press. <https://doi.org/10.1017/9781108868327>
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2020). #LancsBox v. 5.x. [software]. Retrieved on 1 February 2022 from <http://corpora.lancs.ac.uk/lancsbox>
- De Bot, K. (2015). *A history of applied linguistics*. Routledge. <https://doi.org/10.4324/9781315743769>
- Doughty, C. J., & Long, M. H. (Eds.). (2003). *The handbook of second language acquisition*. Blackwell. <https://doi.org/10.1002/9780470756492>
- Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, 7(3), 50–80.
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29, 311–343. <https://doi.org/10.1177/0267658312461497>
- Granena, G., & Long, M. H. (Eds.) (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. John Benjamins. <https://doi.org/10.1075/lllt.35>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>
- Krashen, S. D., Long, M. H., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language learning. *TESOL Quarterly* 13(4), 573–82. <https://doi.org/10.2307/3586451>
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. Longman.
- Lei, L., & Liu, D. (2019). Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics*, 40(3), 540–561. <https://doi.org/10.1093/applin/amy003>

- Long, M. H. (1975). Group work and communicative competence in the ESOL classroom. In M. Burt & H. Dulay (Eds.), *On TESOL '75: New directions in second language learning, teaching and bilingual education* (pp. 211–23). TESOL.
- Long, M. H. (1980). Input, interaction and second language acquisition (Unpublished doctoral dissertation). University of California, Los Angeles.
- Long, M. H. (1981b). Input, interaction, and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences*, 379, 259–278. <https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>
- Long, M. H. (1981a). Questions in foreigner talk discourse. *Language Learning*, 31(1), 135–157. <https://doi.org/10.1111/j.1467-1770.1981.tb01376.x>
- Long, M. H. (1983a). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5(2), 177–193. <https://doi.org/10.1017/S0272263100004848>
- Long, M. H. (1983b). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126–141. <https://doi.org/10.1093/applin/4.2.126>
- Long, M. H. (1983c). Does second language instruction make a difference? A review of research. *TESOL Quarterly*, 17(3), 359–382. <https://doi.org/10.2307/3586253>
- Long, M. H. (1990a). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251–285. <https://doi.org/10.1017/S0272263100009165>
- Long, M. H. (1990b). The least a second language acquisition theory needs to explain. *TESOL Quarterly*, 24(4), 649–666. <https://doi.org/10.2307/3587113>
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). John Benjamins. <https://doi.org/10.1075/sibil.2.07lon>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Long, M. H. (2005). Problems with supposed counter-evidence to the Critical Period Hypothesis. *International Review of Applied Linguistics and Language Teaching*, 43(4), 287–317. <https://doi.org/10.1515/iral.2005.43.4.287>
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. H. (2009). Methodological principles in language teaching. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 373–394). Blackwell. <https://doi.org/10.1002/9781444315783.ch21>
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Long, M. H. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, 36, 5–33. <https://doi.org/10.1017/S0267190515000057>
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based language teaching. *TESOL Quarterly* 26(1), 27–56. <https://doi.org/10.2307/3587368>
- Long, M. H., & Porter, P. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207–227. <https://doi.org/10.2307/3586827>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- Ortega, L. (2021). In celebration of Michael H. Long (1945–2021). *System*. <https://doi.org/10.1016/j.system.2021.102550>

- Serafini, E., Lake, J., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, 40, 11–26. <https://doi.org/10.1016/j.esp.2015.05.002>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the Association for Information Science and Technology*, 24, 265–269.
- Thomas, M. (2013). The doctorate in Second Language Acquisition: An institutional history. *Linguistic Approaches to Bilingualism*, 3, 509–531. <https://doi.org/10.1075/lab.3.4.05tho>
- Van Eck, N. J., & Waltman, L. (2021). VOSviewer (1.6.17) [Computer software]. Universiteit Leiden. Retrieved on 1 February 2022 from <https://www.vosviewer.com/>
- Wang, L., & Zhang, L. J. (2019). Chapter 8. Peter Skehan’s influence in research on task difficulty: A bibliometric analysis using CiteSpace. In Z. (Edward) Wen & M. J. Ahmadian (Eds.), *Task-based language teaching* (pp. 183–195). John Benjamins.
- Zhang, X. (2020). A bibliometric analysis of second acquisition between 1997 and 2018. *Studies in Second Language Acquisition*, 42, 199–222. <https://doi.org/10.1017/S0272263119000573>

Stabilization

A dynamic account

Manfred Pienemann^{1,2}, Frank Lanze², Howard Nicholas³
and Anke Lenzing⁴

¹Paderborn University / ²Linguistic Engineering Co. / ³La Trobe University /
⁴Innsbruck University

We accord with Mike Long's rejection of fossilization as a concept able to describe or explain second language acquisition. And we share his bewilderment (Long, 2003, p. 512f.) that since most studies of fossilization make reference to explanatory factors external to language and cognition, "[s]urprisingly, no one seems to have considered the possibility that if fossilization is, as Selinker (1972) claimed, a cognitive mechanism producing the non-target end-state also called 'fossilization', there is no need for other explanations ...". In this paper we take up Mike's proposition that stabilization might be a more viable alternative to fossilization and propose exactly what Mike asked for, a cognitive mechanism. We demonstrate that in SLA, linguistic simplification, one aspect of stabilization, is a dynamic process that follows its own regularities. We show that it can be modeled in an AI simulation of SLA using the mathematical formalisms of dynamical systems theory that are implied in agent-based modeling. In doing this we show that the formal, mathematical architecture of dynamical systems theory is particularly well suited for a simulation of the cognitive stabilisation mechanism that Mike Long asked for, because agent-based modeling can operate entirely on the basis of the internal dynamics of identifiable stabilization mechanisms, and they can lead to tipping points at which the system may abruptly change direction. Differing from the postmodern DST metaphors that are currently popular among some applied linguists (e.g., Larsen-Freeman, 2006), we have followed Long's (2003) call for an operationalized concept that is testable. In this chapter, our agent-based model is fully operationalized and tested against longitudinal empirical data. Our chapter spans a period of four decades during which the first author and Mike interacted. Many of these interactions left traces in the research described here. Contextualising this research in its 'historical' background may be helpful in tracing the development of ideas that lead up to the AI simulation of SLA reported in this chapter. As the reader will see, many of these ideas were intentionally or unintentionally inspired or provoked by Mike's critique, comments, suggestions and the presence of his critical mind.

1. Mike Long's impact on PT and the way it conceptualizes variability

Everyone who collaborated and interacted with Mike knows what a great source of inspiration he was, not least due to his forceful and witty style of argument as evidenced, for instance, in the title of his 1997 EUROSLA plenary address: "Fossilization: Rigor mortis in living linguistic systems?" What an acute one-line observation about this popular paleontological metaphor – as a brief look at *The digital atlas of ancient life*¹ reveals: "The rapid burial of remains beneath a blanket of sediment is critical to the process of fossilization...". Mike's punch line goes straight to the heart of this badly conceived Jurassic metaphor. Fossils are the preserved remains of dead organisms. Second languages are alive. These two lines summarize the gist of one of Mike's points in his thoughtful critique of the fossilization concept and related research in his contribution to the *Handbook of second language acquisition* (Long, 2003). As a second language is alive, it is impossible to demonstrate that it has stopped developing for good – or in the words of Mike Long (2003, 490) "... that something in a person's make-up is 'permanent' is unfalsifiable during her or his lifetime ...".

When Mike's paper on (or better: against) fossilization appeared in 2003, the notion he debated had been around for over thirty years. During this time, fossilization had had two meanings. It had been viewed by some researchers as a PROCESS leading to a deadlock in second language acquisition despite ample learning opportunities while other researchers used the term to refer to the RESULT of a stalled acquisition process. Stabilization was seen as a precursor to fossilization. In 1972 Selinker (1972, p. 215) had stated the following on the issue: "Fossilizable linguistic phenomena are linguistic items, rules, and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what the age of the learner or amount of explanation and instruction he receives in the TL." As this quotation illustrates and as Mike demonstrated in Long (2003), Selinker's concept defies a testable definition. However, we are now getting ahead of our story that began about 20 years before the publication of Mike's paper. We hope the readers of this chapter will permit a personal account by the first author (MP) of how he encountered Mike's line of argument on fossilization and variation over a period of several years, before we set out our novel contribution to research on stabilization. This retrospective view will allow us to sketch out relevant lines of reasoning in their historical context.

It was in the mid 1980s on one of MP's trips from his home in Sydney to the US on the then standard route via Auckland, Pago Pago and Hawaii that he

1. Retrieved on 15 August 2021 from <https://www.digitalatlasofancientlife.org/learn/nature-fossil-record/the-process-of-fossilization/>

stopped over in Hawaii to discuss SLA research with Mike and his colleagues at the Department of Second Language Studies. Mike had just started collecting data from Ayako, a Japanese woman who had been acquiring ESL naturalistically for 37 years with both motivation and opportunity to learn English. Long (2003) argued that given her “pervasive and persistent errors despite ample opportunity to acquire the target language ...” (p. 509) she initially appeared “... a perfect candidate for a fossilization claim ...” (p. 510).

Mike was intrigued by the impressive gap between Ayako’s decade-long positive learning opportunities and her apparent lack of success. On the same occasion he also demonstrated his recorded evidence of a similar phenomenon with an international visiting scholar from Northern Europe who, in his mental lexicon, had annotated the lexical entry ‘chicken’ for plural (formally similar to ‘ox-en’ or ‘child-r-en’ and functionally similar to ‘glasses’). As a result, this visiting scholar made reference to “two chicken”, “all his chicken” and so on (but never ‘one chicken’). This occurred despite his formal training in linguistics and the massive presence of the modern English form ‘chickens’ in his everyday life. However, unlike Ayako, the visiting scholar instantly changed his lexical entry for this item once he realized his mistake.

What followed were discussions about the role of grammatical awareness and the nature and stability of interlanguage systems. Both Mike and MP were interested in learner variation and the way it relates to the overall system of the developing interlanguage. In the context of the fossilization debate, Mike demanded a testable definition of the notion proposed in Selinker (1972) and found permanence (of linguistic features) to be the only component of the definition that could be tested without ambiguity. Mike also noted that the unit of analysis is not defined for the purpose of empirically testing a linguistic sample for fossilization: “What is *the appropriate unit of analysis* [emphasis in the original]: the whole IL, the module, the linguistic rule, particular forms, words, meanings, collocations, form-function-relationships, ranges of variation, all of these, or something else?” (Long, 2003, p. 491).

It is easy to see that an abundance of variable interlanguage forms would contradict the permanence criterion of Selinker’s original definition of fossilization, and in his 2003 paper Mike did indeed conclude that Ayako – despite her demonstrated lack of learning – displayed far too much variation for her interlanguage to meet the permanence criterion and thus to be fossilized: “For example, plural-s marking, which varies across tasks and time, was supplied correctly only 71 percent of the time in obligatory contexts in free conversation in 1985, and 48 percent of the time in obligatory contexts in free conversation in 1995” (Long, 2003, p. 509).

It turned out – on this occasion and during later visits by Mike to the Language Acquisition Research Centre (LARC) in Sydney that was directed by MP – that despite the shared scepticism about the notion of fossilization and their shared interest in variation, Mike and MP looked at these phenomena from different vantage

points. Mike tested the permanence of linguistic forms in terms of the rate of suppli-
 ance in obligatory contexts (cf. Long, 2003, p. 509) and contrasted variability with
 stability, asserting that "... proof of stability (within discourse domain X, context
 Y, and over time period Z) [is present, for instance,] in a learner's failure to use the
 regular past tense morpheme *-ed* in English appropriately on any verbs ..." (Long,
 2003, p. 491).

In contrast, MP brought to bear a multidimensional perspective on second
 language acquisition developed by Meisel, Clahsen & Pienemann (1981) that is
 based on a dynamic view of language development as espoused in DeCamp's (1973)
 implicational scales analysis, Bickerton's (1971) approach to pidgins and creoles,
 Labov's (1972) variationist linguistics and Bailey's (1973) wave model. The mul-
 tidimensional model of SLA (Meisel, Clahsen & Pienemann, 1981) differentiated
 between a variational and a developmental dimension and utilized descriptive tech-
 niques to identify variation and development in naturalistic SLA corpora. Because
 of the generally low level of accuracy in these corpora, the multidimensional model
 employed an emergence criterion to determine the acquisition of an L2 form based
 on a minimal occurrence of the given form in a defined context.

In the context of the later-developed processability theory and its applicability
 to typologically different languages, the syntax-based emergence criterion of the
 multidimensional model was extended to morphology together with its associated
 requirement for a systematic test of lexical and morphological variability in order to
 exclude the use of formulae (Pienemann, 1998, p. 144ff). Applying the emergence
 criterion to a sample taken from one given learner performing a given task made
 it possible to attribute to this sample a developmental status independent of the
 accuracy of L2 forms contained in the sample. The level of accuracy was thus able
 to be analyzed separately as one aspect of the variational dimension.

Mike was very supportive of the multidimensional model, in particular its
 explanation of development in terms of processing strategies – as evident in his
 2003 paper and in his contribution to the 1991 volume *An introduction to second
 language acquisition research* (Larsen-Freeman & Long, 1991). However, he was
 not convinced that the model contained a sufficient theoretical motivation for the
 variational dimension (cf. Larsen-Freeman & Long, 1991). This evaluation probably
 prevented him from applying the descriptive approach contained in the model to
 his analysis of L2 data.

However, Mike's critique of the multidimensional model formed part of MP's
 motivation to develop the multidimensional approach into one theoretically co-
 herent approach to SLA, addressing development and variation from the same
 psycholinguistic base that applies to all L2s. As a result, MP created processability
 theory (PT) (Pienemann, 1998, 2005). He developed a formal, processing-based
 theory that defines the space within which both development and variation occur,

and he operationalized the emergence criterion for the acquisition of morphology and syntax, applied it to ESL and a host of other second languages and connected it to the centerpiece of PT, the developmental formalism that applies to all second languages.

It may be useful for the purpose of this chapter to briefly summarize the emergence criterion in its PT format. In syntax, the first step in applying the emergence criterion is to specify the syntactic context for the structure (such as Wh-inversion, as in *Where is John?*) that the analyst wants to focus on. This specification may refer to the obligatory context of a syntactic rule of the target language. Alternatively, the analyst may choose to specify non-target linguistic contexts (e.g., WH subj cop X, as in *Where John is?*) for the regularities of the learner language. Any sample of a sufficient size (cf. Pienemann, 1998, Chapter 4) is then scanned for these syntactic contexts, and the analyst (or the software) counts the applications and the non-applications of the rule and calculates the proportion of applications related to the number of contexts. The proportion is reported as a fraction of 1 that can also be converted into a percentage value. The emergence criterion is met when three lexically different rule applications are present in the sample in the same syntactic context. The percentage of rule applications serves to further describe the given structural feature of the learner language. For morphology, an additional criterion must be met. The rule applications must apply to different morphemes from the same paradigm in order to exclude formulae. For instance, the emergence criterion is met for third person -s, if the morpheme appears on three different lexical items in the same sample and if these items also appear with different morphemes from the same paradigm (for example, either 0 or in other cases -ed). From the view of data gathering methodology, this means that samples must be sufficiently rich to permit testing these conditions of the emergence criterion. For a discussion of differences in the exact applications of the emergence criterion see Lenzing, Nicholas, & Roos (2019) and Nicholas, Lenzing, & Roos (2019).

This descriptive apparatus yields highly refined analyses of L2 variation and development when applied to interlanguage corpora. MP's 1998 monograph (Pienemann, 1998) contains a study on ESL interlanguage variation that was inspired by a series of lectures on the role of tasks in language teaching and in language acquisition research delivered in the 1990s by Mike Long at the Language Acquisition Research Centre. In his lectures Mike set out how tasks can be designed to define the natural communicative context for the use of specific target language forms, thus naturally triggering the use of these forms. In the LARC study on variation that was directed by MP, extensive samples of interlanguage speech were collected from six young adult ESL learners in six different tasks covering a total of 720 minutes of recorded speech (= 2,000 turns). All data were collected on the same day. The data were transcribed in full and analysed according to the ESL

developmental patterns established by Pienemann, Johnston, and Brindley (1988) and refined by Pienemann (1998). The aim of the exercise was to see how much the IL of any one learner would vary across the different tasks in relation to (a) stage of acquisition and (b) correctness in obligatory contexts for all of the 16 TL structures contained in the ESL developmental matrix.

The results of the study display the full impact of a multidimensional view of second language acquisition. Looking only at the rates of suppliance in obligatory contexts, most learners had highly variable rates in different tasks for most of the 16 TL structures. In other words, such a one-dimensional view of accuracy suggests maximal IL variation. This view focuses on the distance of the sample from target language correctness. This TL-centered view contrasts with a view from the opposite end, where the focus is the emergence of linguistic forms and regularities in the learner's interlanguage. When the emergence criterion was applied to all 36 samples of the LARC study, it became obvious that the interlanguage of each of the six learners was located on the same stage in the developmental matrix in all six tasks. In other words, the apparent task variability that was a feature of the design of the tasks did not affect the steadiness of the interlanguage system of the learners in terms of language development. MP was able to demonstrate that the variability found in the data in relation to TL-correctness was attributable mainly to the effect of the tasks on lexical choice and the spread of morpho-syntactic development along lexical lines.

As a result, MP formulated the 'steadiness hypothesis' according to which learners' interlanguages will be steady developmentally across different tasks carried out at the same point in time. In other words, the descriptive and explanatory apparatus of PT can reveal and explain the steadiness of interlanguage systems in terms of development while simultaneously accounting for the system's variability in other aspects of the developing grammatical system. And this ability forms the basis of the account of stabilization entailed in PT. In this account, stabilization is understood as a lack of development (measured using the emergence criterion) within structural domain X, in context Y, and over time period Z.²

Before we move on to a more detailed PT account of stabilization, we remind readers that we (i.e., MP and the co-authors of this paper) follow the logic of Mike Long's rejection of fossilization as an operationalizable concept in SLA. As pointed

2. The LARC study and many other SLA studies cited in Pienemann (2015) empirically invalidate two claims by Larsen-Freeman (2006) used to support an argument that SLA is 'chaotic': "There are no discrete stages in which learners' performance is invariant". (Larsen-Freeman, 2006, p. 592) and "Learners do not progress through stages of development in a consistent manner. There is a great deal of variation at one time in learners' performances and clear instability over time". (Larsen-Freeman, 2006, p. 592)

out above, Long (2003) found that the only testable criterion for fossilization was permanence and that it failed that empirical test. It is for this reason that Long (2003) suggested focusing on the study of stabilization in interlanguages instead.

We also follow Long (2003) in his assessment of the inadequacy of explanations of fossilization through factors such as lack of negative feedback, provision of positive feedback, unavailability of or insensitivity to negative feedback, lack of access to UG or L1 transfer. Long (2003, p. 514) boldly states “[w]hat almost all explanations on offer have in common is that they do not work ...”. One problem is that they lack predictive power. Long (2003, p. 514) argues that “[w]hat is sought of explanations is predictive power, not an ability after the fact to describe cases where the proposed causal factors (supposedly) did work, while ignoring those where they did not.” Consistent with this argument, we will focus on the predictive power of processes inherent in the dynamics of language development.

By doing this we will follow a tradition that started in the 1970s in the dynamic accounts of language acquisition and language change referred to earlier and the application of this work to SLA in the multidimensional model (Clahsen, Meisel, & Pienemann, 1981). This tradition continued in the extension of this approach to morphology and typology and its incorporation into a formal linguistic framework in Pienemann’s (1998) processability theory, and it has recently been extended by Lenzing, Pienemann, and Nicholas (in press) to the study of dynamical systems as studied in mathematics and physics.

Let us add an early cautionary note here to avoid possible misunderstandings. The way Lenzing, Pienemann and Nicholas (in press) and the authors of the present chapter conceptualize the application of dynamical systems theory to SLA is fundamentally different from the approach to dynamical systems theory (DST) promoted by Larsen-Freeman (2006) and de Bot and colleagues (e.g., de Bot, Lowie, & Verspoor, 2007) in what has been referred to as “the dynamic approach”. Larsen-Freeman’s and de Bot et al.’s approach is based on metaphors inspired by DST. In contrast, we apply the core of DST, its key mathematical concepts, to SLA processes. However, before we spell out the details of this latest extension of the tradition of dynamic studies of SLA, we will outline how stabilization is conceptualized in PT.

2. Variability, steadiness and stabilization as conceptualized in PT: The ‘wrong track pathway’

Processability theory is based on a hierarchy of processing resources that constitutes the backbone of L2 development. This hierarchy is interpreted through a formal theory of language, LFG (Bresnan, 2001), that has a processing correlate

(incremental procedural grammar (Kempen & Hoenkamp, 1987)). When the hierarchy of processing resources is interpreted through LFG, there is a degree of leeway at every level of the hierarchy in how different grammatical forms conform to the constraints of the given level. This gives rise to variable L2 forms, i.e., to IL variation. For instance, at level 2 of the hierarchy, semantic roles (e.g., ‘agent’, ‘patient’ etc.) are mapped directly onto constituents,³ resulting in canonical word order – and reflecting constraints preventing the learner’s production of any other order of constituents. The simplest and developmentally earliest form of this type of sentence is the equational sentence (e.g., ‘Jane is tall.’). In other words, the processing constraints active at this level force the production of elements in canonical word order. The freedom that remains for the learner is whether or not to produce specific elements such as the semantically redundant copula (i.e., either ‘Jane is tall’ or ‘Jane tall.’) This kind of leeway for the learner is present at every level of the hierarchy, and it is formally described in the way the processing hierarchy is spelt out using LFG formalisms. In this way, development and variation are accounted for by one and the same psycholinguistic framework, as had been demanded by Mike Long. (An up-to-date summary of these core concepts of PT can be found in Pienemann and Lenzing, 2020).

For the purpose of this chapter we will focus on the developmental dynamics of equational sentences. Other developmental trajectories and their variational components that are covered by PT are described in detail in Pienemann (1998) and are summarized in Pienemann and Lenzing (2020).

The developmental schedule of English equational sentences has been described as set out in Table 1 by Johnston (1985 quoted in Pienemann, 1998, p. 177f.) and Pienemann (1998, p. 181):

Table 1. Developmental schedule of equational sentences*

	Structure	Description	Example
1.	N/A (?)	– single constituent:	‘big’/ ‘big?’
2.	N/pro (cop) N/A (?)	– equational S	‘he big’/‘he is big’/‘he big?’
3.	WH pro/N (cop) N/A?	– WH-fronting	‘Where he is?’
4.	Cop pro/N + A?	– copula fronting	‘is he big?’ ‘is she the boss?’
5.	WH cop pro/N/A?	– copula inversion	‘Where is he?’
6.	S whether/Wh pro/N cop N/A	– cancel copula inv.	‘I wonder where he is.’ ‘I wonder whether he is happy.’

* We thank Bronwen Dyson for the observation that there is a need for further refinement of aspects of the relationship between stages 4 and 5. This issue does not materially affect the argument that we are presenting here.

3. The processes that cause canonical word order are formally explained based on LFG in Pienemann (2011).

In Table 1, question marks in brackets indicate that the same word order string can be used with either declarative or interrogative meaning. A question mark without brackets indicates a specific interrogative meaning. In that table, the variational and the developmental dimension have not been separated. Once the two dimensions are separated, a different picture emerges. As is shown in Table 2, the range of options for exploiting the processable leeway varies between two paths that encompass different degrees of presence for the copula, at the extremes, fully supplied (Path 1) or fully omitted (Path 2). This variable space has other consequences that will be detailed below.

Table 2. Developmental schedule of English equational sentences differentiated by variational type

	Structure	Path 1: +copula	Path 2: -copula
1.	N/A (?)	- 'big' / 'big?'	- 'big' / 'big?'
2.	N/A/pro (cop) N/A (?)	- 'he is big'	- 'he big?'
3.	WH pro/N (cop) N/A?	- 'Where he is?'	- 'Where he?'
4.	Cop pro/N + A?	- 'is he big?'	-
5.	WH cop pro/N/A?	- 'Where is he?'	-
6.	S whether/Wh pro/N cop N/A	- 'I wonder where he is.'	-

Table 2 reveals a striking difference between the developmental trajectories of the two variational types (+ and -copula): the variant without the copula does not develop beyond level 3. The reason for the difference is simple. At levels 4, 5 and 6, the copula is fronted (level 4), inverted sentence-internally (level 5) or its inversion is suppressed (level 6). If, in the extreme case, the interlanguage contains no copula (Path 2), none of the later operations can be performed. The outcome of successive variational choices (with or without copula) therefore has repercussions for subsequent development. This is an example of the definition of stabilization provided in Section 1; the Path 2 system stabilizes at level 3. Pienemann (1998) noted that these options for exploiting the leeway occur at every level of the developmental hierarchy and that they accumulate as the learner moves up the hierarchy. Every level of the hierarchy relates to a 'developmental problem' that has to be solved by the learner using the processing resources that are available to her or him at that point. For instance, at level 3 of the hierarchy, ESL learners are able to front wh-words. In the target language (English), the resulting structure (i.e., sentences with a fronted wh-word) requires copula inversion (i.e., 'Where is he?'). However, at this stage learners are not yet able to produce copula inversion. Instead, they have two quite different solutions at their disposal: (1) to continue to use the copula as part of the structure that follows the wh-word BUT in canonical order, thus sticking to the form acquired at level 2 (i.e., 'Where he is?' with pro + cop in canonical order) or

(2) to leave out the copula, thus eliminating the ‘offending’ element that cannot be handled at this stage (i.e., ‘Where he?’) – an approach that simplifies the developmental challenge. We regard this second option as ‘simplifying’ because it removes the challenging feature from consideration. Option 1 also ‘reduces’ the extent of the challenge but does not remove the cause of the challenge from consideration.

This range of options is what Pienemann (1998) refers to when he shows that learners have a general tendency to prefer either variational options that ‘recycle’ structures acquired earlier (option 1 above) or to simplify the problem-structure further (option 2 above). Both solutions result in non-target like forms, but their implications for subsequent options can be profoundly different. The first option allows for further structural development. In contrast, the second option may prevent further development and as a result, in learners who accumulate multiple instances of this option, development tends to stall. Pienemann (1998) argued that learners are consistent in the types of variational choices they make. This claim is supported, for instance, by findings from the ZISA⁴ study referred to earlier (Meisel, Clahsen, & Pienemann, 1981; Clahsen, Meisel, & Pienemann, 1983), a cross-sectional study of German L2 acquisition by adult migrant workers based on extensive informal conversations with 45 informants. In their study, Meisel and his co-workers identified 14 variational features and found that the informants were consistent in the degree of simplification across all 14 variational features (cf. Clahsen, Meisel, & Pienemann, 1983, p. 249ff.) – resulting in clusters of variational features that align with the trends characterising the differences between individual learners. Further evidence of the systematicity of interlanguage variation in a set of longitudinal studies of ESL acquisition was recently added by Dyson (2021). In Dyson (2021), the eight adolescent learners were shown to demonstrate consistency in their use of the simplifying features.

Table 3 gives a rough overview of the distribution of nine of the fourteen variational features that are easy to explain in the context of this chapter. These nine features are included in Table 3 that focuses on the fourteen most highly simplifying learners in the study. A ‘+’ marking in the table indicates that there is evidence of the presence of the features. The learners are arranged from top to bottom in the sequence from fewest features present to most features present. There is an obvious cluster of plus-markings for features 1–9 in Table 3. In fact, considering the arrangement of learners from Rosa I (at the bottom) up to Pedro I, Table 3 conforms closely to the requirements of an implicational table – with a coefficient of scalability of 0.93 (cf. Hatch & Lazaraton, 1991, p. 210). This finding implies that the variational features of these learners not only follow the same clustering but they are related

4. ZISA = Zweitspracherwerb italienischer und spanischer Arbeiter (Second language acquisition of Italian and Spanish workers). The ZISA study was directed by Jürgen M. Meisel

Table 3. Simplifying features in the ZISA study*

	1	2	3	4	5	6	7	8	9
Pedro I.	+	-	+	+	-	-	-	-	-
Toni I.	+	+	+	+	-	-	-	-	-
Lolita S.	+	+	-	+	+	+	-	-	-
Franco I.	+	-	-	+	+	+	-	-	-
Antonio S.	+	(+)	0	-	+	0	-	-	-
Tivo S.	+	-	+	+	+	+	-	-	-
Jose S.	+	(+)	+	+	+	+	-	-	-
Zita P.	-	+	0	+	0	0	-	-	-
Francisco S.	+	+	0	+	+	0	-	-	-
Pascua S.	+	+	0	+	+	0	-	-	-
Eliseo I.	+	+	+	+	+	+	-	-	-
Pasquale I.	-	+	0	+	+	0	-	-	-
Pepita S.	+	+	0	+	+	0	+	-	-
Manuel S.	+	+	+	+	+	+	+	-	-
Dolores S.	+	+	+	-	+	+	+	-	-
Maria P.	-	+	+	+	+	-	+	+	-
Manuel P.	+	+	0	+	(+)	0	+	(+)	-
Luis P.	+	+	+	+	+	+	+	+	+
Rosa I.	+	+	0	+	+	0	+	+	+

Legend for simplifying features

- 1 Omission of subject pronouns
 - 2 Omission of lexical verb in semantically redundant contexts
 - 3 Omission of lexical verb in syntactically complex contexts (stage 4 and above)
 - 4 Omission of verbs overall
 - 5 Omission of the copula
 - 6 Omission of subject pronoun in syntactically complex contexts (stage 3 and above)
 - 7 Omission of the article
 - 8 Omission of prepositions (in semantically redundant contexts)
 - 9 Omission of prep + N (in semantically redundant contexts)
- + = simplifying feature is present
 (+) = simplifying feature is present but in fewer than 5 contexts
 - = simplifying feature is absent
 0 = no context for this simplifying feature

* Table 3 of this chapter is based on Table 23 in Clahsen, Meisel, & Pienemann (1983, 250). In the original (Table 23), cells with no marking were left blank and, formally, should have been marked '-'. In Table 3 the corresponding cells (to the right of all '+'-markings) are marked '-'. The coefficient of scalability has been calculated on the basis of the cells marked '+' or '-'.

implicationally with Rosa I and Luis P at one end of the continuum (most simplifying features included) and Pedro at the other end, someone whose interlanguage includes more than one of the features but is closest to the other learners whose interlanguages include none of the simplifying features.

Remembering that this is not a developmental sequence, this analysis shows that learner variation is not random. Instead, learners vary in accordance with a systematic pattern, and learner languages can be related to relatively coherent variational styles that are placed on a linguistic continuum. At the same time, the fact that Table 3 shows a continuum of learner variation also indicates that the two paths shown in Table 2 represent only the two end regions of the continuum in authentic data and that there are also further alternatives between them. This continuum was also borne out in Dyson's (2021) analysis that documented variation in simplifying options.

According to Pienemann (1998), the complex interaction between variational options and development is part of the dynamics found in second language acquisition. In cases where a consistent trend for a learner is to prefer simplified solutions to developmental problems, an accumulation of such impeding selections ultimately stalls the acquisition process. Pienemann (1998) referred to this preference among learners as the 'bad choice hypothesis'. For the purpose of the present paper we replace the term 'bad choice hypothesis' with the term 'wrong track pathway' in order to avoid any inference of a conscious choice made by the learner, since our argument is that the variational option (pathway) is selected through the dynamics of the acquisition process.

Obviously, the 'wrong track pathway' implies a prediction about L2 development that is reminiscent of the fossilization/stabilization hypothesis, namely the prediction that under certain circumstances L2 development may stall. One of the key differences between the 'wrong track pathway' and the ways in which the fossilization/stabilization hypothesis has been dominantly interpreted is its predictive rather than descriptive character. A second difference is the assumed set of circumstances that determines the stalemate. Overviews of research on fossilization/stabilization such as Han (2004) and Long (2003) quote studies dealing with factors such as corrective feedback, input, universal grammar, learning inhibition, automatization, comprehension, processing, emotional states, focus on form, avoidance, communicative needs, and identity. All of these factors are external to the linguistic dynamics of L2 development. In reference to these factors, Long (2003, p. 512) concludes that "[w]hat almost all explanations on offer have in common is that they do not work – at least not for fossilization ...". And he notes (*ibid.* p. 512): "Surprisingly, no one seems to have considered the possibility that if fossilization is, as Selinker (1972) claimed, a cognitive mechanism producing the non-target-like end-state, also called 'fossilization', there is no need for other explanations ...". It so happens that the 'wrong track pathway' is based in such a cognitive mechanism, and – contrary to the explanations listed above – our claim is that the enacting mechanism is inherent in the dynamics of second language development rather than relying on other influences.

Figure 1 (after Pienemann, 1998, p. 317) affords a graphic illustration of the context of the ‘wrong track pathway’. The points of emergence of stages of the learner language are represented by circles. The levels of the processing hierarchy are represented by horizontal lines. The lines between circles represent the outer bounds of learners’ transition pathways from one stage to the next. Every level of the hierarchy relates to a different developmental problem. In the state at the top, the learner confronts options ranging from the pathway on the left side of the figure to the pathway on the right side of the figure. He or she cannot see the end of the sequence that follows. The line on the right ends at level 3. The sequence on the left continues to at least level 5. In other words, the option followed by the learner at the top determines the subsequent range of options available within the learner’s developmental path, which may (unknowingly) terminate in a dead-end.

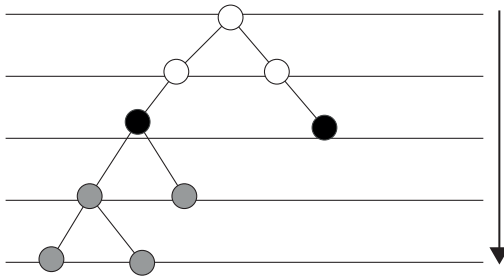


Figure 1. The ‘wrong track pathway’ illustrated as a decision tree

The ‘wrong track pathway’ is conceptually related to Wimsatt’s (1986) notion of ‘generative⁵ entrenchment’ and the developmental lock model. Wimsatt (1986) demonstrated mathematically – with recourse to work of the Nobel laureate Herbert Simon (1962) – that in developmental processes of increasing complexity such as in evolution and embryonic development (Rasmussen, 1987), the preservation of structural features is computationally far less costly than restructuring the system at every level of development.

Rasmussen (1987) presented empirical evidence from the embryonic development of the drosophila that supports Wimsatt’s developmental lock model. One significant consequence of the implications of earlier selections of variational options for the downstream addition of structures is that any unfortunate option that is taken early has far greater repercussions in subsequent development than options that occur later. This relationship is evident in embryonic development in

5. Generative in this context does not carry with it the implications associated with Universal Grammar.

all species. In Wimsatt's model, developmentally earlier structures are understood to be more deeply generatively entrenched than later structures.

The aspect of computational saving inherent in generative entrenchment is captured by Wimsatt's (1986, 1991) 'developmental lock' metaphor, which is based on Herbert Simon's (1962) classic paper, "The architecture of complexity". In that paper, Simon demonstrated that solutions to complex problems can be found more effectively when sub-problems are solved independently, and the solutions to sub-problems are strung together to produce the solution to the overall problem.

Wimsatt's developmental lock is an idealized set of complex problems. The lock consists of ten wheels with ten positions each, very much like an extended version of a combination lock. The total number of possible combinations on this lock is 10^{10} , which, according to Wimsatt, requires 5×10^9 trials to find the correct combination. In this form of unconstrained hypothesis testing, the lock is referred to as a 'complex lock'. In the developmental lock, Wimsatt constrains hypothesis testing by allowing the problem solver to separately address different parts of the combination problem. Rather than having to get all 10 digits right before the combination can be subjected to an empirical test, each wheel can be tested individually in a left-right sequence. The computational advantage of factorising the complex problem in this way is remarkable: Only 50 trials are necessary to find the solution to the developmental lock problem, providing a strict left-right sequence is followed, and each decision is 'locked in'. In other words, in this metaphor, later decisions depend on earlier decisions. If an error is made earlier on it will be very costly to recover from it, because any backtracking will mean that all intervening solutions will be lost.

In the context of the 'wrong track pathway', generative entrenchment is understood as an aspect of the dynamics of developing systems. The key explanatory point that can be derived from the concept of generative entrenchment for language acquisition is that not all structural decisions have to be revised in the developmental process every time a structural change occurs. Initial structural features relevant to one specific feature have the capacity to be locked in and propagate in the developing system and thus determine the ultimate structure without being revisited again and again. In other words, there is a benefit for the learner in that a computational saving is made by laying structures down and keeping them. The alternative would be a developing system in which all processes of structural refinement have to be orchestrated globally for every structural refinement.

The basic thesis inherent in what Pienemann (1988, p. 326 ff.) previously referred to as the 'bad choice hypothesis' – as an integral part of PT – is that different outcomes and developmental paths in language development are due to different developmental dynamics. These different developmental dynamics are caused by

different variational options being selected when learners face developmental problems. This means that the process of development can be fundamentally similar in all learners, with respect to language processing, despite learners having fundamentally different outcomes and different developmental paths. In other words, what we introduce here as the ‘wrong track pathway’ affords an explanation of how learner languages can stabilize, and the descriptive framework of PT serves to operationalize the implications of the options taken and their implicit predictions for stabilization.

The ‘bad choice hypothesis’ did not offer an account of the dynamics that drive the selection of either expanding or limiting variational options. In the following sections we will present an approach that does just that. This approach was inspired by Mike Long’s (2003, p. 512) observation, that “... no one seems to have considered the possibility that if fossilization is ... a cognitive mechanism producing the non-target end-state also called ‘fossilization’, there is no need for other explanations ...”. As implied by Long, we focus on the mechanisms producing the non-target end-state. We hypothesise that a non-target end-state can be brought about by the dynamics of these mechanisms without reference to influences such as universal grammar, learning inhibition, automatization etc. and without reference of any central control. The mechanisms we focus on are cognitive and interactional in nature, and they are very basic. The dynamics of such basic mechanisms can be operationalized in DST (see Mitchell, 2009; Feldman, 2019), because such dynamical systems can model internal growth and decline patterns partly depending on the initial condition but without any central control where tipping points can be reached that are followed by radical changes in the state of the system, including a total cessation of development. As Long (2003) rightly insisted on operationalized concepts that can be empirically tested, our simulation is NOT based on the post-modern DST metaphors that are currently popular among some applied linguists (e.g., Larsen-Freeman, 2006). Instead, our simulation model is fully operationalized and tested against longitudinal empirical data.

Our approach utilizes an agent-based model (ABM) designed by us as a method of calculating and simulating processes within dynamical linguistic systems, because agent based modeling is an effective approach to implementing dynamical systems that are driven by the interaction of very basic processes without central control. Our ABM focuses on the simplification of equational sentences, thus continuing the focus on early L2 systems that we described above. The purpose of this simulation is to formally model the dynamics that determine the selection of variational options in relation to one feature. If simulations such as this can be validated in a comparison with real longitudinal data, they may in the future also serve to explore further aspects of the dynamics of SLA.

However, before we move on to our L2 simulation research and its validation we will outline some of the basics of complex dynamical systems theory on which our agent-based model is based. The reader will notice that our outline of complex dynamical systems theory also includes some basic (high school) mathematics. At this point we acknowledge what Mitchell (2009) had to say about the inclusion of mathematics in her acclaimed general introduction to complexity:

Authors of popular-audience science books are always warned of the following rule: every equation in your book will cut the readership by one-half. I'm no exception – my editor told me this fact very clearly. I'm going to give the logistic map equation here anyway, so half of you who would throw the book out of the window if you ever encountered an equation, please skip over the next line [containing the logistic map equation]. (p. 28, fn. 1)

If this warning also applies to linguists, we suggest following Mitchell's advice and focussing on the key concepts relating to complex dynamical systems that arise from a simple equation rather than trying to reconstruct the way in which they are generated by the equation.

3. Dynamics in SLA

In order to outline the theoretical assumptions underpinning the agent-based model introduced in later sections, we now introduce some key constructs used in the modeling of complex dynamical systems, including sensitivity to initial conditions, tipping points, the logistic map and the notion of chaos. We have elaborated this background in greater detail in Lenzing, Pienemann, & Nicholas (2022) and Pienemann, Lenzing, & Nicholas (forthcoming). These aspects are central to the study of complex dynamical systems.

Both Mitchell (2009, p. 13) and Feldman (2019, p. x) point out that there is no agreement on a definition for the terms 'complex system' and 'complexity'. However, there are a number of common properties shared by complex systems that also apply to our agent-based model. Mitchell (2009, p. 13) summarizes these core characteristics in her definition of the term 'complex system', as "a system in which large networks of components with no central control and simple rules of operation give rise to complex collective behavior, sophisticated information processing, and adaptation via learning or evolution". A 'complex dynamical system' is a complex system that changes over time. It is possible to describe and predict the changing behavior of such a complex dynamical system using mathematics (Mitchell, 2009, p. 15) or simulation models (cf. Wilensky, 1999). Describing and predicting changing behavior is exactly what we do in our simulations of language development and

stabilization. In this section we will first outline some of the key notions associated with complex dynamical systems in the context of mathematical thinking, and we will then go on to show how those notions can be captured in simulation models. We begin with the idea of a dynamical system.

Broadly speaking, “a dynamical system is any mathematical system that changes in time according to a well specified rule” (Feldman, 2019, p. ix). According to Mitchell (2009, p. 16), “[d]ynamical systems theory describes in general terms the ways in which systems can change, what types of macroscopic behavior are possible, and what kinds of predictions about that behavior can be made.”

Henri Poincaré (1913), who is regarded as the father of modern dynamical systems thinking discovered that some dynamical systems display ‘sensitive dependence on initial conditions’. This means that for certain kinds of dynamical systems, tiny differences in their initial conditions lead to huge differences over time. This means that small uncertainties in the initial measurements lead to huge errors in the long run:

[I]t may happen that slight differences in the initial conditions produce very great differences in the final phenomena; a slight error in the former would make an enormous error in the latter. Prediction becomes impossible [...]

(Poincaré, 1913, p. 398)

According to Strogatz (2015, p. 2), with this discovery, Poincaré was “the first person to glimpse the possibility of *chaos* [bold & italics in original], in which a deterministic system exhibits aperiodic behavior that depends sensitively on the initial conditions, thereby rendering long-term prediction impossible”. In the currently dominant applied linguistic accounts of DST, the term ‘chaos’ has been used in its everyday meaning, suggesting that language acquisition is basically chaotic and thus unpredictable (de Bot, 2016, p. 130). This is not the technical meaning of ‘chaos’ in DST. Before we illustrate the notion of ‘sensitive dependence on initial conditions’ with a concrete example, let us first explicate the notion ‘chaos’ as a **technical term**. Feldman (2012, p. 3) defines the term chaos⁶ as follows:

Chaos is a phenomenon encountered in science and mathematics wherein a deterministic (rule-based) system behaves unpredictably. That is, a system which is governed by fixed, precise rules, nevertheless behaves in a way which is, for all practical purposes, unpredictable in the long run.

Feldman (2012, p. 84) argues that “[t]he mathematical use of the word ‘chaos’ does not align well with its more common usage to indicate lawlessness or the complete absence of order. On the contrary, mathematically chaotic systems are, in a sense, perfectly ordered, despite their apparent randomness.”

6. The term itself was introduced by Li & Yorke (1975).

As we have argued elsewhere (Lenzing, Pienemann, & Nicholas, 2022; Pienemann, Lenzing, & Nicholas, forthcoming), the approach to dynamical systems that has emerged in the natural sciences accommodates large measures of determinism and locates ‘chaotic systems’ as only a very small part of the overall possibilities. Here we briefly outline how this apparent contradiction can hold. We do this by illustrating the specific mathematical conditions under which an otherwise stable system becomes chaotic. The first step in this illustration is to consider what happens when a calculation is repeated in specific intervals and the result of the previous calculation is entered into the next calculation as in the example below:

1st calculation:	$0.5 \times (1-0.5)$	= 0.25
2nd calculation:	$0.25 \times (1-0.25)$	= 0.1875
3rd calculation:	$0.1875 \times (1-0.1875)$	= 0.1523437
4th calculation:	0.1523437 etc	

This is the process that underlies a so-called ‘iterated (repeated) function’. The function underlying the above example is the following: $x_{t+1} = x_t (1 - x_t)$

In the above illustration we make use of a specific type of iterated function, namely the so-called ‘logistic map’ (i.e., a specific mathematical equation). This type of iterated function is often regarded as a prime example of a dynamical system.

The logistic map has been applied in different areas of science, such as biology, economy or the social sciences. For instance, in population biology, the logistic map has been used to describe/model the growth of populations, by taking into account factors such as limitations in the population’s habitat and/or limited food resources. We illustrate the application of the logistic map to population growth with the example of a fictitious population of snow hares. We assume that this population is monitored at discrete points in time, namely once a year, over a specific time period. We use the logistic map below:

$$x_{t+1} = rx_t (1 - x_t)$$

It is easy to see that the logistic map (=logistic function) is very similar to the example we gave above – with only one difference. The logistic map also contains a variable ‘ r ’. Applying this formula (i.e., the logistic map) to the calculations above we would only need to multiply the result of each of the above calculations by the value of ‘ r ’.

When the logistic map is applied to calculate population growth, x_t is defined as the “fraction of carrying capacity”, i.e., the ratio of the current population of a species to the maximum population possible in a specific habitat (Mitchell, 2009, p. 305). The parameter r is the growth factor, taking both the birthrate and the deathrate of the population into account – in our example the result of the interaction of the birthrate and the deathrate of snow hares. For mathematical reasons,

the value of the growth factor r can vary between 0 and 4, and, x can vary between 0 and 1 (see May, 1976 for details). Once you enter specific values for the variable x and r into the formula, the following regularities can be observed. If $r < 1$, the population size will decrease, if $r = 1$, it will remain stable and if $r > 1$, the population will grow (see Feldman, 2019, p. 63). Readers can experiment with the logistic map using a simple and intuitive tool on a website⁷ provided by the Santa Fe Institute. This dynamical system is “deceptively simple” (Feldman, 2019, p. 67), but at the same time it exhibits some striking characteristics related to the mathematical notion of chaos.

Depending on the value of the parameter r , the logistic map displays either stable or unstable behavior in the fraction of the maximum possible population that is achieved over time (therefore the maximum is 1.0). According to Feldman (2019, p. 70), the stable behavior is of two types: fixed points or periodic cycles. Fixed points occur if $1 < r < 3$. Periodic cycles occur if r (the growth rate) is larger than 3 and smaller than ≈ 3.569943 . In cases of fixed points, the fraction of the maximum possible population of snow hares grows and reaches a fixed point that is larger than zero. This is illustrated for the first 40 iterates of the function in Figure 2 where $r = 1.9$ and the two different initial conditions are $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green).

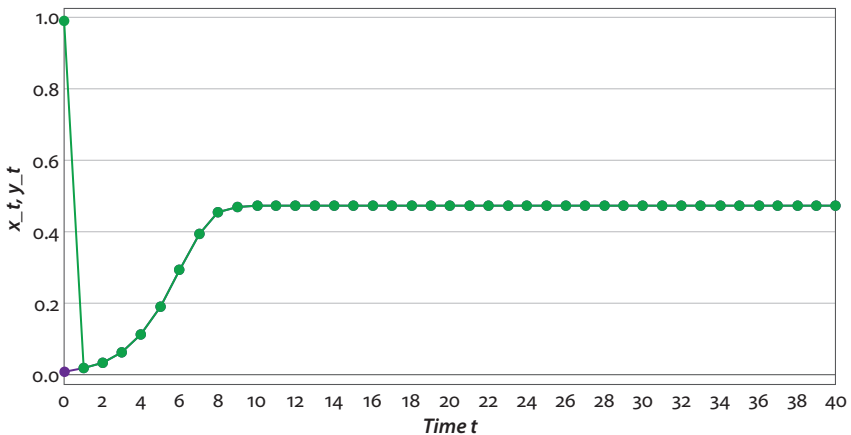


Figure 2. Orbits for the logistic equation with $r = 1.9$ and the two initial conditions $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green) This figure as well as the following ones were created with the program by K. N. Springer. January 2014. © 2014 Santa Fe Institute (ComplexityExplorer.org)

7. https://s3.amazonaws.com/complexityexplorer/DynamicsAndChaos/Programs/time_series.html

In this scenario, the system has an attracting fixed point at ≈ 0.473 . This means that any initial condition where the proportion of the actual population to the maximum possible population is initially greater than 0 and smaller than 1 eventually reaches this “equilibrium value” (Feldman, 2019, p. 68).

The second type of stable behavior of the dynamical system relates to cycles. This is the case for larger r (growth rate) values. If the value of the parameter r is between ≈ 3.0 and ≈ 3.57 , the population also grows, but instead of reaching a stable fixed point, it oscillates between two or more values, depending on the exact value of r . If we consider the case of $r = 3.3$ (Figure 3), we can observe that our snow hare population oscillates between a larger population in one year and a smaller population in the next year. This type of behavior, where the system oscillates between two values, is called a period-2 cycle (Strogatz, 2015, p. 361). Importantly, this behavior is not influenced by the system’s initial condition, as shown in Figure 3 for the two initial conditions specifying the relationship between the actual population and the maximum possible population, $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green). Both initial conditions result in the same later pathway.

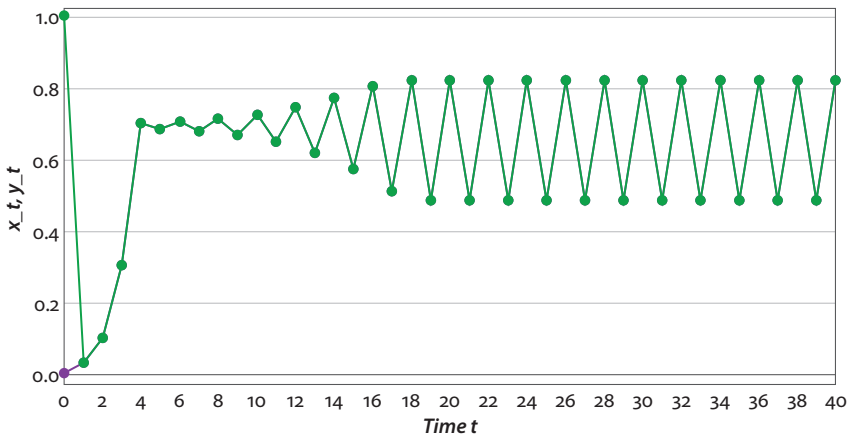


Figure 3. Orbits for the logistic equation with $r = 3.3$ and the two initial conditions $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green)

At larger r values between 3.4 and 3.5, the system undergoes a qualitative change: Again, independent of the system’s initial condition, the period doubles again, and the snow hare population alternates between four different values, i.e., the population size is the same every four years. This period-4 cycle is exemplified with $r = 3.49$ and the two initial conditions $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green) in Figure 4:

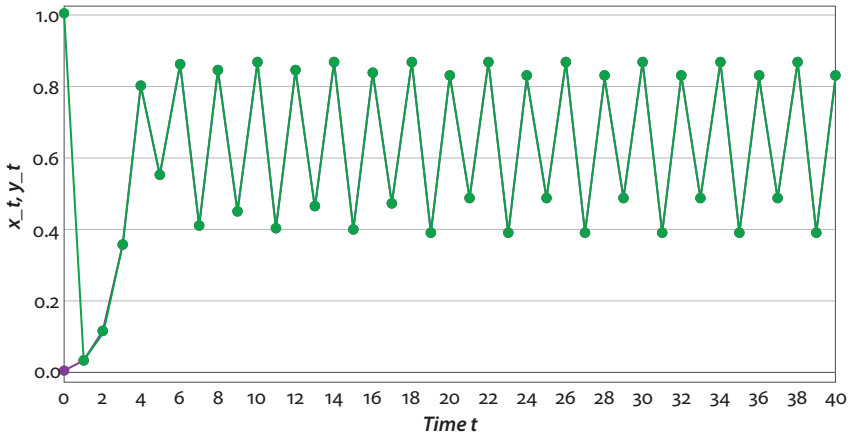


Figure 4. Orbits for the logistic equation with $r = 3.49$ and the two initial conditions $x_0 = 0.01$ (in purple) and $y_0 = 0.99$ (in green) (period-4 cycle)

An increase in r results in further period doublings to cycles of period 8, 16, 32 and so on. A point of period doubling is called a ‘bifurcation’ and it constitutes a tipping point where a small change in an input such as the growth factor dramatically affects the outcome (Scheffer, 2010). Up to (a value just below) the r value of ≈ 3.569946 , the behavior of the logistic map is predictable, as it either reaches a fixed point or oscillates in periodic cycles. However, for values beyond when r approximates 3.569946, the long-term behavior of the system starts to be aperiodic for most r values,⁸ as the sequence no longer settles into fixed points or periodic cycles. Instead, the system displays chaotic behavior, characterized by sensitive dependence on initial conditions, which renders its long-term development unpredictable (see e.g., Mitchell, 2009, pp. 31–32; Strogatz, 2015, pp. 362–363). This chaotic behavior of the logistic map is shown in Figure 5 for $r = 3.9$ and the two quite similar initial conditions $x_0 = 0.2$ (in purple) and $y_0 = 0.2000001$ (in green).

Thus, although the logistic map is a simple equation that is completely deterministic, for certain values of the parameter r , it yields chaotic orbits, as in these cases, the dynamical system is sensitive to initial conditions. This is metaphorically known as the butterfly effect: small differences in initial conditions yield widely diverging trajectories, but unlike the implication of the metaphoric extension, the chaos is limited and its emergence precisely predictable.

8. Although most r values larger than ≈ 3.569946 yield aperiodic orbits, there are also periodic windows for $r > 3.569946$, such that “the orbit diagram reveals an unexpected mixture of order and chaos, with **periodic windows** interspersed between chaotic clouds of dots.” (Strogatz, 2015, p. 363)

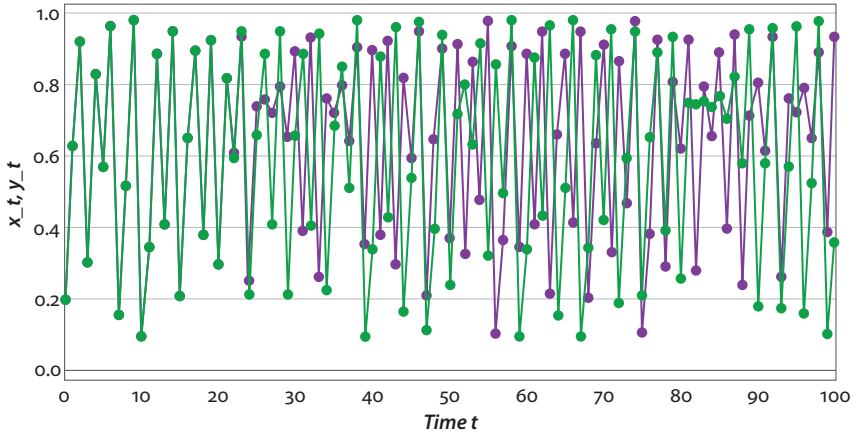


Figure 5. Orbits for the logistic equation with $r = 3.9$ and the two initial conditions $x_0 = 0.2$ (in purple) and $y_0 = 0.2000001$ (in green)

On the one hand, “the presence of chaos in a system implies that perfect prediction *à la* Laplace is impossible not only in practice but also *in principle*, since we can never know x_0 to infinitely many decimal places.” (Mitchell, 2009, p. 33). However, this does not mean that a system with chaotic aspects behaves in a completely random, unpredictable way. Mitchell (2009, p. 34) claims that “unlike colloquial chaos, there turns out to be substantial order in mathematical chaos, in the form of so-called *universal* features that are common to a wide range of chaotic systems [italics in original]”.

As our brief look at a simple mathematical function has shown, the metaphorical use of the term chaos in some popular accounts of DST that focus on the unpredictability of dynamical systems and the complete interconnectedness of all components of a system distract the users of these accounts from the fact that complex dynamical systems are to a large extent predictable and follow well-defined regularities. This insight is essential for our simulation of L2 stabilization, as the mathematical mechanisms underlying DST are used in our simulation to make predictions about developmental trajectories, and these predictions are compared with natural longitudinal data in order to gauge the predictive power of our SLA simulation model.

As we pointed out above, population growth (and decline) – as in our snow hare example – can be modeled using the logistic map that includes a variable for the carrying capacity and a growth factor. The resulting population development can be represented by a curve in a coordinate system that refers to time and population size. However, using mathematical modeling is not the most straight-forward solution for all complex dynamical systems. For instance, if we want to write down an

equation that calculates the movements of a flock of birds, this is nearly impossible. It is hard to incorporate all the interactions into analytical equations. However, the behavior of each individual bird is easy to model. Each bird follows strict rules depending on the other birds and their environment. If we can simulate a number of individual birds in a flock accurately, we can predict the influence of a change of specific conditions of the environment on the movements of the flock of birds. This scenario describes the basic idea of agent-based modeling (ABM). One key component of an ABM is a set of independent ‘agents’ (that could be animals, humans, molecules, words or many other things) that have specific features and that make rule-based ‘choices’ when they act. A second component is that each agent acts independently. The flocking behavior of birds has in fact been implemented in an ABM.⁹ In that ABM, each agent represents a bird and follows the same simple rules that also affect the behavior of the agent’s neighbor. Collectively, the agents display the flocking behavior found in birds, and this occurs without any central control.

Population growth has also been simulated in various ABMs. One example is an ABM that was programmed using NetLogo (Wilensky, 1999), a publicly available agent-based modeling software. This example of an ABM “... explores a simple ecosystem made up of rabbits, grass, and weeds”.¹⁰ This ABM mimics population growth in a way that is similar in effect to the use of the logistic map and the specification of a carrying capacity and a growth factor. But it is driven entirely by the behavior of the individual agents.

ABMs are used in many branches of science, engineering, in the social sciences and in economics to simulate real systems, especially “... when prototyping or experimenting with the real system is expensive or impossible.” (Borshchev and Filippov, 2004, p. 1) When designing ABMs, the modeler does not define the global behavior of the system (dynamics),

[i]nstead, the modeller defines behaviour at [an] individual level, and the global behaviour emerges as a result of many (tens, hundreds, thousands, millions) individuals, each following its own behaviour rules, living together in some environment and communicating with each other and with the environment.

(Borshchev & Filippov, 2004, p. 6)

Agent-based modeling has two distinct advantages for our stabilization project: (1) It permits a more straight-forward translation of the dynamics of language acquisition than mathematical modeling, and (2) it may ultimately permit the modeling of SLA dynamics in cases where data collection in a “...real system is [too] expensive or

9. <http://ccl.northwestern.edu/netlogo/models/Flocking>

10. <http://ccl.northwestern.edu/netlogo/models/RabbitsGrassWeeds>

impossible” (Borshchev & Filippov, 2004, p. 1). As we will show below, ABMs share several features with mathematical modeling, in particular their capacity to model sensitivity to initial conditions, tipping points and non-linearity.

4. Simulating linguistic simplification in a dynamical system using an Agent-Based Model

As outlined in Section 2, we extend the account of the L2¹¹ developmental dynamics inherent in the ‘bad choice hypothesis’ by focussing on the internal dynamics of the simplification process involving the inclusion or exclusion of the copula. In this section we will present an ABM-based simulation of this simplification process (non-supply of the copula) called ‘the wrong track pathway model’ or simply the ‘WTP model’. We implemented a model of linguistic simplification in NetLogo¹² (Wilensky, 1999). The model is informed by influences that we will outline below. It is presented in an interface that both permits the identified controlling conditions to be entered (in the top half of the interface) and shows the results of the simulation (in the lower part) (see Figure 6).

This simulation is based on the idea that this simplification process is not primarily driven by external forces. Instead, the simplification process is assumed to reflect a series of options whose selections are subject to their own dynamics. These dynamics result from the intersection of learner and interlocutor selections in combination with other features of how the learner processes information. We will demonstrate in this section that certain mental factors that limit human information processing cause an amplification of the process of supplying or not including the copula once a tipping point is reached. We will show how the internal dynamics

11. A reviewer raised the issue of: What prevents L1 learners from relying on the dynamics inherent in our model? As many have acknowledged, nearly all healthy L1 learners will eventually become ‘native speakers’ of the TL, only a minority of L2 learners (after a critical age of about 6) will attain native speaker competence in the TL. Many hypotheses have been put forward to explain this phenomenon (cf. Meisel, 2011; Hyltenstam & Abrahamsson, 2003). And several authors have observed fundamental differences in the neurological basis of L1 and L2 processing. It is beyond the scope of this chapter to develop a set of hypotheses that are capable of explaining these differences and to implement them in an extended simulation model. What we can state here is that in the context of the extensive research efforts on differences and similarities in L1 and L2 acquisition, L1 learners have been shown to have an advantage in either the learning mechanism or in access to knowledge or skills, and these factors would need to be included in a simulation of the dynamical system of L1 acquisition that would prevent the vast majority of L1 learners (except pathological cases) from stabilizing.

12. Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. <http://ccl.northwestern.edu/netlogo/>

of this process are able to tip the selection of an option in relation to the impending developmental problem in one direction or another.

The world of the WTP model that we are exploring here is necessarily reduced. It focuses on a single developmental step relating to level 2 of the PT hierarchy and, within this stage, it focuses on equational sentences. In the WTP model, words are agents as defined in agent-based modeling, which means that they are identifiable entities existing as part of a dynamic network. The dynamic process that they are involved in assigns them the features +cop(ula) and -cop(ula). This assignment occurs as the words are produced in sentences containing two words. In every iteration, a sentence (i.e., an initiation by the learner or a response by the interlocutor) is produced by forming a link between the two words selected for the sentence. The link between the two words in the same sentence is that each of the words in that sentence is marked in the same way, either as including '+' or not including '-' cop.

In the interface, the number of words contained in every simulation run using the model can be set by the user (between 10 and 10,000). Every word has a unique identity. As indicated above, a sentence can be either an initiation or a response. In each iteration a sentence is formed by linking two words selected on a random basis. After every iteration, the marking choice (+ or -cop) made for the sentence in that iteration is stored as a '-' or '+' value attached to each of the two words that have been used. When a sentence that is used in a subsequent iteration of the simulation forms a link with a particular word,¹³ the established value of that word (number of plusses and/or minuses that that word has accumulated in its previous uses) biases the sentence-level choice to be made for +/- cop for that sentence. As a result, any choice (+/- cop) made for the sentence (and hence the words it links) in any one iteration further biases the value assigned in the selection made in a learner's subsequent iteration.

The user of the interface can control the number of iterations that will be carried out. The patterns in the interactions between the learner and the interlocutor influence the calculation of the simplification index. The simplification index is updated after every iteration. The index is based on the average of the value of the choices between +/- cop for all items in the learner's lexicon. The index is calculated as a fraction of 1 such that 1 reflects 100% simplification and 0 reflects no simplification at all. The calculation is carried out iteratively after each of the runs of the simulation (i.e., separately for each of the iterations of the sequence of values) and it is plotted as an orbit with the number of iterations on the x-axis and the values of the simplification index on the y-axis. The orbit is displayed in a window at the bottom of the interface.

13. Words are selected randomly.

When running a simulation, the user can set the initial simplification rate by setting a value between 0 and 1 on a slider positioned under the total and initial lexicon size windows in the upper part of the interface. The user can also control the learner's as well as the interlocuter's general approach to simplification by operating the respective sliders positioned at the bottom of the two columns at the right of the interface.

Figure 6 shows the interface of a simplified version of the WTP model. The full version contains a large number of parameters that space in this chapter does not allow us to explain. It may suffice to mention that those parameters include (but are not limited to) the following aspects of SLA:

- the proportion of sentences produced by the learner and the interlocutor,
- the uptake of values from the interlocutor,
- the uptake values from the learner's own production,
- the uptake of lexical items from the interlocutor,
- the recycling of lexical items from the learner's own production.

For the purpose of this chapter we have included in the illustration only those parameters that we are focussing on.



Figure 6. WTP model with neutral settings and 'other-simplification' set at 0.49

In principle, there are two classes of parameters shaping the simulation. (1) The first set are those that specify aspects of the ‘dialog’ between the simulated learner and a simulated interlocutor who may be a native speaker of the target language or a non-native speaker. In this version, the first of these parameters addresses simplification. It indicates the degree of ‘learner-simplification-bias’ (the extent to which the learner is inclined to simplify [use -cop] sentences) and ‘other-simplification-bias’ (the extent to which the interlocutor is inclined to simplify sentences, potentially using ‘foreigner talk’).¹⁴ (2) The second set of parameters operationalize certain mental behaviors of the learner. In this version, the two behaviors included are (2a) ‘entrenchment’ (the point at which the learner implicitly decides that there is no need to pay further attention to features of the language they encounter) and (2b) ‘forgetting’ (how extensively and on what basis the learner forgets the particular words and their associated copula values, for example whether anything that has been learned can be forgotten at any time or whether forgetting starts to occur only after a certain number of words have been learned). The box at the bottom displays the simplification index as a black graph. The most recent simplification index value is displayed in the small window above the graph.

A series of simulations revealed that with the class 2 parameters turned off, any setting of the class 1 parameters yields a ‘simplification index’ that stabilizes at a specific point after a limited number of iterations. This is due to the way the index is calculated. Class 1 parameters feed 1:1 into the simplification index. For instance, in the situation portrayed in Figure 6, the initial simplification is set at 0.5 and the simulation sets the ‘initial state’ at 0.498 simplification (by a process that involves a degree of chance). If the ‘other-simplification-bias’ (i.e., the linguistic simplification of the interlocutor) is set at 0.49 a simplification index of about 0.49 emerges with all other parameters in neutral (or ‘off’) position. This value remains constant even though the learner’s simplification bias is 0, i.e., a bias to consistent inclusion of the copula. In such a scenario, with only class 1 parameters active, this implies that from iteration to iteration words are assigned the values specified by the selected interlocutor parameters.

A fundamentally different process is initiated once the continuous assignment of values to words is allowed to be influenced by intervening factors such as ‘forgetting’. In this case, the learning of the values of words will start to interact with the forgetting of the values of words, resulting in a non-linear development of the index. Figure 7 displays this kind of behavior for the intervening variable ‘forgetting’. The class 1 parameters are set as follows: ‘learner-simplification-bias’ = 0 (the learner

14. Both native and non-native speakers may simplify their talk when addressing a non-native speaker. This use of a language, has been referred to as use of a ‘simplifying register’ or (in the case of native speakers) ‘foreigner talk’ (see for instance Clyne, 1981).

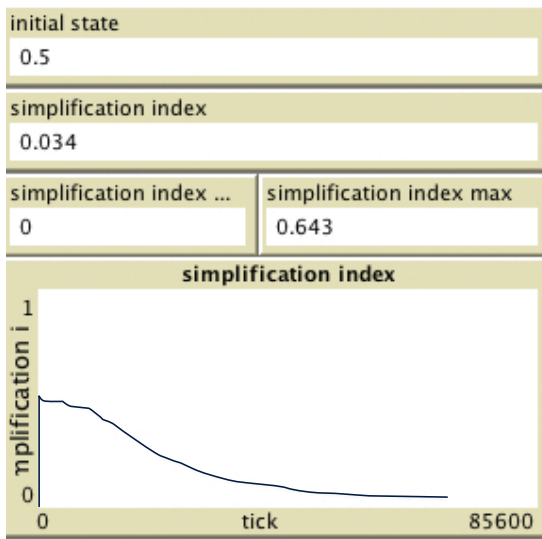


Figure 7. WTP model with ‘other-simplification’ set at 0.49 and ‘forget-behavior’ set to ‘progressive’

supplies the copula consistently), ‘other-simplification-bias’ = 0.49 (the interlocutor omits the copula on 49% of occasions) and ‘initial simplification’ (the pattern of simplification that the learner had established prior to the commencement of the simulation) is 0.50.

In Figure 7, the parameter ‘forgetting’ has been set to ‘progressive’. In this mode the simulated learner may forget the values of words before having reached the set lexeme memory size (i.e., 100 that has been carried over from Figure 6 to Figure 7). This has the effect that selected records (i.e., the values of specific words) are deleted from the database, and the simplification index is re-calculated without the ‘forgotten’ words in the next iteration. And this process continues from iteration to iteration.

As can be seen in Figure 7, the simplification index that is created in a simulation with these interacting parameters displays a continuous, negative growth that levels off at about 0.03. This result stands in stark contrast to the simulation shown in Figure 6 where, as indicated, all parameters except ‘forgetting’ are set in exactly the same way as in Figure 7. In other words, activating the parameter ‘forgetting’ causes a continuous and bounded (negative) growth pattern for the simplification index such that the learner pattern no longer directly reflects the input.

When we repeat this simulation with one minor difference from Figure 7 in the settings – with other-simplification set at 0.51 instead of 0.49 – the graph

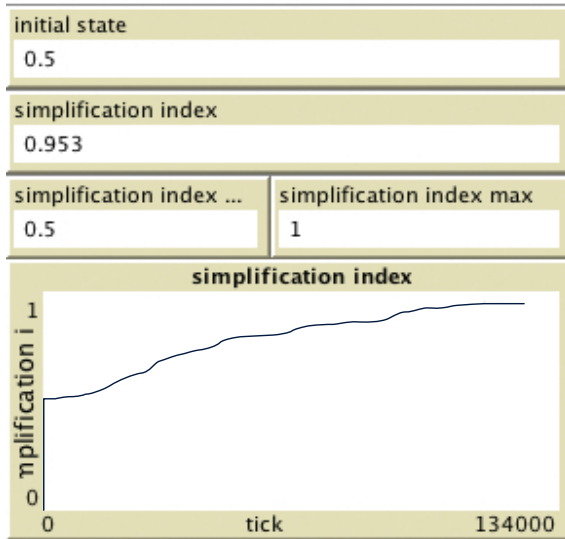


Figure 8. WTP model with ‘other-simplification’ set at 0.49 and ‘forget-behavior’ set to ‘progressive’

representing the development of the simplification index moves in the opposite direction, as shown in Figure 8.

In cognitive psychology, Ebbinghaus (1885) discovered the ‘learning curve’, a set of empirically validated behavioral characteristics of learning and forgetting organisms that includes a steep initial incline that eventually levels off at the upper bound of the carrying capacity. A curve with these features is also referred to as a logistic curve, and it can be generated by a logistic map, as originally developed by Verhulst (1838) in the context of modeling population developments that we discussed in Section 3. A logistic map is suitable for modeling development up to an upper bound, because it is designed for circumstances in which the potential maximum size is limited.

As these simulations show, the parameter ‘forgetting’ sets in motion a dynamic development of the simplification index, resulting in a logistic ‘growth’ – where ‘growth’ can entail either a steady and bounded increase or a steady and bounded decrease of the index. We have shown how the index increases steadily with ‘other-simplification’ initially set at 0.51 or above, and it decreases steadily with ‘other-simplification’ initially set at 0.49 or below. We have further shown that forgetting has a highly selective effect as it has the capacity to amplify any preference for –cop or +cop to the extreme, given a sufficiently long process (i.e., a sufficient number of iterations). The surprising outcome of this process is that with an initial

interlocutor simplification bias smaller than 0.5 the learner ends up simplifying less than the interlocutor. And inversely, with an initial interlocutor bias greater than 0.5 the learner ends up simplifying more than the interlocutor,¹⁵ which suggests that the learner's ultimate achievement is not solely determined by the features of the input when class 2 parameters are included. In each case the learner amplifies a pattern observable in the input and so exceeds the extent of the pattern that the input models.

A very similar phenomenon occurs with another intervening variable that has been labelled using a different interpretation of the term 'entrenchment' (see Watkins et al. 2013).¹⁶ As indicated, the interpretation of entrenchment that we use for this purpose comes from social psychology, and it refers to a behavior where individuals who potentially have several choices in relation to a particular decision/opinion will stick to one choice after they have encountered a certain number of individuals who have opted for this choice, and they do not change their specific choice even after encountering other choices or individuals who have made different choices. (cf. Watkins et al. 2013, <https://www.ecologyandsociety.org/vol18/iss2/art32/>).

This interpretation of 'entrenchment' is the second variable that (like forgetting) sets in motion two counter-acting lines of development, related to (1) the continuous and cumulative marking of words with +cop and -cop and (2) the increase or decrease of values depending on certain tipping points. The upper bound in the simulation displayed in Figure 9 is around 0.7. This is due to the chosen values for the parameters 'entrenchment-threshold' (defining the number of values required for a lexical entry before entrenchment is activated), 'entrenchment activation' (defining the proportion of plus and minus values in a lexical entry that is required for entrenchment to occur) and 'entrenchment punishment' (defining the degree of entrenchment). Whereas 'forgetting' can result in either the maximum (1) or the minimum (0) index, the entrenchment result demonstrates that the growth patterns associated with these kinds of processes can be simulated in ways that do not necessarily target either of these extreme bounds.

15. The computational process underlying the growth phenomenon is the following: the simplification bias that translates into a flat line at a fixed point with only class 1 parameters converts into a developmental trend when – due to the forgetting process – the same value (i.e., 1) is subtracted from the minus and/or plus markings of the activated words, while the learning process is still ongoing. In effect, even the smallest change in the bias above or below the midpoint sets off a dynamic trend that mimics a learning curve and results in a plateau close to 1 or 0 depending on the direction of the trend.

16. Watkins et al.'s notion of entrenchment is not identical with Wimsatt's notion of generative entrenchment although both notions utilize the same metaphor.

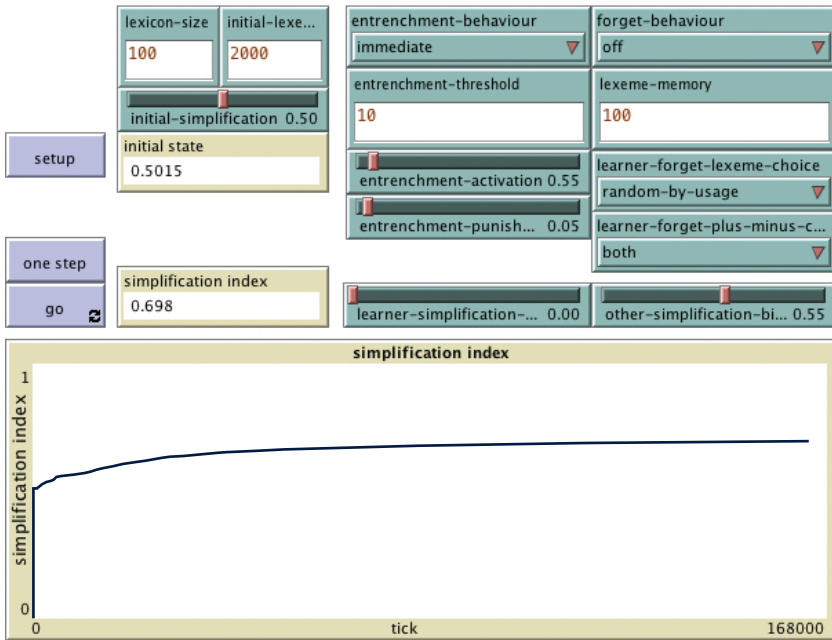


Figure 9. Entrenchment

So far our simulations have focused on just two of many possible parameters affecting second language acquisition. The purpose of our simulations was to study – in a simulated longitudinal setting – aspects of the internal dynamics of the process that could potentially lead to an enduring preference for the inclusion or exclusion of the copula. Focussing on individual parameters with the experimental exclusion of other factors, our simulations have shown that, in this dynamical system, antagonistic processes such as learning vs. forgetting or learning vs. entrenchment have the capacity to amplify very small preferences. The direction of the amplification that the simulation reveals depends on the initial value of the intervening variable. The magnitude of the amplification increases with the number of iterations but also levels off at some point. A minute difference in a parameter value can tip the balance in one direction or the other. In other words, definitionally significant characteristics of dynamical systems can be simulated.

Before we consider the implications of these findings for the choices learners make when faced with developmental problems, it may be useful to recapitulate the following three fundamental facts about variational features that we documented in Section 2. (1) The choice between alternative features has repercussions for later development. (2) Learners are consistent in their variational choices, but (3) the choice of variational features is not black-or-white; instead it is shaped by complex networks of influence.

The presence of complex networks of influence means that we need to consider the amplifying effect of dynamic, antagonistic processes. Our simulations show that with parameters such as forgetting or entrenchment as intervening variables, over time the initial proportions of the two variants will gradually change, in favour of one. The exact curves that might result from the interaction of such parameters could only be predicted if it were possible to include all of the relevant parameters and precisely indicate the timing of the activation of the parameters for the individual learner.

The trend we have seen so far for these parameters is that ‘forgetting’ (of previous examples) will increase the simplification index for values greater than 0.5 and decrease the simplification index for initial values smaller than 0.5. Entrenchment will increase all values greater than 0.5 if set on a low activation level, with a low ‘punishment’ level accompanied by a low threshold level. In other words, our simulations demonstrate that dynamic processes related to the two mental limitations of the learner that we have specified amplify simplification indices that are initially moderate, thus moving the index toward the area of one or other of the extremes. Our prediction is that this process will have repercussions when the interlanguage system reaches the next developmental level. Our claim is that when the learner is faced with a developmental challenge, the dynamic interactions reflected in the simplification index will, by their very nature, pre-empt the choice of structure that is due to emerge (i.e., WH pro/n (cop) N/A? – ‘Where he is?’ or – ‘Where he?’). An index that indicates a dominance of simplification will signal structures without the copula, that will therefore provide no resources on which further development could be based. As we pointed out above, this option leads into a developmental dead-end, one that we believe demonstrates a view of stabilization that is testable.

5. Empirical validation

Our hypotheses and our simulations focused on a specific aspect (equational sentences) within a specific developmental segment of second language acquisition. We focused on the developmental segment whose beginning is defined by the appearance of the first equational sentences and whose ending is marked by the appearance of equational sentences with fronted WH-words. We simulated the dynamics of copula insertion within this segment, yielding specific amplification processes instigated by specific initial states and specific values of the chosen parameters.

We will now proceed to compare our simulations with empirical data from longitudinal studies of second language acquisition. The data required for the validation process need to fit the constraints that were applied to the simulations, i.e.,

contain rich, extensive and varied examples. We will focus on the following five longitudinal studies that satisfy these requirements:

- Hakuta's (1974) study of Uguisu's acquisition of ESL (L1 = Japanese, observation from 5;4 – 6;5)
- Haznedar's (1997) study of Erdem's acquisition of ESL (L1 = Turkish, observation period: 18 months, 4;3 at the start of the study)
- Nicholas' (1987) study of Cindy's acquisition of GSL (L1 = English, covering 8 months, 3;4 at the start of the study)
- Pienemann's (1980, 1981) studies of Luigina's and Concetta's acquisition of GSL (L1 = Italian, observation period: 14 months, 8 years at the start of the study).¹⁷

In the studies by Nicholas (1987) and by Pienemann (1980, 1981) the L2 is German, whereas in the studies by Haznedar (1997) and Hakuta (1974) it is English. However, the development of equational sentences is identical in ESL and GSL as illustrated below with German examples included in the same pathways:

	Structure	Path 1: +copula	Path 2: –copula
1.	N/A (?)	– 'groß'/ 'groß?'	– 'groß'/ 'groß?'
2.	N/A/pro (cop) N/A (?)	– 'er ist groß'/ 'er ist groß?'	– 'er groß'/er groß?'
3.	WH pro/N (cop) N/A?	– 'Wo er ist?'	– 'Wo er?'
4.	Cop pro/N + A?	– 'ist er groß?'	–
5.	WH cop pro/N/A?	– 'Wo ist er?'	–
6.	S whether/Wh pro/N cop N/A	– 'Ich frage, wo er ist.'	–

Hakuta (1974)

The dynamics of the omission of the copula developed as follows in Hakuta's (1974) study of Uguisu's acquisition of ESL, ranging from initially very high omission to no omission at all at various later times:

Weeks	13	15	17	18	19	25
Index	0.90	0.30	0.00	0.20	0.00	0.00

Our simulation software affords an analysis of the match of these values with the values produced in the corresponding simulations. Figure 10 provides an overview

17. Conversational and task-based data were audio-recorded in individual sessions with the children lasting up to one hour each. All data were fully transcribed and analysed for the acquisition of word order, constituent structure and suppliance of obligatory elements in line with the principles set out in Meisel, Clahsen, & Pienemann (1981).

of these comparisons of the naturalistic longitudinal L2 data with the simulations in the form of curves representing the simplification index (i.e., rate of copula omission) at the corresponding points in time.

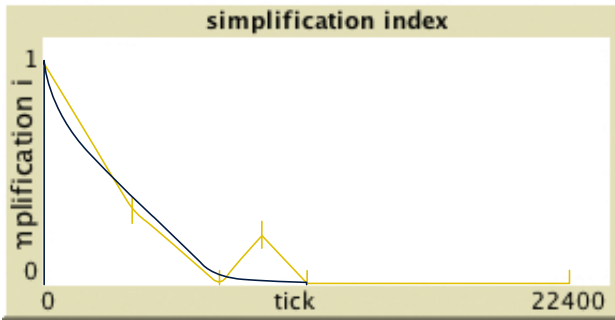


Figure 10. Simplification index in simulation (black line) and in Uguisu's longitudinal data (yellow line)

The software accurately relates the number of weeks of observation in the naturalistic data to the same relative temporal distances in terms of iterations (i.e., 'ticks'). We ensured that the same initial state was chosen in our simulations as in the corresponding naturalistic data, in the case of Uguisu, 0.9. In the simulation, the dynamics that followed from these initial states were entirely driven by the iterative process based on the chosen parameters.

A visual inspection of Figure 10 reveals that the simulated and the real curves are very similar. This impression is quantified in the following statistics that attribute an 'excellent' fit to both sets of data and the corresponding simulations:¹⁸

RMSE:	0.088
NRMSE:	0.001
PS:	0.93 (Excellent)

The calculations underlying these analyses are explained in footnote twenty.¹⁹ As mentioned above, the temporal distances between times of data collection (weeks) were lined up equidistantly with the number of iterations. The initial states of the naturalistic data were entered as initial states for the simulations. 'Sim-Idx', 'Diff'

18. The full reference data for this calculation are provided in the appendix.

19. The above analyses are based on the following goodness-of-fit statistics.

– The Root-Mean-Squared Error (RMSE) that is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n}}$$

and ‘Diff²’ refer to steps in the calculation of the Root-Mean-Squared Error (RMSE), the Normalized Root-Mean-Squared Error (NRMSE) and the Performance Score (PS). Values are provided for the RMSE, NRMSE and the Performance Score at the end of the analysis.

An ‘excellent’ Performance Score was calculated for the simulations of this longitudinal study. This means that the simulation agrees almost perfectly with the Hakuta’s (1974) natural data. This result underscores the significance of ABM simulations as a potential alternative or complementary approach to the use of real data in situations where real data are difficult to obtain.

Haznedar (1997)

We subjected the data contained in Haznedar’s (1997) study of the acquisition of ESL by Erdem to the same analytical steps as those of Uguisu’s ESL acquisition (Hakuta, 1974). The relevant observation period ends in sample 18, as Erdem acquired Yes/No Inversion for equational sentences (stage 4) in sample 11 (see Haznedar, 1997, p. 328).²⁰ In Haznedar’s data set, the omission of the copula showed the following dynamics for recordings spread out according to the number of days from the commencement of recordings:

Days	31	56	70	86	94	126
Index	1.0	1.0	0.06	0.13	0.13	0.16

These data were compared with our simulation that was based on the same initial state (i.e., 1.0). The corresponding curves for the natural data and the simulation are displayed in Figure 11.

-
- The Normalized Root-Mean-Squared Error (NRMSE) that is defined as

$$\text{NRMSE} = \frac{\sum (S_i - O_i)^2}{\sum O_i^2}$$

O = Learner data, S = Simulation data

- And the Performance Score (PS) calculated as a Nash-Sutcliffe Score (Nash and Sutcliffe, 1970)

$$\text{PS} = 1 - \frac{\langle (x_c - x_m)^2 \rangle}{\langle (x_m - x_R)^2 \rangle}$$

Xc = simulation data, Xm = learner data, XR = mean Xm

²⁰ On page 328 Haznedar (1997) displays Appendix D-1 that contains a table with the raw numbers for “Number and Percentage of Yes/No questions”. The table reveals that Yes/ No questions were acquired in session 11.

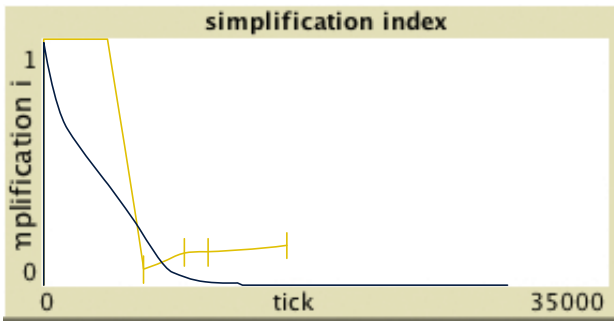


Figure 11. Simplification index in simulation (black line) and in Erdem's longitudinal data (yellow line)

The goodness of fit was calculated for these curves in the same way as above,²¹ and a Performance Score of 0.632 resulted from this calculation, which is judged as 'good' by (Nash & Sutcliffe, 1970). This result is not as strong as that achieved with Hakuta's (1974) data, but still suggests that the model provides a satisfactory simulation of Haznedar's (1997) report of Erdem's copula use patterns.

RMSE: 0.217
 NRMSE: 0.013
 PS: 0.66 (Good)

Nicholas (1987)

We repeated this procedure for the longitudinal study by Nicholas (1987) of Cindy's GSL development with the following results, also using days to separate the recordings:

Days	21	23	28	30	34	36	41	43	48	50	55
Index	1.0	0.33	0.33	0.5	0.27	0.24	0.16	0.15	0.22	0.13	0.13
RMSE:	0.164										
NRMSE:	0.007										
PS:	0.75 (Good)										

Similar to the result for Haznedar (1997), this result indicates that the model provides a satisfactory simulation of Cindy's copula use patterns. A PS value of 0.75 is close to the threshold for an 'excellent' value (i.e., 0.8).

21. Again, the full reference data for this calculation are provided in the Appendix.

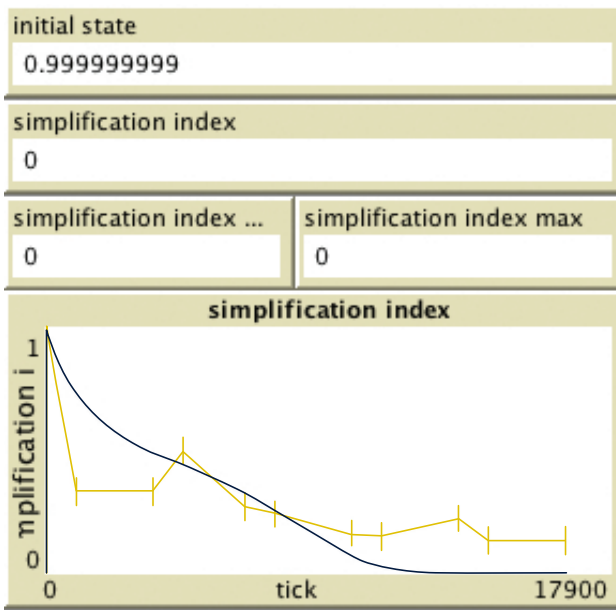


Figure 12. Simplification index in simulation (black line) and in Cindy's longitudinal data (yellow line)

Pienemann (1980, 1981)

The analytical procedure described above was also applied to the two longitudinal studies on GSL acquisition conducted by Pienemann (1980, 1981) – with the following results with the recordings separated according to weeks:

Luigina

Week	49	55	60	64	85
Index	1.0	0.91	0.5	0.13	0.0

Concetta

Week	39	47	53	58	62
Index	0.17	0.0	0.0	0.0	0.0

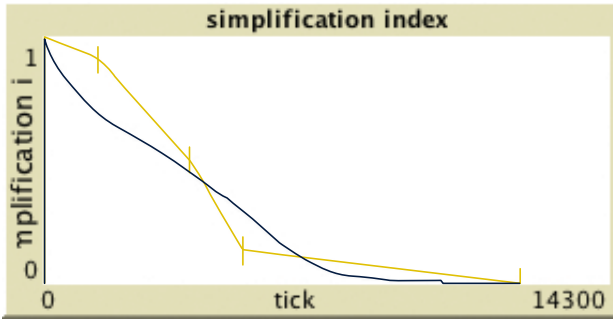


Figure 13. Simplification index in simulation (black line) and in Luigina's longitudinal data (yellow line)

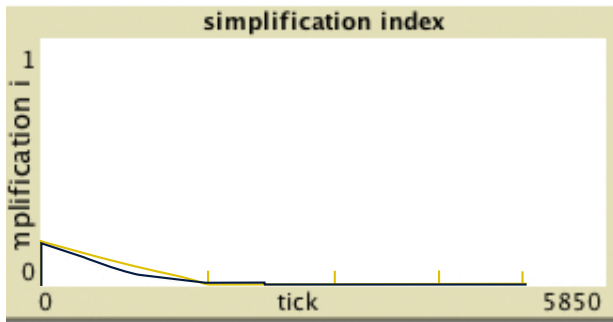


Figure 14. Simplification index in simulation (black line) and in Concetta's longitudinal data (yellow line)

Luigina

RMSE: 0.124
 NRMSE: 0.001
 PS: 0.867 (Excellent)

Concetta

RMSE: 0.004
 NRMSE: 0.000
 PS: 1.0 (Excellent)

Summing up, a comparison of our simulations with five longitudinal studies yielded the following results:

- Uglusu: RMSE: 0.088 NRMSE: 0.001 PS: 0.93 (Excellent)
- Erdem: RMSE: 0.217 NRMSE: 0.013 PS: 0.66 (Good)
- Cindy: RMSE: 0.164 NRMSE: 0.007 PS: 0.75 (Good)
- Luigina: RMSE: 0.124 NRMSE: 0.001 PS: 0.867 (Excellent)
- Concetta: RMSE: 0.004 NRMSE: 0.000 PS: 1.0 (Excellent)

As this overview shows, the data from three of the five longitudinal studies show an ‘excellent’ match with the simulation data, and two studies show a ‘good’ match. The extent of congruence that we have been able to demonstrate suggests that our simulations provide a valid representation of real, longitudinal data.

In all five learners included in our validation study the simplification index developed towards zero. Our hypotheses about generative entrenchment in L2 development predict that in subsequent stages all five learners will develop their L2s within the non-stabilizing strand, thus moving further in the developmental hierarchy. Pienemann’s (1980, 1981) studies confirm this hypothesis. Concetta’s interlanguage develops up to level 5 of the PT hierarchy during the 62-week observation period, and Luigina produces ‘WH pro cop?’ (i.e., *‘Where he is’), a structure that characterizes the non-stabilizing direction. The other three learners reported by Hakuta (1974); Nicholas (1987) and by Haznedar (1997) also developed further in the non-stabilized direction.

In other words, all five (child) informants included in our validity study do not stabilize in their L2 acquisition. It is interesting to note that Dyson (2021) found in her longitudinal study of eight adolescent learners of English that the younger learners did not appear to stabilize in their L2 acquisition, although their rate of development varied. In the context of this study it would have been interesting to include stabilizing learners as well – including adult learners. However, relevant longitudinal data are very hard to come by. This was also noted by Mike Long in his 2003 paper that we quoted extensively above. The main reason for this difficulty is that conducting longitudinal studies of SLA is very labor-intensive and even more so when working with adults in uncertain and often vulnerable life circumstances. One needs to start with more informants than one expects to analyze because inevitably informants may move away, fall ill or simply get tired of the researcher. Data elicitation sessions may yield no relevant data. The recording equipment may fail. Data may be lost due to technical problems or due to human error. Data collection may start too late or end too early for the linguistic phenomena the researcher may be interested in.

In the absence of this kind of longitudinal data, it may be instructive to revisit the extensive cross-sectional data set collected and analysed by Meisel and his co-workers (Meisel, Clahsen, & Pienemann, 1981; Clahsen, Meisel, & Pienemann, 1983) and to complement our re-analysis with suitable simulations where needed. We quoted this study above to demonstrate that learners’ interlanguages contain clusters of similar variational features. Pienemann (1998, p. 327) showed in a re-analysis of data contained in the study that long-term residents with GSL and a high simplification index (including extensive copula omission) do not show evidence of having developed further than level 3 – despite their long-term residency. Further details are displayed in Table 5 that also shows that the interlanguage of many learners with SHORT periods of residence has a considerably lower rate

of copula omission.²² In other words, the high omission rate is not an indicator of IL development but it may be a key factor in a developmental stalemate.

Table 5. Informants from the ZISA study (Clahsen, Meisel, & Pienemann, 1983) at or below level 3

A. Informants with long periods of residence

Name	years in Germany	PT level	omission rate (copula)
Dolores S	15	3	0.90
Montse S	14	3	0.88
Pepita S	14	2	1.00
Eliseo I	12	3	0.86
Antonio S	11	2	1.00
Pascqua S	7	2	0.75
Pasquale I	7	2	0.86
Rosa I	7	2	0.88

B. Informants with short periods of residence

Name	Years in Germany	PT level	Omission rate (copula)
Zita P	1	3	--
Manuel P	1	3	0.67
Miguel S	1	3	0.40
Estefania P	3	3	0.44
Rosemarie S	1	2	0.43
Francisco I	2	2	1.00

These data add plausibility to the hypothesis that L2 development may stall in highly simplifying learners (Pienemann's (1998) 'bad-choice hypothesis'). In this paper we have shown how what we now refer to as the 'wrong track pathway' is determined by the dynamics of SLA that shape the direction of further development. In the case of equational sentences, extensive copula omission structurally excludes the further development of equational sentences with a copula, thus pre-empting the development of Y/N questions and WH-inversion in equational sentences.

This likely impact of copula omission rates on further development is obvious in all those cases where the omission rate is 1 when contexts for the new sentence structure (XScopY)²³ appear. In the case of structures that are close to 1, one needs

22. Note, however, that Francisco I entirely omits the copula with a very short residence period. In other words, high omission is not solely a correlate of long residence either.

23. This structure appears at level 3 as shown in Table 2 above where the relevant line is as follows:
"3 WH pro/n (cop) N/A? – 'Where he is?' – 'Where he?'"

to look at the dynamics inherent in the development of sentence structures. In those cases (Pepita, Antonio, Francisco in Table 5) there will be competition between XScopY and XSøY²⁴ for a short period. We assume that this competition will be subject to the same kind of dynamics as those discussed above for the insertion or omission of the copula. And we hypothesize that due to the dynamics of the competition process, XSøY will prevail.

In order to test this hypothesis, a simulation of such a dynamic competition process was carried out with 0.88 as the initial state and 0.88 as the simplification bias (thus imitating Montse S and Rosa I) and with the parameter ‘forgetting’ active. The outcome of this simulation is displayed in Figure 15 as an orbit representing the development of the proportion of XScopY and XSøY.

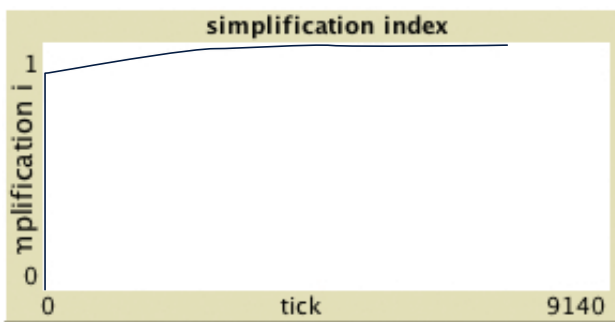


Figure 15. Simulated prediction of the copula use pattern of Montse S and Rosa after the data collection point

Figure 15 shows that after a relatively small number of iterations the orbit approaches 1, indicating that, for the given values, XSøY will prevail. We also carried out a simulation with the copula omission rate from the last session of the relevant observation period in Haznedar’s (1997) study (i.e., 0.16) in order to complement our simulations with a situation found at the other end of the continuum when omission rates are close to but not equal to zero when level 3 structures emerge. We found that the orbit representing the development of the proportion of XScopY and XSøY reaches 0 after a relatively small number of iterations – as shown in Figure 16. We take these observations as supporting evidence for our hypothesis that the direction of further development at choice points is determined by the internal dynamics of the acquisition process.

24. XSøY denotes equational sentences without a copula as in ‘Where he?’.

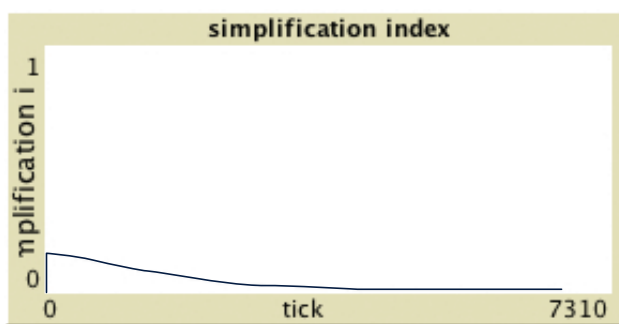


Figure 16. Simulation-based prediction of Haznedar's further copula use pattern after the recorded period

However, we need to add to this conclusion that our simulations have included only a few selected (Class 2) variables (specifically 'forgetting' and 'entrenchment') in order to make this presentation of our core hypotheses manageable in the context of a book chapter. Naturally, the effect of these variables may be altered by other variables that may be effective at particular points in time during L2 development. One example of such an intervening variable is the amount of L2 simplification produced by the interlocutor. The purpose of this chapter was to present a first step in testing our hypotheses about the constraining effect of learner variation on developmental pathways by focusing on a small set of key variables affecting some of the dynamics of SLA.

6. Conclusion

In this paper we focused on the contested issue of fossilization and stabilization. We agree with Long's rejection of the concept of fossilization and his preference for stabilization. We have offered a definition of stabilization as a lack of development (measured using the emergence criterion) within structural domain X , in context Y , and over time period Z . In line with Long's argument for the need of an approach to stabilization that yields predictive power, we have presented the 'wrong track pathway' as a dynamic approach towards explaining stabilization and we have pointed out that in its original form the 'bad-choice hypothesis' did not spell out how learners arrive at good or bad choices. Therefore we have added a new component to what we have termed the 'wrong track pathway' that is based on the dynamics leading up to the point of (implicit) decision-making. We explored these dynamics for the development of equational sentences in a set of ABM-based

simulations. These simulations demonstrated that initial trends in simplification will be amplified by mental processes such as forgetting or entrenchment leading to a state of the interlanguage that permits only one of two strands for further development – with one pathway leading to a dead-end.

We demonstrated the validity of these ABM-based simulations of L2 development by comparing the simulations with empirical longitudinal data and by calculating the goodness of fit for this comparison, which turned out to be excellent. In our view, this exercise demonstrates that ABM-based simulations can be a powerful tool for exploring the dynamics of second language acquisition. Given that agent-based models constitute a method of calculating and simulating processes within dynamical linguistic systems, the successful application of ABMs to second language development also implies that we have identified an area of SLA in which dynamical systems theory can be applied in its original mathematical framing.

If we can consolidate the validity of ABM-based simulations of L2 development we may in the future be in a position to use these kinds of simulations as an additional source of insights about the dynamics of L2 development – particularly in cases where natural data are hard or impossible to come by.

Acknowledgements

We would like to thank Bruno Di Biase, Susanne Carroll, Bronwen Dyson and Jürgen M. Meisel for their thoughtful comments on an earlier version of this paper. We are grateful to Xia Li, statistical consultant at La Trobe University for her confirmation of the appropriateness of the statistical procedures that we have used.

References

- Bailey, C.-J. (1973). *Variation and linguistic theory*. Center for Applied Linguistics.
- Bickerton, D. (1971). Inherent variability and variable rules. *Foundations of Language*, 7, 457–492.
- Borshchev, B., & Filippov, A. (2004). *From system dynamics and discrete event to practical agent based modeling: Reasons, techniques, tools*. The 22nd International Conference of the System Dynamics Society, 25–29 July 25, Oxford, England.
- Bresnan, J. (2001). *Lexical-functional syntax*. Blackwell.
- Clahsen, H., Meisel, J. M., & Pienemann, M. (1983). *Deutsch als Zweitsprache. Der Zweitspracherwerb ausländischer Arbeiter* [German as a second language. The second language acquisition of immigrant workers]. Gunter Narr.
- Clyne, M. (Ed.). (1981). Foreigner talk. *International Journal of the Sociology of Language*, 28.

- De Bot, K. (2016). Multi-competence and dynamic systems. In V. Cook & Li Wei (Eds.), *The Cambridge handbook of linguistic multi-competence* (pp. 125–141). Cambridge University Press. <https://doi.org/10.1017/CBO9781107425965.006>
- De Bot, K., Lowie, W. M., & Verspoor, M. H. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21. <https://doi.org/10.1017/S1366728906002732>
- DeCamp, D. (1973). What do implicational scales imply? In C.-J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English* (pp. 141–148). Georgetown University Press.
- Dyson, B. P. (2021). *Dynamic variation in second language acquisition. A language processing perspective*. John Benjamins. <https://doi.org/10.1075/palart.8>
- Ebbinghaus, H. (1885). *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Duncker & Humblot. *English translation: Memory: A contribution to experimental psychology*. Columbia University, 1913.
- Feldman, D. P. (2012). *Chaos and fractals. An elementary introduction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199566433.001.0001>
- Feldman, D. P. (2019). *Chaos and dynamical systems*. Princeton University Press.
- Hakuta, K. (1974). Learning to speak a second language: What exactly does the child learn? In D. Dato (ed.), *Georgetown University round table on languages and linguistics 1975* (pp. 193–207). Georgetown University Press.
- Han, Z. H. (2004). *Fossilisation in adult second language acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781853596889>
- Hatch, E., & Lazaraton, A. (1991). *The research manual. Design and statistics for applied linguists*. Heinle and Heinle.
- Haznedar, B. (1997). Child second language acquisition of English: A longitudinal case study of a Turkish-speaking child (Unpublished doctoral dissertation). Durham University.
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition research* (pp. 539–587). Blackwell. <https://doi.org/10.1002/9780470756492.ch17>
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258. https://doi.org/10.1207/s15516709cog1102_5
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619. <https://doi.org/10.1093/applin/aml029>
- Larsen-Freemans, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. Longman.
- Lenzing, A., Nicholas, H., & Roos, J. (Eds.), (2019). *Widening contexts for Processability Theory: Theories and issues*. John Benjamins. <https://doi.org/10.1075/palart.7>
- Lenzing, A., Pienemann, M., & Nicholas, H. (2022). Lost in translation? On some key features of dynamical systems theorizing invoked in SLA research. In K. Kersten & A. Winsler (Eds.) *Understanding variability in second language acquisition, bilingualism and cognition*. Routledge.
- Li, T.-Y., & Yorke, J. (1975). Period three implies chaos. *The American Mathematical Monthly*, 82(10), 985–992. <https://doi.org/10.1080/00029890.1975.11994008>

- Long, M. H. (2003). Stabilization and fossilization. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition research* (pp. 487–536). Blackwell.
<https://doi.org/10.1002/9780470756492.ch16>
- May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261(5560), 459–467. <https://doi.org/10.1038/261459a0>
- Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in second language acquisition. *Studies in Second Language Acquisition*, 3(2), 109–135.
<https://doi.org/10.1017/S0272263100004137>
- Meisel, J. M. (2011). *First and second language acquisition*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511862694>
- Mitchell, M. (2009). *Complexity. A guided tour*. Oxford University Press.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
[https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Nicholas, H. (1987). A comparative study of the acquisition of German as a first and as a second language (Unpublished doctoral dissertation). Monash University. Retrieved on 2 February 2022 from https://monash.hosted.exlibrisgroup.com/MON:au_everything:catau21186631810001751
- Nicholas, H., Lenzing, A., & Roos, J. (2019). How does PT’s view of acquisition relate to the challenge of widening perspectives on SLA? In A. Lenzing, H. Nicholas, & J. Roos (Eds.), *Widening contexts for Processability Theory: Theories and issues* (pp. 391–398). John Benjamins.
<https://doi.org/10.1075/palart.7.17nic>
- Pienemann, M. (1980). The second language acquisition of immigrant children. In S. W. Felix (Ed.), *Second language development. Trends and issues* (pp. 41–56). Narr.
- Pienemann, M. (1981). *Der Zweitspracherwerb ausländischer Arbeiterkinder* [The second language acquisition of immigrant children]. Bonn: Bouvier.
- Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. John Benjamins. <https://doi.org/10.1075/sibil.15>
- Pienemann, M. (2005) (ed.) *Cross-Linguistic Aspects of Processability Theory*. John Benjamins: Amsterdam.
- Pienemann, M. (2011). The psycholinguistic basis of PT. In M. Pienemann & J.-U. Kefßler (Eds.), *Studying Processability Theory* (pp. 27–49). John Benjamins.
<https://doi.org/10.1075/palart.1.03the>
- Pienemann, M. (2015). An outline of processability theory and its relationship to other approaches to SLA. *Language Learning*, 65(1), 123–151. <https://doi.org/10.1111/lang.12095>
- Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition*, 10, 217–243.
<https://doi.org/10.1017/S0272263100007324>
- Pienemann, M., & Lenzing, A. (2020). Processability Theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition. An introduction* (3rd ed., pp. 162–191). Routledge. <https://doi.org/10.4324/9780429503986-8>
- Pienemann, M., Lenzing, A., & Nicholas, H. (forthcoming). Can dynamic systems theory explain SLA? Unpublished manuscript, University of Paderborn, Germany.

- Poincaré, J. H. (1913). *The foundations of science: Vol I, Science and hypothesis; Vol II, The value of science; Vol III, Science and method* (G. Halsted, Trans.). The Science Press. (Original works published in 1902 (Vol I), 1905 (Vol II), 1908 (Vol III)).
- Rasmussen, N. (1987). A new model of developmental constraints as applied to the Drosophila system. *Journal of Theoretical Biology*, 127(3), 271–299. PMID:3431126. [https://doi.org/10.1016/S0022-5193\(87\)80107-8](https://doi.org/10.1016/S0022-5193(87)80107-8)
- Scheffer, M. (2010). Foreseeing tipping points. *Nature*, 467, 411–412. <https://doi.org/10.1038/467411a>
- Selinker, L. (1972). Interlanguage. *IRAL*, 10, 209–231. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106, 467–482.
- Strogatz, S. H. (2015). *Nonlinear dynamics and chaos. With applications to physics, biology, chemistry and engineering (2nd ed.)*. Westview.
- Verhulst, P.-F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique*, 10, 113–121. (cited after Nicolas Bacaër, 2011).
- Watkins, C., Massey, D., Brooks, J., Ross, K., & Zellner, M. L. (2013). Understanding the mechanisms of collective decision making in ecological restoration: An agent-based model of actors and organizations. *Ecology and Society*, 18(2), 32. <https://doi.org/10.5751/ES-05497-180232>
- Wilensky, U. (1999). NetLogo. *Center for Connected Learning and Computer-Based Modeling*, Northwestern University, Evanston, IL. Retrieved on 2 February 2022 from <http://ccl.northwestern.edu/netlogo/>
- Wimsatt, W. (1986). Developmental constraints, generative entrenchment and the innate-acquired distinction. In W. Bechtel. (ed.), *Integrating scientific disciplines* (pp. 185–208). Martinus Nijhoff. https://doi.org/10.1007/978-94-010-9435-1_11

Appendix 1

===== reference data Uglusu =====

Entrenchment: off, forget: progressive + default, initial state 0.9, other simplification 0.0

===== reference data =====

	Time	Ref-Idx	Tick	Sim-Idx	Dif to Ref	Dif ²	Dif to Ref-Mean	Dif ²	Dif to 0.5	Dif ²	
0	->	13	0.9	0	0.899	0.001	0	0.666	0.443	0.399	0.159
1	->	15	0.3	3099	0.392	-0.092	0.008	0.159	0.025	-0.108	0.012
2	->	17	0	6199	0.041	-0.041	0.002	-0.193	0.037	-0.459	0.211
3	->	18	0.2	7748	0.011	0.189	0.036	-0.223	0.05	-0.489	0.24
4	->	19	0	9298	0.004	-0.004	0	-0.23	0.053	-0.496	0.246
5	->	25	0	18597	0	0	0	-0.233	0.054	-0.5	0.25

RMSE: 0.088

NRMSE: 0.001

PS: 0.93 (Excellent)

===== reference data Erdem=====

entrenchment: threshold 3, activation 0.95, punishment 0.05

forget: progressive + default settings, other simplification 0.0, initial state 1.0

===== reference data =====

	Time	Ref-Idx	Tick	Sim-Idx	Dif to Ref	Dif ²	Dif to Ref-Mean	Dif ²	Dif to 0.5	Dif ²	
0	->	31	1	0	1	0	0	0.587	0.344	0.5	0.25
1	->	56	1	16455	0.594	0.406	0.165	0.181	0.033	0.094	0.009
2	->	70	0.06	25671	0.333	-0.273	0.074	-0.081	0.007	-0.167	0.028
3	->	86	0.13	36202	0.081	0.049	0.002	-0.332	0.111	-0.419	0.176
4	->	95	0.13	42126	0.01	0.12	0.014	-0.403	0.162	-0.49	0.24
5	->	126	0.16	62532	0	0.16	0.026	-0.413	0.171	-0.5	0.25

Ref-Mean: 0.413

RMSE: 0.217

NRMSE: 0.013

PS: 0.66 (Good)

===== reference data Cindy=====

Entrenchment : off, Forget: immediate, other simplification: 0.0, initial state : 1.0

===== reference data =====

	Time	Ref-Idx	Tick	Sim-Idx	Dif to Ref	Dif ²	Dif to Ref-Mean	Dif ²	Dif to 0.5	Dif ²	
0	->	21	1	0	1	0	0	0.685	0.47	0.5	0.25
1	->	23	0.33	970	0.708	-0.378	0.143	0.394	0.155	0.208	0.043
2	->	28	0.33	3398	0.476	-0.146	0.021	0.161	0.026	-0.024	0.001
3	->	30	0.5	4369	0.43	0.07	0.005	0.115	0.013	-0.07	0.005
4	->	34	0.27	6311	0.308	-0.038	0.001	-0.007	0	-0.192	0.037
5	->	36	0.24	7282	0.237	0.003	0	-0.078	0.006	-0.263	0.069
6	->	41	0.16	9709	0.055	0.105	0.011	-0.26	0.067	-0.445	0.198
7	->	43	0.15	10680	0.017	0.133	0.018	-0.298	0.089	-0.483	0.234
8	->	48	0.22	13107	0.003	0.217	0.047	-0.312	0.097	-0.497	0.247
9	->	50	0.13	14078	0.001	0.129	0.017	-0.314	0.098	-0.499	0.249
10	->	55	0.13	16506	0	0.13	0.017	-0.314	0.099	-0.5	0.25

RMSE: 0.16

NRMSE: 0.007

PS: 0.75 (Good)

===== reference data Luigina =====

Entrenchment: off, forget: progressive + default, initial state 1.0, other simplification 0.0

===== reference data Luigina =====

	Time	Ref-Idx	Tick	Sim-Idx	Dif to Ref	Dif ²	Dif to Ref-Mean	Dif ²	Dif to 0.5	Dif ²	
0	->	49	1	0	1	0	0	0.492	0.242	0.5	0.25
1	->	53	0.91	1213	0.724	0.186	0.035	0.216	0.047	0.224	0.05
2	->	60	0.5	3336	0.471	0.029	0.001	-0.037	0.001	-0.029	0.001
3	->	64	0.13	4549	0.333	-0.203	0.041	-0.175	0.031	-0.167	0.028
4	->	85	0	10919	0.002	-0.002	0	-0.506	0.256	-0.498	0.248

RMSE: 0.124

NRMSE: 0.001

PS: 0.867 (Excellent)

===== reference data Concetta =====

Entrenchment: off, forget: progressive + default, initial state 0.17, other simplification 0.0

	Time	Ref-Idx	Tick	Sim-Idx	Diff	Diff ²
0	->	39	0.17	0	0.17	0.001
1	->	47	0	1745	0.009	-0.009
2	->	53	0	3055	0.001	-0.001
3	->	58	0	4146	0	0
4	->	62	0	5019	0	0

RMSE: 0.004

NRMSE: 0

PS: 1 (Excellent)

Enhanced incidental learning

Ilaria Borro

University of Portsmouth

This chapter will focus on theory and research into enhanced incidental learning (EIL), which is an approach to facilitating learning from input. The chapter will draw comparisons between EIL and input enhancement (IE). Input enhancement (Sharwood-Smith, 1981, 1993) consists of overt increases in the perceptual salience of target items, and is intended to improve noticing. The possibility that input enhancement may only result in improved explicit knowledge will be examined. In contrast, EIL does not entail changes to the input itself, but covert modifications of the conditions under which the input is processed (Long, 2017). There will be discussion of existing empirical results which suggest EIL may give rise to implicit knowledge by means of detection rather than noticing.

1. Introduction

The close psycholinguistic relation between language use for genuine communication on the one hand and L2 acquisition on the other is one of the guiding principles of Mike Long's work. This concept led to pedagogical proposals involving a limited amount of explicit instruction in order to keep learner focus on meaning and real information exchange. Focus on form in task-based language teaching is one of the main examples of such a technique, as it involves a brief, reactive, in context switch of attention from meaning to form often occurring as a function of communication needs (Long, 2015). In a truly communicative learning environment, focus on form calls for explicit instruction, with conscious noticing as one of its possible goals.

Mike Long's recent proposal, enhanced incidental learning, indicated a further step towards an even less explicit form of instruction, which aims to bypass conscious attention, instead bringing into play unconscious detection. In his own words, enhanced incidental learning conditions are designed to be "the least interventionist, but still effective, form of instruction" (Long, 2017, p. 36), in order to generate implicit knowledge. This chapter seeks to present the latest developments in enhanced incidental learning, including its pedagogical rationale (§ 2), psycholinguistic foundations (§ 3), along with the existing empirical studies of its potential (§ 4).

2. Enhanced incidental learning and input enhancement: pedagogical needs and definition

Language instructors must create the conditions that enable acquisition of the huge amount of knowledge and skills needed for foreign language use. However, there is often insufficient time to do so on a language course. It is therefore a core duty for SLA research to provide instructors with psycholinguistically sound and empirically grounded information on optimal use of the limited time available. The proposal of enhanced incidental learning (EIL) originates from such a line of reasoning, with the aim of contributing to the fulfilment of this duty.

In order to focus the needs of pedagogy, the section below opens with a definition of the aim of language instruction. The strengths and shortcomings of existing techniques are then discussed. Thereafter, the design and definition of enhanced incidental learning conditions are considered.

In the SLA literature of the last few decades there is broad agreement that the aim of language instruction is the acquisition of implicit rather than explicit knowledge of language (Long, 2017). Indeed, most of the various approaches to contemporary research on SLA agree on the difference of conscious and controlled knowledge to subconscious and automatic knowledge (Whong et al., 2014; VanPatten, Smith, & Benati, 2020). The former, explicit knowledge, is created intentionally by committing notions to memory, while acquisition of the latter, implicit knowledge, takes place unconsciously, i.e. without the subject being aware of the process. This distinction has been adopted and acknowledged since the 1980s, and is effectively defined by Paradis (2004, p. 47):

Explicit knowledge is qualitatively different from implicit competence. Explicit knowledge is conscious awareness of some data (utterances) and/or of their explicit analysis (structure). Implicit competence, on the other hand, is a set of computational procedures (of which the speaker is unaware) that generates sentences (which serve as data from which linguists or reflective speakers may construct a grammar – a set of rules – that becomes part of one’s explicit knowledge).

Desirable aspects of language use such as fluency, spontaneous speech and listening comprehension, operate as a function of implicit competence¹ (Bialystok, 1978; Long, 2017; Whong et al., 2014). Moreover, implicit knowledge is deeper, more durable, faster, and more efficient. Being automatic, it is less vulnerable to the effects of a concomitant memory load, fatigue, stress, or noise (Paradis, 2004). Therefore,

1. In the present contribution, the terms ‘implicit knowledge’ and ‘implicit competence’ are used interchangeably.

when a learner uses implicit knowledge of language, more attentional resources are freed for concentrating on content, the actual goal of communication.

Accordingly, the priority of implicit over explicit knowledge is widely (though not universally, see e.g., DeKeyser, 1994, 2015) acknowledged. Nevertheless, there is less agreement on how to acquire it. In particular, one of the most debated areas concerns the relative amounts of implicit and explicit learning conditions that learners should be exposed to.

A methodological and terminological note is necessary here. From a theoretical perspective, learning can be either explicit or implicit, i.e., it can take place with or without awareness of the process and of the linguistic features learned. However, it is hard if not impossible for researchers and instructors in a classroom context to directly determine the cognitive processes taking place in the learners' minds. Moreover, from an experimental perspective, measurement of the level of consciousness at the point of learning is a relatively new research area, which is still developing and refining its tools (e.g., Lee & Revesz, 2020; Rebuschat et al., 2015). Therefore, when researchers and instructors design pedagogical techniques, it is essential to take into account that "pedagogically-induced attention may or may not fit in with the learner attention" (Gass, 1999, p. 321). This caveat should always be considered when dealing with levels of consciousness in relation to learning.

That said, it is possible to design conditions which are more likely to result in either a conscious or an unconscious learning process i.e., respectively, intentional and incidental learning conditions. In intentional-learning conditions, learners are explicitly asked to learn the target linguistic structures, therefore the learning process takes place with awareness, and is likely to result in explicit knowledge. On the other hand, in incidental learning conditions, L2 speakers are involved in meaning-focused activities without an explicit request to learn formal aspects of the language. Therefore, learning is likely to take place unconsciously as a byproduct of communication and may result in the creation of implicit knowledge.

Considering the above-mentioned priority of implicit knowledge as a goal for language instruction, incidental learning conditions would appear to be preferable and commonly adopted. However, this is rarely the case. Even though the focus on content is highly beneficial for the development of language skills, incidental and implicit learning are extremely slow processes, and this is a serious issue when considering the limited time available in language courses. Moreover, learners are often required to perform form-focused, metalinguistic-oriented tests, where the communicative skills acquired through incidental learning conditions are not the most useful or appreciated. On the other hand, the intentional attention allocation to form involved in explicit instruction leads to faster and more effective learning when explicit, pencil-and-paper testing is concerned (§ 3). Nevertheless, two observations are necessary. First, the knowledge mainly measured in offline tests

is explicit in nature and concerns language forms, thus being of little if any use in real-life, real-time communication. Second, it is unviable to devote time to intentionally learning the huge amount of information needed to effectively use a language. By way of example, considering vocabulary, it has been reported that as many as 9000 word families are needed for a speaker to successfully deal with newspapers and novels (Nation, 2006). Learning such a number of words explicitly is unlikely to be an option.

The considerations discussed so far show that both purely incidental and purely intentional learning conditions have severe limitations. Therefore, there is a clear need for pedagogical techniques capable of overcoming them, so as to facilitate language acquisition. Put differently, the learning process needs to be sped up while being kept incidental, i.e., likely to result in more efficient creation of implicit knowledge. The aim is to raise and direct learners' attention, while at the same time keeping the level of consciousness below the level of awareness and involving the speakers in meaningful, communicative activities (for a deeper psycholinguistic reflection on these issues, see § 3).

An attempt in this direction has been made in the few last years by manipulating salience through input enhancement (IE). The following section describes how, despite promising premises, such a technique seems not to achieve the set goal.

Sharwood-Smith (1991) introduced the concept of input enhancement, a practice meant to raise learner's consciousness of the formal aspects of language through the manipulation of perceptual salience. Such a practice is seen as a resource the learner (especially the mature learner) can exploit to accelerate the process of learning how to communicate in the target language.

The term 'input enhancement' comprises a wide range of interventions, which can be negative (a form of corrective feedback) or positive and have various degrees of explicitness and elaboration. For the present contribution, it is relevant to focus on positive enhancement, which aims to emphasize correct forms so as to trigger conscious noticing and therefore a change in the knowledge the learner has of the structure. There are numerous ways to promote such emphasis.

Textual aspects such as frequency can be manipulated, artificially increasing the number of occurrences of a given structure or vocabulary item (input flood). On a different level, morphemes, words and phrases can be made typographically more salient through bolding, underlining, italicizing, changing the font type or size, etc. Aurally, salience can be increased by borrowing strategies naturally adopted by native speakers when dealing with language learners such as adding pauses before and after the target items, increasing volume, and slowing the reading pace. Finally, technology provides researchers and practitioners with new tools, such as interactive links to glosses and translations.

When considering the above-mentioned incidental/intentional learning condition dichotomy, IE cannot fit in either category. In a review of pedagogical approaches to formulaic language teaching, Pellicer-Sánchez and Boers (2019) created a third category specifically for input enhancement, namely *semi-incidental* learning conditions. Indeed, numerous studies involving attention-drawing techniques such as input enhancement still expect the subject to focus mainly on a text's meaning and content. This condition does not qualify as intentional, as no explicit instruction or request to learn the enhanced FSs is provided. At the same time, enhancement is meant to shift the subjects' attention from content to form, fostering noticing (Schmidt, 2010), therefore, neither is the 'incidental' category applicable.

Indeed, IE has been employed and experimented as a means to accelerate learning and proved to be somewhat effective (for a more detailed review of the empirical literature available, see § 4). However, the few experimental studies utilizing online assessment tools found that instruction involving IE resulted in the creation of explicit rather than implicit knowledge (Borro, 2021; Sonbul & Schmitt, 2013; Toomer & Elgort, 2019). In other words, despite the apparent potential for expediting the learning process while keeping it incidental, IE seems to trigger intentional learning instead, which is consistent with its declared aim of improving conscious noticing.

This is where the proposed enhanced incidental learning (EIL) comes into play. Long (2017, 2019; Jordan & Long, 2022) posits exposing learners to instructional conditions specifically designed to trigger unconscious detection rather than conscious noticing, i.e., genuine incidental learning. This technique manipulates the conditions under which input is processed rather than the input itself, in order to affect the cognitive processes involved and in all likelihood generate implicit knowledge. Therefore, enhanced incidental learning conditions are implemented by pursuing "the least interventionist, but still effective, forms of instruction" (Long, 2017, p. 36), of which the learners are supposed to remain largely unaware. In order to create such conditions, a number of pedagogical tools may be utilised. Fundamentally, enhanced incidental conditions amount to adding unobtrusive enhancement devices to purely incidental conditions. Such devices are mostly aural, and change whether the input is meant either for a whole group of students or occurs in real-time teacher-student interaction. In the first case, overall changes are mostly made to the way input is experienced by means of bimodal or multimodal presentation. For instance, learners can read a text while simultaneously listening to it (bimodal exposure), or they can view and listen to relevant tv contents, maybe with added captions (trimodal exposure). Further unobtrusive devices to foster even more detection can be added, inspired by research on natural native-to-nonnative communication strategies (foreigner talk discourse, FTD, Long, 1983). These include slower pace, one-beat pauses before and after critical

items, increased stress or volume, prosodically highlighted word grouping and elaboration through redundancy.

It should be noted that some of these devices have been employed in the IE framework (e.g., slower pace, pauses or increased volume for aural enhancement). However, the crucial difference is that even though EIL employs enhancement devices, the aim is never to force a switch of conscious attention from meaning to form; rather, devices are selected in order to have the lowest possible level of noticeability, with the goal of triggering and improving unconscious learning processes.

When it comes to redundancy, this technique is part of input elaboration, an unobtrusive pedagogical intervention consisting of modifications to the interactional structure of the conversation (Long, 2019). The advantage of this technique is that texts are made more comprehensible, and therefore processable, while at the same time preserving the elements and features of the language, which thus become available for incidental learning. Indeed, in both written and aural texts, complex grammatical construction and low-frequency vocabulary are not avoided, but made comprehensible by adding redundancy, repetitions, paraphrase, synonyms, more overt marking of grammatical and semantic relations and full NPs (noun phrases) in lieu of pronouns.

Input elaboration is consistent with the enhanced incidental learning framework as it is unobtrusive and aimed at boosting incidental learning while providing no explicit instruction. Accordingly, in teacher-student(s), real-time interaction, enhanced incidental learning conditions mainly entail negotiation for meaning, which results in a tailored, comprehensibility- and content-driven learning process.

This kind of modification is potentially capable of keeping the learners' focus on meaning, while at the same time enhancing the internal process of unconscious detection and therefore the creation of implicit knowledge. A deeper reflection of the psycholinguistic motivations and underpinnings of enhanced incidental learning compared to input enhancement is the focus of the next section.

3. Enhanced incidental learning and input enhancement: Psycholinguistic rationale and implications

This section accounts for the psycholinguistic rationale of enhanced incidental learning, in the context of notions such as the relative role of implicit and explicit instruction, the interface debate, the function of salience for acquisition and the noticing-detection debate.

The varying views on how languages are taught best ultimately reflect the relative value accorded to implicit and explicit learning, that is, their relative psycholinguistic role in second language acquisition.

Implicit learning can be defined as a process that, with no conscious learning effort, generates a tacit and abstract knowledge from a complex, rule-governed environment, knowledge which is representative of the structure of the environment and can be applied implicitly and generalized to novel circumstances (Reber, 1967, 1989). The emphasis on the lack of intention is confirmed in later definitions of implicit learning (e.g., Williams, 2005; Hulstijn, 2005), as well as the unconscious nature of the knowledge gained, which becomes apparent in changes in the behavioral responses of the subjects, without the subjects being aware of such changes (Godfroid, 2016).

Conversely, explicit learning is a conscious effort to learn notions about the language, ascertaining whether the input presents regularities and voluntarily working out the rules (Hulstijn, 2005), and is the most likely outcome of explicit instruction.

The efficacy of explicit instruction and learning has been investigated and demonstrated in numerous studies and meta-analyses (e.g., Goo et al., 2009; Norris & Ortega, 2000). The empirical literature shows that teaching the language explicitly results in significant improvement after a relatively short time and the effects of explicit instruction are more sizable when compared to implicit learning conditions. However, these results suffer from a measurement-related bias as most studies and meta-analyses only include the effects of explicit instruction on explicit knowledge, measured through pencil-and-paper, offline tests. Predictably, explicit instruction is highly beneficial when explicit knowledge of the taught structures is measured. However, comparatively few studies to date included online assessment in their design. Therefore, a highly relevant part of the picture, specifically the effects of instruction on the creation of implicit knowledge, is not considered.

Since the desired goal for instruction is implicit competence, findings about the creation of explicit knowledge through explicit instruction would support this kind of intervention only if the explicit knowledge gained was somehow capable of becoming implicit.

The existence of such a direct interface is subject to debate. On the one hand, scholars maintaining a strong-interface position claim that no implicit knowledge can be created without pre-existing explicit knowledge of the structure (e.g., DeKeyser, 2015). In contrast, the no-interface position assigns no role whatsoever to explicit instruction for the acquisition of a real-time, efficient use of the second language. Indeed, the existence of a direct interface between explicit and implicit knowledge is not supported by empirical evidence. Linguistic and neurological studies (Paradis, 1994) demonstrate a qualitative difference between implicit and explicit knowledge, which are created through different processes and stored in different areas of the brain. Therefore, it is impossible for the intentionally created explicit knowledge to become implicit. On the contrary, there is empirical evidence showing that explicit instruction can even have a detrimental effect on

the development of implicit knowledge and competences. Various studies have found that adding explicit instruction to a reading led to weaker content comprehension (Boers et al., 2017; Choi, 2017; Lee, 2007). In other words, the treatment undermined communication and the development of reading skills. Furthermore, Toomer and Elgort (2019) found that exposing learners to noticeable, typographical input enhancement (bolding the target items in a reading text, see *infra*, § 4) boosted the creation of explicit knowledge, but prevented the acquisition of implicit knowledge, which would have been gained without the explicit instructional intervention. The authors hypothesize that forcing the intentional attention allocation to the target items through explicit instruction changes the nature of the learning process and thus hampers its implicit, tacit, and unconscious nature.

These observations confirm that exposing learners to explicit instruction might not be the best solution when the aim is the creation of implicit knowledge. Indeed, there is broad agreement that involving learners in meaningful use of the language is either the only (Krashen, 1982) or the main (e.g., Bialystok, 1978; Ellis, 2005; Paradis, 2004) way of promoting implicit knowledge, because such conditions trigger the unconscious processes that create representations in implicit memory (Ellis, 1994, 2005).

Despite this, the dominant position in the literature does not exclude the utilization of a limited and controlled amount of explicit instruction, which is considered to have a beneficial role for language acquisition. According to the weak-interface position, a first stage of explicit instruction is necessary for some aspects of the language. The resulting explicit knowledge does not become implicit, but rather facilitates the statistical creation of implicit competence through frequency of production and reception. In other words, even though explicit instruction cannot have a direct role in the creation of implicit knowledge, explicit knowledge of the surface forms can be useful as a model for practice, which in turn creates implicit competence. While the learner practices the form under the guidance of explicit knowledge, implicit learning takes place because the underlying processes can establish themselves (Paradis, 2004).

The necessity of an initial, limited explicit intervention is due to some features of pure implicit learning rendering it potentially misleading and flawed. Implicit learning comprises priming routines or chunks of representations. This is an automatic process and is carried out by what Ellis calls 'zombie agents', i.e., processors that execute routines without and beyond our conscious control (Ellis, 2005). The processing strategies and L1 routines automatically driving these systems usually make the processing fast and efficient in the L1, but may actually hamper the correct processing of the input in the L2, e.g., ignoring relevant clues and/or diverting the subject's attention towards inappropriate features. Here the two core concepts of salience and blocking come into play.

Salience is one of the main factors that can affect learners' level of consciousness, therefore, it needs to be carefully considered when the aim is the creation of implicit knowledge. Salience lacks a clear definition in second language acquisition (Gass, Spinner & Behney, 2018), but when descriptions are attempted, they mainly deal with the perception of stimuli. Cho and Reinders (2013) report that "salience refers to the ease with which learners can perceive given input" (p. 134). Likewise, Wulff (2019) defines salience as the "general perceived strength of a stimulus" (p. 24). Similarly, VanPatten and Benati (2010) point to the "degree to which something catches a person's attention" (p. 143).

A number of classifications have been put forward in order to better define salience features and interactions with language perception and learning. The main distinction which is relevant for the present line of argument is that between perceptual and constructed salience. Perceptual salience is intrinsic in the features of a given linguistic form such as stress, accent, or any perceptual prominence capable of attracting attention. Ellis (2018) names this phenomenon psychophysical salience, indicating that it "arises in sensory data from contrasts between items and their context" (p. 21).

Constructed salience, on the other hand, takes place when a linguistic stimulus is intentionally made more prominent and noticeable by an external intervention. In instructed SLA, this practice is frequent in the case of linguistic constructions with so low a psychophysical salience that they can go undetected and unnoticed. For instance, crucial form-function relationships are often non-salient in the language stream. This is true across languages as it corresponds to the known least-effort principle (Zipf, 1949) according to which, the more frequently a form is used, the more often speakers abbreviate it in the attempt to minimize articulatory effort. This leads to the shortening of the most frequent words in languages, which consist of crucial items such as grammatical-function words and bound inflections. Such words become short and therefore non-salient, i.e., hard to perceive in the input and thus unlikely to be learnt without a pedagogic intervention.

Regarding blocking, experimental studies (e.g. Cintrón-Valentín & Ellis, 2015) demonstrate that in the presence of redundancy (e.g. a lexical and a morphological cue both indicating when an action takes place), the subjects extract the communicative content from the most salient cue (usually the lexical one), and this blocks their attention from the less salient (usually the morphological one), in the process preventing the noticing and learning of the less salient feature. The phenomenon of blocking is based on an associative process, which shifts attention as a result of previous experience: when the learner implicitly knows that a stimulus is associated with a certain outcome, it becomes harder to learn a different stimulus for the same outcome. The previous experience driving such an attentional bias can either be related to salience (as in Cintrón-Valentín & Ellis (2015) above-mentioned

example), processing strategies (Lee & Benati, 2007; VanPatten, 2004) or the L1. These phenomena clearly indicate that language teaching is necessary in order to direct the processing strategies of language learners (Benati, 2005).

In addition to the theoretical issues treated so far, one more aspect needs to be considered in regard to the relative roles of implicit and explicit learning of an L2. One of the main reasons the acquisition of a second language does not take place efficiently through implicit learning alone, as L1 acquisition does instead, is related to so-called critical periods. It has been demonstrated that children's capacity for instance learning (i.e., picking up form-meaning incidentally) and more generally for implicit learning begins to weaken around puberty, at the same time the capacity for explicit learning gets stronger (Granena, 2016; Granena & Long, 2013; Long, 2017). Although implicit learning does remain an option for adults, it is unlikely to be the sole means of successful and accurate acquisition of a second language after puberty.

To sum up, the intrinsic features of implicit learning seem to point to the necessity for a pedagogical intervention aimed at directing the learning process. Such interventions have mostly been designed to involve constructed salience and result in the conscious registration of stimuli.

Input enhancement is one such technique and its rationale is highly consistent with Ellis' (2001) Law of Contiguity, which describes language acquisition as a process that must start with conscious noticing of language structures (Schmidt & Frota, 1986; Schmidt, 1990, 2010), in order to "seed" (Ellis, 2005, p. 320) the new form. Once the construction is seeded in explicit memory, then implicit learning through statistical tallying can take place. At this point, to promote and speed up statistical learning, pedagogy should expose the subjects to exaggerated input, so that new implicit knowledge can be created below the level of awareness. Increasing the psychophysical salience of target forms by means of input enhancement (e.g., typographical enhancement) in a text is a direct implementation of this theory, as enhancement makes the stimulus more noticeable, therefore triggering the creation of explicit memory, while the text provides abundant, contextualized input.

While the need for noticing is the dominant view in the literature, there is debate on this issue, with scholars arguing for the possibility for learning to take place entirely below the level of awareness, i.e., at the detection level.

Tomlin and Villa (1994) defined detection as the cognitive registration of stimuli, which is necessary and sufficient for further processing and learning to take place. Crucially, detection takes place below the level of awareness, in contrast with conscious noticing. Empirical studies report mixed results, with some findings supporting the possibility for learning to take place at the detection level (e.g., Williams, 2005; Leung & Williams, 2011, 2012; Rebuschat & Williams, 2012), while others show the need for noticing (e.g., Hama & Leow, 2010). The debate is still

open, and since noticing seems to result mainly in explicit knowledge, pedagogical techniques avoiding conscious attention allocation are worth investigating.

In the framework of enhanced incidental learning it is hypothesized that unobtrusive enhancement devices added to incidental learning conditions can direct attention to non-salient aspects of the language while keeping the learning process at the level of detection. In other words, the purpose of enhanced incidental learning is to make unconscious, implicit learning more efficient by means of learning conditions triggering unconscious detection rather than conscious noticing.

In summation, from a psycholinguistic perspective, enhanced incidental learning takes account of the aim of language instruction (implicit knowledge) and the most likely way to reach it, i.e., incidental learning conditions. At the same time, flaws in implicit learning mechanisms are acknowledged, and the need for instruction to direct the process is met. In this context, enhanced incidental learning is based on the hypothesis of detection being both necessary and sufficient for learning, i.e., of the possibility for learning to take place entirely below the level of awareness.

Empirical findings about this possibility are mixed, rendering it necessary to gather more data. Studies investigating enhanced incidental learning provide insights on this topic as well, as the next section shows. This makes research of enhanced incidental learning relevant not only at a pedagogical level, but also for a deeper understanding of the psycholinguistic processes leading to language acquisition.

4. Enhanced incidental learning and input enhancement: Empirical evidence

4.1 Input enhancement

The effects of input enhancement on language learning have been investigated in numerous experiments. Most of them focused on the effectiveness of typographical enhancement for grammar learning and report mixed results, with some studies finding significant effects (e.g., Cintrón-Valentín & Ellis, 2015; Jahan & Kormos, 2015; Issa et al., 2015; LaBrozzi, 2016; Lee, 2007) while others report no benefits (e.g., Indrarathne & Kormos, 2016; Izumi, 2002; Loewen & Inceoglu, 2016; Winke, 2013).

Despite the mixed results on L2 development, empirical findings show more agreement when it comes to the effects of typographical enhancement on attention allocation. Three main tools have been employed in order to investigate how typographical enhancement affects attention, specifically notetaking (Izumi, 2002),

think-aloud protocols (Bowles, 2003; Leow, 2001; Leow et al., 2003) and eye tracking (Cintron-Valentin & Ellis, 2015; Indrarathne & Kormos, 2016; Issa et al., 2015; Loewen & Inceoglu, 2016; Simard & Foucambert, 2013; Winke, 2013). The majority of these studies confirmed that typographical enhancement increased the amount of attention going to the target items, often triggering intentional learning.

Regarding vocabulary, studies present a higher degree of agreement on the significant and positive impact of enhancement in written input (e.g., Aloitabi, 2020; Boers et al., 2017; Peters, 2012; Szudarski & Carter, 2016). It should be noted that all the mentioned studies accounted for explicit knowledge gains only, as the posttests were not designed to record implicit competence.

Three studies to date employed not only offline but also online measurements, aiming to assess gains in both explicit and implicit knowledge and these merit a more extensive consideration. Sonbul and Schmitt (2013) employed a counterbalanced, within-subject design to compare three experimental conditions for collocation learning, namely enriched (input flood: three occurrences in a short passage), enhanced (same as enriched condition, with bold and red font added), and decontextualized (collocations presented individually on slides to be memorized). Forty-two subjects carried out two offline tests to assess receptive and productive explicit knowledge, while implicit knowledge was measured by means of priming. The test battery was repeated immediately after the treatment and then again two weeks later. Results of the offline tests showed that all three experimental conditions led to significant and durable learning at both productive and receptive levels. However, regarding implicit knowledge, both immediate and delayed priming sessions resulted in no measurable gains for any of the instruction conditions.

Toomer and Elgort (2019) ran a conceptual replication and extension of Sonbul and Schmitt's (2013) study. They applied the same within-subject counterbalanced design to a larger sample, as they exposed their 62 participants to three learning conditions, specifically reading only (no typographical enhancement), bolding, and bolding-plus-glossing. Compared to Sonbul and Schmitt (2013), the treatment was longer and spaced, and target collocations occurred a total of nine times. These choices were meant to boost statistical learning and implicit knowledge gains. As in the previous study, testing sessions took place immediately after the treatment and then again two weeks later, and included both offline and online measures. Explicit knowledge was measured by offline recall and recognition tests, while lexically-primed decision was used to detect the creation of implicit knowledge. Analysis confirmed Sonbul and Schmitt's outcomes as the experimental groups developed significant explicit knowledge of the target collocations, with the enhanced condition more effective than the reading-only condition. Findings on implicit knowledge extend and partially confirm those of the previous study. No priming effect was shown as a result of typographical enhancement, while implicit

knowledge emerged for collocations in the reading-only condition. The fact that learners were able to develop implicit knowledge in this study and not in Sonbul and Schmitt's (2013) can be explained by the longer, spaced nature of the treatment, as well as the increased frequency of the target items. The lack of implicit knowledge as a result of the enhanced condition, however, raises a more complex issue. A possible interpretation is that typographical enhancement made the target collocations more salient (as the explicit knowledge gains demonstrate), as a result triggering conscious noticing and intentional learning and thus explicit knowledge. Conversely, the reading-only condition did not interfere with the unconscious process of implicit learning, and could therefore result in the creation of implicit knowledge. In order for this interpretation to be confirmed, a finer observation of the learning process is necessary and this can be gained through online measures such as eye tracking and a consciousness assessment with verbal protocols. The third study addressing both explicit and implicit knowledge gains with input enhancement employs such tools and supports the hypothesis, as reported at the end of this section (Borro, 2021).

The fact that input enhancement mainly results in explicit knowledge confirms the need for a less obtrusive pedagogical technique, specifically designed to trigger detection and the creation of implicit knowledge. Enhanced incidental learning conditions constitute such a technique.

4.2 Enhanced incidental learning

Few studies to date have focused explicitly on enhanced incidental learning. Much of the empirical evidence of its effectiveness comes from studies investigating the techniques utilized in enhanced incidental learning such as bimodal and multimodal exposure and input elaboration through redundancy. As in the input enhancement literature, studies addressing implicit knowledge gains are unfortunately rare.

Input elaboration has proved beneficial for language learning. In order to investigate its effectiveness, empirical studies have compared it to both simplified and genuine (i.e., authentic, non-elaborated and non-simplified) texts.

Urano (2000) compared the effects of exposure to genuine, simplified, and elaborate texts on 40 Japanese learners of English. Outcomes from meaning- and form-recognition tests showed that elaborated input triggered incidental vocabulary learning, while simplification did not.

With a more complex design, Chung (1995) exposed 484 learners of L2 English to nine passages in one of five versions: genuine, simplified, lexically elaborated, structurally elaborated and both lexically and structurally elaborated. Outcomes from both immediate and delayed form- and meaning-recognition tests showed subjects exposed to elaborated input outperformed the genuine-text group.

A similar pattern emerged from a study by Kim (2006), who compared elaboration conditions to typographical enhancement and a genuine, baseline text. The between-subject design involved 297 EFL learners and showed the elaboration to be significantly more effective than the baseline conditions. At the same time, no difference was found between elaboration and exposure to the more explicit typographical enhancement, which supports the effectiveness of unobtrusive devices.

Elaborated text results in larger vocabulary gains than non-modified input also in O'Donnell's (2009) study, where she exposed a 197-participant sample to literary texts in one of the two conditions.

Godfroid and colleagues (2013) contributed to this line of research with eye-tracking data about attention allocation, collected from a sample of 28 learners of English as a second language. Comparing a genuine and an elaborated version of a reading text, they found elaboration to be effective in drawing the participant's attention to the target vocabulary items. Unfortunately, it is not possible to state whether this attention was conscious or unconscious, since the design did not include any verbal protocol.

Fewer studies have investigated the effects of elaboration on auditory input, and they present mixed results: while Loschky (1994) found no advantage of elaborated text over the unmodified baseline, Toya's (1992) outcomes from a vocabulary posttest show that the elaborated-text group outperformed subjects exposed to genuine input. There were no further investigations on the effects of modification on auditory input until a recent work by Kobayashi Hillman (2020). Vocabulary posttests performed by her 106 L1-Chinese speakers of Japanese as an L2 showed auditory elaborated input to be advantageous compared to both the simplified and the unmodified versions.

Some of the mentioned studies, *inter alia*, investigated the effects of elaboration on comprehension, confirming its benefits in this respect as well (Chung, 1995; Kobayashi Hillman, 2020; O'Donnell, 2009; Oh, 2001; Yano et al., 1994). This constitutes an important addition, as comprehension is required for learning to take place (see Lichtman & VanPatten, 2021 for a review about this statement from Krashen on).

Summing up, the empirical literature on input elaboration confirms two main points underpinning the EIL framework. First, the need for instructional intervention directing and improving incidental learning from reading or listening, which is less effective in the case of natural, unmodified input. Second, the fact that unobtrusive devices such as those employed in elaboration are indeed effective, while not involving any explicit intervention.

A different technique constituting one of the core pedagogical indications of the EIL hypothesis is the employment of bimodal and multimodal exposure, which has been increasingly investigated in the last few years.

Bimodal exposure, i.e., reading-while-listening, is beneficial for incidental learning in many respects. First, the aural component provides a superimposed pace, which discourages the learners from stopping and intentionally learning unknown words or phrases. In Horst and colleagues' (1998) words, "reading aloud created the circumstances for incidental acquisition by precluding opportunities for intentional word learning" (p. 211).

Another advantage of reading while listening concerns content comprehension. The availability of prosody and sound-symbol correspondence provides learners with an additional tool for text comprehension (Chang, 2009; Chang & Millet, 2015; Tekmen & Daloglu, 2006). In an eye-tracking study, Conklin and colleagues (2020) investigated the alignment between gaze and audio, i.e., whether subjects fixated on a word at the same time they heard it. Statistical analysis showed that this alignment was scarce in general, but significantly better for subjects with a lower vocabulary size. This finding can confirm the notion that audio is an important resource for L2 readers experiencing comprehension problems. In turn, a better comprehension of the context provides increased cues and attentional resources for inferring the meaning of unknown items. As Malone (2018) pointed out, listening while reading allows the learners to make the cognitive resources needed for phonological decoding available for form and meaning processing instead. A behavioral confirmation for the psycholinguistic effect of bimodal exposure emerges in studies that investigated learners' preferences among exposure modality (e.g., Brown et al., 2008; Chang, 2009). Indeed, students claimed to be more comfortable in the reading-while-listening mode as compared to reading only and listening only.

With a different and complementary perspective, Barclay (2021) investigated the effects of different modes on cognitive burden and knowledge decay, finding bimodal exposure advantageous in this respect as well. The advantage of these beneficial features is demonstrated in the empirical studies with exposure modality as an independent variable, i.e., comparing reading while listening with reading only or listening only.

Malone (2018) rigorously designed experimental procedures meant to ensure incidental conditions when addressing learning of vocabulary through reading-only and reading-while-listening. Eighty participants read four graded readers on timed slides on a computer screen, with or without the aural component. Form-recognition and form-meaning connection tests showed that reading while listening was significantly more beneficial for incidental vocabulary learning. Teng (2018) confirmed these results in a study assessing four dimensions of vocabulary knowledge: form recognition, grammar recognition, meaning recall and collocation recognition. Outcomes from 60 experimental subjects showed that reading while listening led to larger knowledge gains than reading only for all four dimensions investigated.

Webb and Chang (2012) focused on the effect of repeated reading, as their 82 subjects read or read and listened to a short text several times over two seven-week periods. Dependent measures were provided by modified vocabulary knowledge scales, which were applied to a pretest-posttest design. Outcomes showed bimodal exposure to be significantly more effective than reading only for vocabulary knowledge gains.

Webb and Chang (2020) also designed the only study comparing the effects of reading-only, reading- while-listening and listening-only on incidental learning of formulaic sequences. Their 112 experimental subjects were randomly divided into three experimental groups (reading only, listening only and reading while listening) and a control group. Results from offline immediate and delayed posttests demonstrated an advantage for bimodal exposure over both reading-only and listening-only modalities.

A study by Godfroid (2016) addressed the effects of implicit instruction, i.e., bimodal exposure, on implicit grammar knowledge. Thirty-eight L2 learners were exposed to reading while listening to examples of a German grammar structure without explicit instruction, while an additional task (sentence-picture matching) kept them focused on meaning. Their reaction times and performance in a word-monitoring posttest demonstrated implicit knowledge gains and interviews revealed that 33 out of 38 participants remained unaware of the process. These outcomes confirm the potential of learning below the level of awareness, thus supporting EIL's psycholinguistic rationale.

The benefits of reading-while-listening are confirmed for multimodal exposure, i.e., viewing, listening, and reading, as in captioned videos, although no study employing online measurement tools is available. This positive impact is believed to be due to captioning breaking down speech into words, and therefore favoring word recognition, which is necessary for comprehension and for identification of novel lexical items. Moreover, this frees attentional resources which can be employed for grammar learning (Lee & Revesz, 2018). Indeed, a meta-analysis including 18 studies (Montero Perez et al., 2013) proved captioning to be beneficial for language learning at both the vocabulary and grammar level.

These advantages are demonstrated by studies comparing the effects of bimodal (non-captioned videos) and multimodal (captioned videos) on learning. Cintron-Valentin and colleagues (2019) investigated the effects of captioned vs. non-captioned videos on both vocabulary and grammar learning. Vocabulary posttests outcomes show a clear advantage for captions, while only some of the targeted grammar structures benefitted from such treatment. Lee and Revesz (2020) carried out a similar study, exposing 72 Korean students learning English as a second language to either captioned or non-captioned videos. Results from

grammar posttests showed that captions were beneficial compared to non-captioned treatment.

The empirical research reported so far shows the potential of the tools EIL employs, but it is not specifically motivated by the EIL hypothesis. The only complete study to date expressly aimed at demonstrating the effectiveness of enhanced incidental learning is by Borro (2021), who compares it to input enhancement.

The study investigates the acquisition of L2 idiomatic expressions, involving 83 Chinese learners of Italian L2. The experiment measured gains in both explicit and implicit knowledge, respectively through pencil and paper tests and a self-paced reading test, which were repeated immediately after the treatment and again three weeks later. Moreover, levels of attention and awareness were assessed by means of eye tracking and retrospective verbal reports. No pretest was performed so as not to hamper the incidental nature of learning. With a between-subject design, participants were randomly assigned to four experimental groups and a control group, controlling for proficiency and working memory. All the participants were exposed to the same reading-while-listening graded reader, where the ten target items were embedded with the artificially increased frequency of seven occurrences each. The enhanced-incidental-learning condition was operationalized as the addition of unobtrusive aural enhancement (a one-beat pause before and after each target item) of the first two occurrences of the target items. The other experimental conditions included typographical enhancement, a combination of aural and typographical enhancement, and the plain text with no added enhancement (increased frequency only).

The study findings show that pedagogical treatments involving typographical enhancement resulted in intentional learning and the creation of explicit knowledge. Conversely, the EIL condition led to the same additional allocation of attention to the target items, which in this case, crucially took place below the level of awareness. Such attention allocation was absent in subjects reading the plain text with no enhancement. Even though results from the implicit knowledge test at the product level were inconclusive, eye tracking measures showed a growing familiarity with the target idioms throughout the treatment for all subjects exposed to enhancement devices, but not for the increased-frequency-only group. This means that learning took place consciously as a result of the instructional intervention, with typographical enhancement and unconsciously in EIL conditions. These findings are relevant both psycholinguistically and pedagogically. As for the former, the study reports the possibility for learning to take place below the level of awareness, adding to the noticing-detection debate. From a pedagogical angle, instructors can benefit from two main findings: first, typographical enhancement results in explicit knowledge, as already shown in previous research. Second, EIL conditions

involving unobtrusive enhancement seem capable of triggering more learning than reading-while-listening alone, all the while keeping the level of consciousness below the awareness threshold and making implicit knowledge gains more likely.

5. Conclusion and directions for future research

The present contribution sought to illustrate the theory and practice of enhanced incidental learning compared to a widely spread pedagogical technique with an apparently similar rationale, i.e., input enhancement.

Acknowledging implicit knowledge as the main aim for language instruction implies the need to develop and validate pedagogical techniques capable of boosting its acquisition. Explicit instruction proves beneficial for explicit rather than implicit knowledge gains, while purely implicit learning is exceedingly slow and likely flawed. Therefore, a limited amount of instruction is needed in order to speed up and direct the learning process while keeping it as implicit as possible. Input enhancement constitutes a possibility in this regard, as it aims to increase the salience of targeted items and trigger noticing without involving explicit instruction. Despite these promising premises, recent empirical data show that IE results mainly in explicit rather than implicit knowledge.

EIL is thus proposed as an even less obtrusive form of instruction, as it is consistent with the hypothesis of learning taking place at the detection level, i.e., below the awareness threshold, and resulting in implicit knowledge. This hypothesis is confirmed in the most comprehensive study to date specifically addressing EIL, which on the one hand showed the possibility for learning to take place unconsciously in EIL conditions and on the other confirmed explicit knowledge as the main outcome of IE.

More empirical data is needed on this promising subject. Future experimental studies may analyse pedagogical tools aimed at setting EIL conditions, such as bimodal and multimodal exposure, unobtrusive aural enhancement and redundancy. In order to collect evidence of learning at the detection level, the effects of these techniques need to be assessed at both the process and the product level. Specifically, at the process level, measures of attention such as eye tracking and pupillometry should be adopted and their outcomes triangulated with verbal reports about consciousness and measures of knowledge gains. Moreover, posttests need to include the assessment of implicit knowledge, acknowledging the directions and standards the methodological literature is currently proposing (e.g., Godfroid & Kim, 2021).

References

- Alotaibi, S. (2020). The effect of input mode, input enhancement and number of exposures on the learning and processing of binomials in the L2 (Unpublished doctoral dissertation). University of Nottingham.
- Barclay, S. (2021). *The role of form presentation mode and learning aptitude in foreign language vocabulary learning and decay*. Presentation at the British Association of Applied Linguistics Vocabulary Special Interest Group Conference. University of Sheffield, 27 May 2021.
- Benati, A. G. (2005). The effects of processing instruction, traditional instruction and meaning-output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9(1), 67–93. <https://doi.org/10.1191/1362168805lr1540a>
- Bialystock, E. (1978). A theoretical model of second language learning. *Language Learning*, 28(1), 69–83. <https://doi.org/10.1111/j.1467-1770.1978.tb00305.x>
- Boers, F., Demecheleer, M., He, L., Deconinck, H. & Eyckmans, J. (2017). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*, 27(2), 448–469. <https://doi.org/10.1111/ijal.12141>
- Borro, I. (2021). Enhanced incidental learning of formulaic sequences by Chinese learners of Italian L2 (Unpublished doctoral dissertation). University of Portsmouth.
- Bowles, M. (2003). The effects of textual enhancement on language learning. An online/offline study of fourth semester students. In P. Kempchinsky & C. E. Pineros (Eds.), *Theory, practice and acquisition: Papers from the 6th Hispanic Linguistics Symposium and the 5th Conference on the Acquisition of Spanish and Portuguese* (pp. 395–411). Cascadilla Press.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136–163.
- Chang, A. (2009). Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories. *System*, 37(4), 652–663. <https://doi.org/10.1016/j.system.2009.09.009>
- Chang, C.-S., & Millett, S. (2015). Improving reading rates and comprehension through audio-assisted extensive reading for beginner learners. *System*, 52, 91–104. <https://doi.org/10.1016/j.system.2015.05.003>
- Cho, M., & Reinders, H. (2013). The effects of aural input enhancement on L2 acquisition. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 133–148). University of Hawai'i, National Foreign Language Resource Center.
- Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. *Language Teaching Research*, 21(3), 403–426. <https://doi.org/10.1177/1362168816653271>
- Chung, H. (1995). Effects of elaborative modifications on second language reading comprehension and incidental vocabulary learning. *University of Hawai'i Working Papers in ESL*, 14, 27–61.
- Cintrón-Valentín, M., & Ellis, N. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Language Learning*, 37(2), 197–235.
- Conklin, K., Alotabi, S., Pellicer-Sánchez, A., & Vilkaitė-Lozdiene, L. (2020). What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second Language Research*, 36(3), 257–276. <https://doi.org/10.1177/0267658320921496>

- DeKeyser, R. M. (1994). How implicit can adult second language learning be? In J. Hulstijn & R. Schmidt (Eds.), *Consciousness in second language learning* (pp. 83–96). *AILA Review*, 11, 83–96.
- DeKeyser, R. M. (2015). Skill Acquisition Theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 94–112). Routledge.
- Ellis, N. (1994). *Implicit and explicit learning of languages*. Academic Press
- Ellis, N. (2001). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 2–66). Blackwell.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.
<https://doi.org/10.1017/S027226310505014X>
- Ellis, N. (2018). Saliency in usage-based SLA. In S. Gass, P. Spinner & J. Behney (2018). *Saliency in second language acquisition*. New York, NY: Routledge, pp. 21–40.
- Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2), 319–333. <https://doi.org/10.1017/S0272263199002090>
- Gass, S., Spinner, P., & Behney, J. (2018). *Saliency in second language acquisition*. Routledge.
- Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition*, 38(2), 177–215.
<https://doi.org/10.1017/S0272263115000388>
- Godfroid, A., & Kim, K. (2021). The contributions of implicit-statistical learning aptitude to implicit second-language knowledge. *Studies in Second Language Acquisition*, 43(3), 606–634.
<https://doi.org/10.1017/S0272263121000085>
- Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, 35(3), 483–517. <https://doi.org/10.1017/S0272263113000119>
- Goo, J., Granena, G., Novella, M., & Yilmaz, Y. (2009). *Implicit and explicit instruction in L2 learning: Norris and Ortega (2000) revisited and updated*. Paper presented at the Second Language Research Forum, Michigan State University.
- Granena, G. (2016). Age differences, maturational constraints, and implicit and explicit L2 learning. *Brill Research Perspectives on Multilingualism and Second Language Acquisition*, 1(1), 9–44.
- Granena, G., & Long, M. H. (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. John Benjamins. <https://doi.org/10.1075/llt.35>
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207–223.
- Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. Introduction. *Studies in Second Language Acquisition*, 27(2), 129–140.
<https://doi.org/10.1017/S0272263105050084>
- Indrarathne, B., & Kormos, J. (2016). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition*, 39(3), 401–430.
<https://doi.org/10.1017/S027226311600019X>
- Issa, B., Morgan-Short, K., Villegas, B., & Raney, G. (2015). An eye-tracking study on the role of attention and its relationship with motivation. *Eurosla Yearbook*, 15(1), 114–142.
<https://doi.org/10.1075/eurosla.15.05jss>
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541–577.
<https://doi.org/10.1017/S0272263102004023>

- Jahan, A., & Kormos, J. (2015). The impact of textual enhancement on EFL learners' grammatical awareness of future plans and intentions. *International Journal of Applied Linguistics*, 25(1), 47–66. <https://doi.org/10.1111/ijal.12049>
- Jordan, G. & Long, M. (2022). *English language teaching now and how it could be*. Cambridge Scholars Press.
- Kim, Y. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *TESOL Quarterly*, 40(2), 341–373. <https://doi.org/10.2307/40264526>
- Kobayashi Hillman, K. (2020). Effects of different types of auditory input on incidental vocabulary learning by L2 Japanese learners (Unpublished Ph.D. dissertation). College Park, MD: University of Maryland.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon.
- LaBrozzi, R. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, 20(1), 75–91. <https://doi.org/10.1177/1362168814561903>
- Lee, J., & Benati, A. G. (2007). *Delivering processing instruction in classrooms and virtual contexts: Research and practice*. Equinox.
- Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal*, 102(3), 557–577. <https://doi.org/10.1111/modl.12503>
- Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, 42(3), 625–651. <https://doi.org/10.1017/S0272263120000108>
- Lee, S. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, 57(1), 87–118. <https://doi.org/10.1111/j.1467-9922.2007.00400.x>
- Leow, R. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania*, 84, 496–509. <https://doi.org/10.2307/3657810>
- Leow, R., Grey, S., Matijuan, S. & Moorman, C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning*, 13, 93–108.
- Lichtman, K., & VanPatten, B. (2021). Was Krashen right? Forty years later. *Foreign Language Annals*, 54(2), 283–305. <https://doi.org/10.1111/flan.12552>
- Loewen, S., & Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. *Studies in Second Language Learning and Teaching*, 6(1), 89–110. <https://doi.org/10.14746/ssllt.2016.6.1.5>
- Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5(2), 177–193. <https://doi.org/10.1017/S0272263100004848>
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Oxford: Wiley-Blackwell.
- Long, M. H. (2017). Geopolitics, methodological issues, and some major research questions. *ISLA*, 1(1), 7–44. <https://doi.org/10.1558/isla.33314>
- Long, M. H. (2019). Optimal input for language learning: Genuine, simplified, elaborated, or modified elaborated? *Language Teaching*, 53(2), 169–182. <https://doi.org/10.1017/S0261444819000466>

- Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Second Language Acquisition*, 16(3), 303–323.
<https://doi.org/10.1017/S0272263100013103>
- Malone, J. (2018). Incidental vocabulary learning in SLA. Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, 40(3), 651–675.
<https://doi.org/10.1017/S0272263117000341>
- Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720–739.
<https://doi.org/10.1016/j.system.2013.07.013>
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528. <https://doi.org/10.1111/0023-8333.00136>
- O'Donnell, M. E. (2009). Finding middle ground in second language reading: Pedagogic modifications that increase comprehensibility and vocabulary acquisition while preserving authentic text features. *The Modern Language Journal*, 93(4), 512–533.
<https://doi.org/10.1111/j.1540-4781.2009.00928.x>
- Oh, S. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, 35(1), 69–96. <https://doi.org/10.2307/3587860>
- Paradis, M. (2004). *Neurolinguistic theory of bilingualism*. John Benjamins.
<https://doi.org/10.1075/sibil.18>
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit language learning* (pp. 393–419). Academic Press.
- Pellicer-Sánchez, A., & Boers, F. (2019). Pedagogical approaches to the teaching and learning of formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.), *Understanding formulaic language: A second language perspective* (pp. 153–173). Routledge.
- Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *Language Learning Journal*, 40(1), 65–79.
<https://doi.org/10.1080/09571736.2012.658224>
- Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology*, 118(3), 219–235. <https://doi.org/10.1037/0096-3445.118.3.219>
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 317–327. [https://doi.org/10.1016/S0022-5371\(67\)80149-X](https://doi.org/10.1016/S0022-5371(67)80149-X)
- Rebuschat, P., Hamrick, P., Sachs, R., Ziegler, N., & Riestenberg, K. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, 37, 299–334.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Schmidt, R. W. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of Classic 2010, Singapore, December 2–4*. Singapore (pp. 721–737). National University of Singapore, Centre for Language Studies.
- Schmidt, R., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Newbury House.

- Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 2, 159–168. <https://doi.org/10.1093/applin/2.2.159>
- Sharwood-Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research*, 7(2), 118–132.
- Sharwood-Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15(2), 165–179. <https://doi.org/10.1017/S0272263100011943>
- Simard, D., & Foucambert, D. (2013). Observing noticing while reading in L2. In J. M. Bergleithner, S. N. Frota & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 207–226). National Foreign Language Resource Center, University of Hawaii at Manoa.
- Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, 63(1), 121–159. <https://doi.org/10.1111/j.1467-9922.2012.00730.x>
- Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, 26(2), 245–265. <https://doi.org/10.1111/ijal.12092>
- Tekmen, E., & Daloglu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals*, 39(2), 220–243. <https://doi.org/10.1111/j.1944-9720.2006.tb02263.x>
- Teng, F. (2018). Incidental vocabulary acquisition from reading-only and reading-while-listening: A multi-dimensional approach. *Innovation in Language Learning and Teaching*, 12(3), 274–288. <https://doi.org/10.1080/17501229.2016.1203328>
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and SLA. *Studies in Second Language Acquisition*, 16(2), 183–204. <https://doi.org/10.1017/S0272263100012870>
- Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, 69(2), 405–439. <https://doi.org/10.1111/lang.12335>
- Toya, M. (1992). Form of explanation in modification of listening input in L2 vocabulary learning (Unpublished doctoral dissertation). University of Hawai'i, Honolulu.
- Urano, K. (2000). Lexical simplification and elaboration: Sentence comprehension and incidental vocabulary acquisition (Unpublished doctoral dissertation). University of Hawai'i, Honolulu.
- VanPatten, B. (Ed.). (2004). *Processing Instruction*. Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410610195>
- VanPatten, B., & Benati, A. G. (2010). *Key terms in second language acquisition*. Continuum.
- VanPatten, B., Smith, M., & Benati, A. G. (2020). *Key questions in second language acquisition*. Cambridge University Press.
- Webb, S., & Chang, A. (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition*, 1–22. <https://doi.org/10.1017/S0272263120000297>
- Webb, S., & Chang, C.-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *The Canadian Modern Language Review*, 68, 276–290. <https://doi.org/10.3138/cmlr.1204.1>
- Whong, M., Gil, H.-G., & Marsden, E. (2014). Beyond paradigm: The 'what' and the 'how' of classroom research. *Second Language Research*, 30(4), 551–568. <https://doi.org/10.1177/0267658314526749>

- Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, 27(2), 269–304. <https://doi.org/10.1017/S0272263105050138>
- Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee, 2007, with eye-movement data. *Studies in Second Language Acquisition*, 35(2), 323–352. <https://doi.org/10.1017/S0272263112000903>
- Wulff, S. (2019). Acquisition of formulaic language from a usage-based perspective. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.), *Understanding formulaic language: A second language perspective* (pp.19–37). Routledge.
- Yano, Y., Long, M., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, 44, 189–219. <https://doi.org/10.1111/j.1467-1770.1994.tb01100.x>
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press.

The uncertainty principle in second language acquisition

Stefano Rastelli
Università di Pavia

This chapter presents three themes that I have discussed with Mike Long on different occasions in recent years: The Discontinuity Hypothesis, the necessity to study the ‘intra-language’ (in addition to the ‘interlanguage’) and the uncertainty principle. The latter is the idea that abstract rules and statistical regularities are entangled states of mind in a L2 learner’s competence, to the extent that researchers can never know whether learners are recognizing a morphosyntactic form (due to its frequency) or they are generating that form by a rule (due to abstract ‘labels that predate the input’). Uncertainty implies that what we can uncover about a learner’s implicit competence depends on how such competence is measured.

1. Introduction: Mike and science

On more than one occasion I happened to hear Mike Long questioning the scientific rigor of SLA research in general (of course his and mine included). This can surprise only those who never had the chance to discuss with Mike in relaxed environment mood (e.g., at dinner) about the advances in the sciences of nature. Mike was amazed by the subjects his son was engaging with at college, even at an introductory level: bio-engineering and bio-mechanics, brain and molecular sciences, particle physics, genetics, and so forth. His admiration – and subsequent exclamation “that’s real science!” – was neither due to an inferiority complex nor to science envy (see Gregg et al., 1997, p. 543). That enthusiasm came from Mike’s Popperian rationalism, the very belief that human knowledge – including that of how languages are acquired – can advance only if the hypotheses are tested under falsifiable (explicit and replicable) conditions. Rationalism in SLA boils down to three points: (i) non-testability is the trademark of ‘nonempirical ideas’;¹ (ii) even

1. I borrowed the term from Hulstijn (2013, p. 515): “All scientific disciplines must create room for ideas or theories that do not yet lend themselves to empirical testing, but for a discipline to

remarkable individual differences in second language (L2) learning can *never* be taken as counterevidence to generalizability of findings – as Gregg et al. (1997, p. 545) put it, “we don’t try to build a theory [...] of the acquisition of French by the butcher’s daughter” (on this topic also see Jordan, 2004; Pallotti, 2021); (iii) considering complexity in SLA does not defy the seek for abstract principles, because “idealization [...] is a misleading term for the only reasonable way to approach a grasp of reality” (Chomsky, 1995, p. 7). These points represent exactly what SLA researchers have learned from the sciences of nature, since the dawn of the discipline. Hulstijn (2013, p. 516) reports that Long (1993) had counted up to between 40 and 60 theories of SLA, and that he had found this count worrisome. Probably Mike was not concerned with the number per se, but with the lack of predictive power of such theories. Even now that alternatives to rational enquiry have become popular and perhaps predominant in our field, Mike always remained firm in the view that SLA finds its place within language science, “which is in turn part of cognitive science” (Long, 2015, p. 572), and that future SLA students should be systematically trained in experimental methods and techniques.²

2. Rationalism continued

Rationalism and rational inquiry should not be necessarily anchored to SLA ideas of the past, as new ones too can be proven true or false. Long (2016) welcomed three recent ideas on L2 development which – in Mike’s words – could challenge ‘received wisdom’ in SLA:

1. L2 development might not correspond to a unique continuous function, e.g., a process leading to *only* one change of state in a learner’s mental knowledge over time. Unlike many theories that depict SLA as the single learning trajectory going from ‘not-knowing’ to ‘knowing’ (regardless of whether or not such

develop fruitfully it is crucial that nonempirical ideas do not outnumber empirical ones [...] If an academic discipline is characterized by too many nonempirical ideas and too few empirical ideas, it runs the risk of losing credit in the scientific community at large (and in society).”

2. “For that, they [SLA students] need a solid grounding in the field’s core knowledge bases – breadth, as well as depth – and for the same reason, a reasonable level of rigorous training in qualitative and quantitative research methods. For SLA and many areas of applied linguistics, the research methods component should include a course dealing specifically with L2, not just general research methods. These include the creation, validation, and use of measures of elicited imitation, grammaticality judgments, reaction-times, eye-tracking, language accuracy, language complexity, language acquisition, and language use” (Long, 2015, p. 571).

- trajectory is linear, monofactorial etc.), the Discontinuity Hypothesis depicts SLA as a process going from not-knowing to knowing twice, statistically and grammatically. These are two states of knowledge – one based on frequency, the other based on frequency-independent, abstract computations – that are entangled and superposed in a learner’s competence (Rastelli, 2014, 2019, 2021);
2. Future experimental research on L2 development should supplement the traditional analysis of learners’ interlanguages (where the focus is on ‘between-systems’ variance § 4) with the analysis of learners’ *intra-language* (where the focus is on within-item variance). The latter is the study of how perfectly target-like forms are generated (by rule or by rote). Indeed, the convergence of interlanguages to the norms of the target language is not the end of the story. L2 development may continue even when the forms produced by learners seem indistinguishable from the forms produced by native speakers (Rastelli, 2020, 2021, forthcoming; Rastelli & Murakami, 2022);
 3. L2 learners’ mastery of even the deepest aspects of syntax and morphology might not be ‘once and for all’. Current theories of L2 competence – whether generative or usage-based – should live up to the challenge posed by psychological and neurobiological theories of rehearsal (i.e., performance science and the science of voluntary movements). In human decisions, as well as in sports, those highly trained proceduralized routines that are assumed to underlie real time performances could be at least partly re-assembled (see § 5).

In the remainder of this chapter, I discuss these ideas, their falsifiability and their common origin, the knowability issue. The uncertainty principle that is referred to in the title of this chapter represents a fully – albeit nontraditional – rationalist answers to the question of ‘how do we know what learners know?’.

3. Second language development could be twofold

Long (2016) provides a review of the discontinuity hypothesis (DH, Rastelli, 2014, 2019, 2021, forthcoming). According to the DH, adults can learn part of L2 morphosyntax twice, in two different ways. The same item can be learned as the product of generation by a rule or as a modification of a template already stored in memory. These learning modalities, which are often seen as opposed in language theory, are entangled and superposed in adult SLA. Learners may resort to grammatical rules and statistical templates under different circumstances during language processing. L2 development would consist of the shift from statistical learning (SL) to grammatical learning (GL). Such shift could be an abrupt qualitative change, and not a gradual transition to the next stage in a continuous developmental process. For it

to happen, learners need to encounter a statistically critical number of instances of a form or structure/construction. This critical number does not only concern the mere repetition of the whole form/structure, but also the number of instances in which its subcomponents (e.g., stem and affix, auxiliary and past participle, article and noun etc.) are encountered disjointly in different contexts. For example, in order for the Italian auxiliaries *essere* 'be' and *avere* 'have' in the perfective past to be learned, learners are expected to encounter process those auxiliaries with different verbs. The more diverse its structured 'variation set', the faster and more stably the meaning and function of the auxiliary will be learned. Once such quantitative threshold for acquisition is reached and crossed, learners are able to perceive regularities in the features the form/structure instances share, and to conceptualize – always implicitly – the motivation behind those regularities to apply a rule to novel instances. "Crucially, the new grammatical representations do not displace previously acquired statistical rule(s). Rather, the sudden shift to GL is marked by *geminatio*: dual statistical and grammatical representation of an item or structure at two cognitive levels in underlying competence. The two learning processes, SL and GL, and the two mental representations for the same L2 phenomena, statistical rules and grammatical categories, continue to exist side by side. Discontinuity is not the same as nonlinearity. Discontinuity is when a representation geminates, such that when language acquisition is plotted over time on Cartesian axes, there is a twofold limit (two points) on the y axis instead of one" (Long, 2016, p. 2).

A feature of the DH Mike was particularly interested in was how it can be falsified. The difference between SL and GL is qualitative (neurophysiological) and cannot be revealed by any measure of automatization alone, such as reaction times. Indeed, both statistical and grammatical knowledge are implicit and highly automatized, thus the speeding up (or slowing down) of learners' decisions in behavioral tasks cannot discriminate between the statistical and the grammatical nature of the underlying knowledge. Another approach would be to look at the developmental shift between N400 and P600 components as revealed by the Event-Related Potentials (ERP) technique, as the N400-P600 shift would signal the grammaticalization of previously stored statistical knowledge (Osterhout et al., 2006). However, the presence of the P600 component cannot be solely and univocally linked to grammaticalization, but also to post-processing revision stage, to wrap-up effects, and to a number of distinct sub-components, each of which may reflect different kinds of cognitive processes. A proper – perhaps complementary – approach would be to utilize eye-tracking to analyze statistical and grammatical patterns of reading behavior. SL and GL likely provide the language parser with different instructions about how and where to look in online processing. Rastelli (forthcoming) compared how Italian auxiliaries *avere* 'have' and *essere* 'be' in past compound tenses (e.g., *è arrivata* 'she arrived / has arrived' vs. *ha lavorato* 'she worked / has worked') are

read and processed by the same L1 Chinese-L2 Italian learners over time. SL was estimated by looking at whether learners' response speed and accuracy to sentence acceptability depended on both absolute frequency of the auxiliary + past participle collocation and the backward transition probabilities (BTP) between the auxiliary and the past participle. GL was estimated as a function of learners' length of fixations or regressions over the auxiliary immediately before learners' decisions on sentence acceptability. The results of the study showed that learners changed the way they read the same sentence over time, by shifting their attention from the statistical pivot of the sentence (the lexical part of the compound, i.e., the past participle *arrivata* 'arrived') to the grammatical pivot (the auxiliary and its abstract frequency-independent properties). Such results may be indicative of the existence of a developmental shift from SL to GL. However, they can say nothing about the actual superposition of statistical and grammatical knowledge in a L2 learner's competence, that is, if and when learners and speakers shift from one modality to another during language processing, nor about the fact that the shift is abrupt.

4. Second language acquisition research could benefit from the study of intra-language

If the same L2 item can be learned statistically, as a modification of a template (or productive schema-construction) stored in memory, and grammatically, as the product of generation by a rule, then SLA researchers should pay attention not only to the differences *between* the items produced by learners over time in different interlanguage systems, but also to the differences existing *within* items that appear already perfectly target-like in learners' production. Analyzing target-like forms and not only how forms (and functions) change over time would represent an important addition to the study of L2 development. While the study of the interlanguage concerns how forms-functions mapping change across different systems ('between-systems variance'), the study of the *intra-language* investigates how target-like forms are generated, statistically or grammatically ('within-item variance'). Rastelli (2021) and Rastelli and Murakami (2022) propose to study the intra-language by utilizing a statistical method based on contingency measure ΔP (delta-pi). The two studies aim at analyzing morpheme-to-lexeme contingency in an L2 Italian corpus in order to discriminate between the statistical and the grammatical nature of target-like morphosyntactic forms (of the Italian past perfective). The hypothesis is that lexeme-morpheme contingency changes within the same target-like form over time. Initially, the perfective morpheme will be contingent upon an exceptionally limited number of lexemes. 'Exceptionally' means that the morpheme is not distributed as one would expect given the statistical distribution

of all other morphemes in the corpus. This abnormal morpheme-to-lexeme reliance indicates that the morpheme is not (yet) productive and that the corresponding morphosyntactic form – although being perfectly target-like – is probably represented as an unanalyzed chunk in learners' competence. The studies demonstrated that the contingency of lexeme–morpheme associations diminished as learners' proficiency level progressed. This could be due to the fact that learners started to conceive the morpheme independently of the lexeme and starts to extend it to the available lexicon. Admittedly, the analysis of ΔP is not informative about when and how the passage between SL and GL would occur.

5. Language competence could be permanently re-assembled and rehearsed during online performance

Performance science (e.g., Gigerenzer, 2007; Raab & Gigerenzer, 2015) and motor system neuroscience (i.e., the study of voluntary movements; Ajemian et al., 2013; Bizzi & Ajemian, 2015) – albeit from totally different perspectives – converge in suggesting that both intuitions behind immediate human decisions (e.g., how Italy goalkeeper Donnarumma saved the last penalty in the winning shoot-outs against England during the Euro final) and the neural networks supporting highly proceduralized motor sequences in sports are less stable and reliable than it is maintained in neurocognitive models, where memories and knowledges are defined by fixed patterns of synaptic weights that are consolidated and stabilized after learning (e.g., during training, exercise or sleep, see Ullman, 2020, p. 131). If such hypothesis is correct, one consequence would be that even the most highly practiced grammar rules – just like the most routinized human decisions and actions – even those highly trained, automatized and stored in long term memory are not immutable always at a speaker's disposal, just waiting to be retrieved. Yet, at least the motor components of such proceduralized competences – those that are based on nonstationary synaptic patterns that fluctuate coherently (Ajemian et al., 2013) – would be subject to continuous re-assembly, re-calibration, and rehearsal by both speakers and language learners.

Performance science – a branch of cognitive psychology dedicated to the study of human decisions – can provide SLA researchers with a novel conceptual framework to explain what happens when human beings must make immediate decisions without much time for thinking and under conditions of uncertainty. This framework is probably also applicable to SLA, to some extent at least. 'Uncertainty' defines situations where not all the alternatives, consequences, and probabilities are known, and/or where the available information is not sufficient to reliably estimate these conditions. Similarly, speaking and understanding an L2 is a situation in

which real time decisions (e.g., which words to utter, how to segment and interpret the stream of sounds) are made in conditions of high psychological pressure and uncertainty. Indeed, often times learners must speak and understand in real time without knowing – or knowing only partially – how to express meanings.

On a more general (cognitive and metacognitive) level, the crucial question is to what extent learners' feel that reliable, pre-determined knowledge of language rules – whether they are induced or deduced – is a pre-requisite in order to speak and understand a L2 in real time. This is a question that seems as radical as it seems bizarre, considered what 'knowing a language' means to us: does learners' performance really depend on their previous knowledge of language rules and how good is the performance system at applying such rules in real time? The alternative would be that learners – especially when under communicative pressure – can be satisfied with partial heuristics which consist of favoring the simpler over the correct and using foresight, gut feelings and all sort of things that have nothing to do with any idea of grammaticality, replicability and consistency in language behavior. Let us admit for a moment the possibility that L2 learners can temporarily disconnect their performance from their admittedly partial – albeit stable – knowledge of the second language, especially when such knowledge is too complex. A realistic scenario – not an explanation – would be that L2 learners (as a subset of intelligent human beings) – at some developmental stages at least – are designed to automatically ignore complexity, to prune the decision algorithms to the very minimum and to disregard the probabilities that a form of the language or a given interpretation are – or sound – correct to native speakers. Athletes during the final phases of a hard-fought close competition, as well as doctors who make the right instantaneous, intuitive decision in intensive care units, may be in a comparable situation to L2 learners trying to string together words in a sentence. In such situations, they could abandon the protocols and rely on intuitions and foresight – 'gut feelings', to use Gigerenzer's (2007) words – rather than on cold reasoning, deduction, and probability *calculus* relative to the correctness to what they are doing. Their decisions would then be based on approximative, 'fast-and-frugal', moment to moment heuristics, which are neither based on all the necessary information nor on the sum of their past experience with similar situations. According to the performance science approach, human beings often rely on immediate, momentaneous heuristics in an adaptive way in order to make accurate decisions in situations of uncertainty (Raab & Gigerenzer, 2015; Hafenbrädl et al., 2016; Gigerenzer, 2020). They do not rely on conscious knowledge of probabilities, because knowledge of trade-off between probabilities and outcomes is often too costly and time-consuming. Fast-and-frugal heuristics instead are much simpler strategies that ignore part of the available information, and their rationality is only co-determined by the environment and the contingent situation. As to SLA, the adaptive heuristics

might drive learners to think that what proved correct in a situation, may not be in another similar situation. Contrary to the assumption that people under pressure stick to well-established and successful routines rather than trying out new ones, performance science and the science of voluntary movements may suggest that novel moves based on instinct, improvisation and intuition – the result of permanently re-organizing synaptic patterns – may prevail especially in situations when there is no time to think. The knowledge underlying language performance too may be partly subject to continuous re-assembly, re-calibration, and rehearsal. It is important to stress that – as in the case of the motor system neuroscience – performance science is *not* a theory of SLA and may sound rather speculative and in need of operationalization in view of empirical testing. Yet, both these approaches can be useful to investigate the relationship between the existence of mental language rules and what language speakers and learners can actually do with those rules.

6. The knowability issue

The three ideas sketched above stem directly from the knowability issue, which is the question of ‘how do we know what learners know?’ (e.g., Lakshmanan & Selinker, 2001; Lardiere, 2003). A rationalist – albeit nontraditional – answer is that any assessment of L2 learners’ competence is always a function of a measurement. The construct ‘L2 competence’ itself depends on how a learner’s competence is measured. This poses serious problems to hypotheses testing and falsifiability in SLA. For example, if one is interested in SL and in assessing how frequency impacts SLA, the existence of features that predate a learner’s exposure to the input will be totally eclipsed and lost upon any measurements based on frequency scores. Indeed, frequency scores will precisely and exclusively create the empirical object(s) they are designed to create and will filter out what does not conform to such measurement. In contrast, if one is interested in GL and how learners’ innate intuitions shape language development, high or low acceptability rates – for example – will create the empirical object that exactly fits such measurement. In this case, the amount of learners’ previous experience with the rated items will be totally lost over measurement concerning a learner’s intuitions on language. Statistical and grammatical knowledge cannot be assessed simultaneously – not even by the most refined regression models – because in some domains of the language (morphosyntax and all ‘combinatorial grammar’) these are entangled and superposed states of mind within a learner’s competence. ‘Entangled’ means that SL and GL represent complementary quantities of the same mental object (a learner’s competence of the language). Complementarity between frequency and the grammar in the domain of morphosyntax (regular alternance between stems and affixes) arises because the existence of a rule always implies the

existence of regularities in the input, and the other way round. When a grammar rule of morphosyntax is systematically applied by speakers of a language, this automatically produces regularities in the input of that language, which can be detected by a rule-independent statistical device any learner is equipped with. Conversely, when a learner's 'pattern-sensitive device' (as Ellis (2005) calls it) can track and record the regularities in the input, maybe after a certain statistical threshold is reached, a mental rule is abstracted bottom-up in that learner's mind. Later on, such inductive rule can be extended to novel items of the language, purely on the basis of analogy or other cognitive mechanisms. In sum, frequency can be both the cause and the consequence of the existence of a rule, and conversely, a rule can be both the cause and the consequence of the frequency and distribution of an item in the input. Frequency and grammar are dimensions that cannot be separated from one another in a learner's mind. As soon as a language is used in the real world by real human beings, each and every element of that language will assume two values (a.k.a. 'magnitudes') simultaneously: one is grammatical (abstract or input independent), the other statistical. What SLA researchers currently lack is metrics that can be used to measure such complementary values together and with arbitrary precision. Actually, when evaluating the contribution of statistics and grammar to language acquisition, most past and present SLA approaches tend to frame frequency and grammar (i.e., SL and GL) within the innatism vs. non-innatism debate. In this context, SL opposes GL, as the former is "learning by data" and the latter "learning by rules." The resulting knowledge is, respectively, input-driven (if learning depends on frequency and distribution of words) or knowledge-driven (if abstract, frequency-independent categories drive learning). Of the two hardline positions, one holds that first and second language acquisition is entirely statistical, a gradual distillation of frequencies and co-occurrence patterns (e.g., Goldberg, 2008; Chater & Christiansen, 2010), the other that the "soul of the language" does not use statistics at all (e.g., Mehler et al., 2006; Dekydsporter, 2009) because statistical facts are secondary and derivative in SLA, with little or no direct explanatory force (Bley-Vroman, 2002). The idea described in this section is that statistical procedures and grammatical (frequency-independent) representations are entangled dimension which cohabit and superpose in an adult learner's competence.

7. The uncertainly principle and second language development

The observation that entangled dimensions of the same object cannot be measured simultaneously and that the object of any experiment becomes real and observable (i.e., enters a state of reality) by researchers only when it is measured (by only showing the features that fit that precise measurement) echoes the formulations of

the uncertainty principle in quantum mechanics theory of the last century. Both Bohr's and Heisenberg's versions of such principle maintained that it is impossible to know the details of a system without perturbing it. Namely, it is not possible to simultaneously determine the value of two complementary variables within the same system with arbitrary precision. In the case of SLA, the system that is perturbed by measurement is a learner's mental competence, and the complementary variables that cannot be measured simultaneously are grammar rules and statistical regularities. The uncertainty principle in SLA poses limits to the rational inquiry of 'what we can now about what learners know', but it also enlightens novel areas of research and new connections with the sciences of nature, quantum physics included.

Finally, it should be recalled that uncertainty is neither irrationalism nor a post-modernist confutation of the primacy of scientific inquiry in SLA research. Yet, the uncertainty principle introduces a new object for the scientific inquiry in the SLA field. This object is measurement. Modern science is struggling with the idea that Any measurement creates (ontologizes) only those features of the phenomenon that fit measurement. Why should SLA be immune from such debate? A learner's (and speaker's) competence of the language presents at least two such complementary features: frequency and the grammar. Falsifiability of SLA theories must encompass the measurement issue. In the case of the DH, falsifiability depends on our capacity of imagining new experiments (maybe mental experiments, similar to Wheeler's 'delayed choice' experiment) in which learners (and native speakers) can represent identical things in two ways, statistically and grammatically.

8. Arrivederci

If I ever had the chance to sit down again with Mike at dinner in Tuscany (or anywhere else on earth he ever liked), I would timidly try to convince him that uncertainty is inherent in future rationalist SLA research and that the uncertainty principle must be dealt with, as it has been already in other fields of science (e.g., cognitive psychology, vector semantics etc.). To bend his indulgent sarcasm, I would insist that uncertainty does not belong to irrationalism or post-modernism. Rather, it is rationalism updated and taken to its extreme consequences. It means acceptance of the limits of our work. It is also a source of fresh and new ideas that can finally put aside the innatism (generative) vs. non-innatism (usage-based) contention, the infinite debate that many of us have never felt comfortable with.

References

- Ajemian, R., D'Ausilio, A., Moorman, E., & Bizzi, A. (2013). A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits. *Proceedings of the National Academy of Sciences (PNAS)*, 110(52) E5078–E5087. <https://doi.org/10.1073/pnas.1320116110>
- Bizzi, E., & Ajemian, R. (2015). A hard scientific quest: Understanding voluntary movements. *Daedalus*, 144(1), 83–95. https://doi.org/10.1162/DAED_a_00324
- Bley-Vroman, R. (2002). Frequency in production, comprehension and acquisition. *Studies in Second Language Acquisition*, 24(2), 209–213. <https://doi.org/10.1017/S027226310200205X>
- Chater, N., & Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science*, 34(7): 1131–1157. <https://doi.org/10.1111/j.1551-6709.2009.01049.x>
- Chomsky, N. (1995). *The minimalist program*. The MIT Press.
- Dekydsprotter, L. (2009). Second language epistemology. *Studies in Second Language Acquisition*, 31(2), 291–321. <https://doi.org/10.1017/S0272263109090317>
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–352. <https://doi.org/10.1017/S027226310505014X>
- Gigerenzer, G. (2007). *Gut feelings. Short cuts to better decision making*. Penguin Books.
- Gigerenzer, G. (2020). How to explain behavior. *Topics in Cognitive science*, 12(4), 1363–1381. <https://doi.org/10.1111/tops.12480>
- Goldberg, A. (2008). Universal grammar? Or prerequisites for natural languages. *Behavioral and Brain Sciences*, 31(5), 522–523. <https://doi.org/10.1017/S0140525X0800513X>
- Gregg, K., Long, M. H., Jordan, G., & Beretta, A. (1997). Rationality and its discontents in SLA. *Applied Linguistics*, 18(4), 538–558. <https://doi.org/10.1093/applin/18.4.538>
- Hafenbrädl, S., Waeger, D., Marewski, J. N., & Gigerenzer, G. (2016). Applied decision making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition*, 5, 215–231. <https://doi.org/10.1016/j.jarmac.2016.04.011>
- Hulstijn, J. H. (2013). Is the second language acquisition discipline disintegrating? *Language Teaching*, 46(4), 511–517. <https://doi.org/10.1017/S0261444811000620>
- Jordan, G. (2004). *Theory construction in second language acquisition*. John Benjamins. <https://doi.org/10.1075/llt.8>
- Lakshmanan, U., & Selinker, L. (2001). Analysing interlanguage: How do we know what learners know? *Second Language Research*, 17, 393–420. <https://doi.org/10.1177/026765830101700406>
- Lardiere, D. (2003). The comparative fallacy revisited: A reply to Lakshmanan and Selinker (2001). *Second Language Research*, 19, 129–143. <https://doi.org/10.1191/0267658303sr160a>
- Long, M. H. (1993). Assessment strategies for SLA theories. *Applied Linguistics*, 14, 225–249. <https://doi.org/10.1093/applin/14.3.225>
- Long, M. H. (2015). First person singular: Building the road as we travel. *Language Teaching*, 48(4), 561–574. <https://doi.org/10.1017/S0261444815000282>
- Long, M. H. (2016). Stefano Rastelli: Discontinuity in second language acquisition. The switch between statistical and grammatical learning. Book review. *Applied Linguistics*, 37(4), 590–597. <https://doi.org/10.1093/applin/amw011>
- Mehler, J., Peña, M., Nespors, M., & Bonatti, L. (2006). The soul of the language does not use statistics: Reflections on vowels and consonants. *Cortex*, 42(6), 846–854. [https://doi.org/10.1016/S0010-9452\(08\)70427-1](https://doi.org/10.1016/S0010-9452(08)70427-1)

- Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs and event-related potentials: A means for exploring the neurocognition of second language processing. In P. Indefrey & M. Gullberg (Eds.), *The cognitive neuroscience of second language acquisition* (pp. 199–230). Blackwell.
<https://doi.org/10.1111/j.1467-9922.2006.00361.x>
- Pallotti, G. (2021). Cratylus' silence. On the philosophy and methodology of Complex Dynamic System Theory in SLA. *Second Language Research*. <https://doi.org/10.1177/0267658321992451>
- Raab, M., & Gigerenzer, G. (2015). The power of simplicity: A fast-and-frugal heuristics approach to performance science. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2015.01672>
- Rastelli, S. (2014). *Discontinuity in second language acquisition: The switch between statistical and grammatical learning*. *Multilingual Matters*. <https://doi.org/10.21832/9781783092475>
- Rastelli, S. (2019). The Discontinuity Model: Statistical and grammatical learning in adult second language acquisition. *Language Acquisition*, 26(4), 387–415.
<https://doi.org/10.1080/10489223.2019.1571594>
- Rastelli, S. (2020). Contingency learning and perfective morpheme productivity in L2 Italian: A study on lexeme–morpheme associations with ΔP . *Corpus Linguistics and Linguistic Theory*, 16(3), 459–486. <https://doi.org/10.1515/cllt-2019-0071>
- Rastelli, S. (2021). Intra-language: The study of L2 morpheme productivity as within-item variance. *International Review of Applied Linguistics IRAL*. <https://doi.org/10.1515/iral-2020.0082>
- Rastelli, S. (forthcoming). Statistical and grammatical processing of auxiliaries in L2 Italian: An eye-tracking study.
- Rastelli, S., & Murakami, A. (2022). Apparently identical verbs can be represented differently. Comparing L1-L2 inflection with contingency-based measure ΔP . *Corpora*, 17/1, 97–121.
<https://doi.org/10.3366/cor.2022.0236>
- Ullman, M. T. (2020). The Declarative-Procedural Model. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition. An introduction* (3rd ed., pp. 128–161). Routledge. <https://doi.org/10.4324/9780429503986-7>

Tandem interaction enhancement

Manipulating NS-NNS semi-spontaneous conversation to promote focus on form

Elena Nuzzo

Università Roma Tre

This chapter discusses the applicability of the construct of focus on form, one of Michael Long's most prominent contributions to the field of instructed second language acquisition, in the very particular context of tandem learning. In tandem interactions, two partners learning each other's language assume in turn the role of the native-speaker teacher. These low-structured learning environments would benefit from the implementation of focus-on-form pedagogical procedures including, but not limited to, corrective feedback. An external manipulation on the part of the tandem organizer, which I call 'interaction enhancement', is needed to push the partners to systematically and confidently apply focus-on-form techniques to their exchanges.

1. Introduction

In many educational institutions all over the world, language instructors have only a few teaching hours per week. A minimal part of these hours is usually devoted to spoken interaction, and learners' interlocutors are unlikely to be native speakers of the language they are learning. There might be differences among institutions and countries, but having conversation with native speakers of the target language is probably more an exception than the rule for most high school and university students in foreign language learning contexts (Loewen & Sato, 2018). For this reason, some teachers decide to complement curricular activities with (e-)tandem programs.

Language tandem is a learning arrangement in which two people with different mother tongues work together to learn each other's languages (Brammerts & Calvert, 2003). Through tandem learning, L2 learners can expand their opportunities to engage in meaningful and goal-oriented communication with native or near-native speakers of their target language as well as establish relationships with

people from other cultures and develop intercultural competences. The advent of online videoconferencing tools has dramatically increased the opportunities for language students and instructors to take advantage of tandem practices, either outside of a traditional classroom or within a classroom. With the evolution of technology, and after the massive digital shift caused by the pandemic in 2020, the upcoming decade is likely to witness a further expansion of these virtual exchanges.

The type of native-nonnative speaker (NS-NNS) interaction which occurs within the tandem setting “pertains neither to the clear (and socially institutionalized) hierarchy of teacher-learner interaction nor to the exact symmetry of peer interaction where language learners share the same L1 and the same target L2” (Scheuer & Horgues, 2020, p. 147). At some point in the conversation, the native speaker of the language being spoken serves as a model and tutor towards their nonnative partner, thus temporarily assuming the duties of a teacher. How this teaching role should be fulfilled by the members of the partnership is not clearly defined, although a number of studies have tackled the issue by exploring the use of peer corrective feedback in tandem exchanges.

Scholars and practitioners interested in tandem learning have mainly focused on the intercultural value of these NS-NNS exchanges (Ware & Pérez Cañado, 2007, p. 107), whereas instructed second language acquisition (ISLA) research has privileged teacher-student and NNS-peer interaction in classroom or NS-NNS interaction in laboratory settings (Loewen & Sato, 2018, p. 293). This chapter aims to promote a dialogue between the two research traditions, assuming that tandem is a learning environment in which external manipulations can have positive effects on second language acquisition. Specifically, it is suggested that tandem interactions would benefit from the systematic implementation of focus-on-form procedures, although the construct of focus on form was originally developed for classroom-based – and essentially teacher-to-student – interactions. As tandem partners may not spontaneously integrate these procedures into their exchanges, I argue that focus on form needs to be promoted through strategies of ‘interaction enhancement’ in tandem programs.

2. Tandem learning contexts

Language tandems provide a unique collaborative learning environment where two persons with different L1s learn each other’s native languages in reciprocal cooperation. The recent shift to a multilingual approach to language learning with the crisis of the very concept of ‘native speakerness’ (Kramsch, 2009; Woodin, 2020; cf. also Hylthestam & Abrahamsson, this volume) has been challenging this key assumption of tandem learning (see the discussion in Tardieu & Horgues, 2020b).

However, for the purposes of this chapter, I refer to the prototypical tandem situation where participants are native (or near native) speakers of each other languages.

The two tandem partners function, in turn, as a second language learner and as an expert and tutor in the native language, with the shared purpose of enhancing their proficiency in the L2 (Brammerts, 1996; Brammerts & Calvert, 2003; Little & Brammerts, 1996). The roles switch as the language of the conversation changes. Therefore, tandem learning entails a dynamic asymmetrical relation between the partners, where role reversal makes the communicative situation minimally hierarchical as the participants “feel they are in the same boat” (Scheuer & Horgues, 2020, p. 148).

This very particular momentary language expertise asymmetry is counterbalanced by the principle of reciprocity, whereby both learners should not only contribute equally to their language exchanges in quantitative terms and in terms of reciprocity of respective L1/L2 input, but also in terms of commitment. This means that they should “devote the same length of time to each language and to invest the same quantity of energy and involvement in their partner’s success” (Reymond & Tardieu, 2001, p. 21).

Another basic principle underpinning tandem learning is learner autonomy which dictates that each tandem participant is responsible for directing their own learning. Learners set their own objectives and decide how to reach them, and plan meetings together with their partners according to their needs. This does not imply that language instructors and tutors do not have a role in organizing activities and assisting participants. As Helmling (2002) pointed out, learner autonomy may be guaranteed and even enhanced by a certain form of control, monitoring and guidance performed by the institution. In other words, the principle of autonomy is not incompatible with institutionalized forms of tandem learning. As a matter of fact, tandem activities are often integrated and acknowledged by higher education institutions, or even regarded as a mandatory component of the language courses (Tardieu & Horgues, 2020b, p. 270).

Tandem learning has been part of foreign language education for over 50 years in Europe. Its origins can be traced back to collaborative practices introduced by the *Office Franco-Allemand de la Jeunesse* (Franco-German Youth Office) which aimed at fostering cooperation between French and German youths through linguistic and cultural exchanges in holiday camps during the late 1960s (Tardieu & Horgues, 2020a, p. 1).

The early years of the 21st century witnessed the integration of new video-call technology in tandem learning. So, in addition to traditional in-person tandems and email- or chat-based written exchanges introduced in the 1990s, videoconferencing-supported tandem programs have expanded learners’ opportunities to engage in meaningful oral communication with native speakers of their target language.

Several projects in Europe have helped teachers experiment and implement various forms of online tandem learning from extra-curricular to class-embedded. However, e-tandem (also called teletandem, cf. Telles & Vassallo, 2006) learning has always been a peripheral activity, carried out by a relatively small group of convinced practitioners. It was the onset of COVID-19 in 2020 that drew the attention of university and school management to e-tandems, and virtual exchange in general, on a much larger scale. After the sudden spread of distance learning caused by the pandemic, it is likely that all forms of virtual exchange, including e-tandem learning, will become an important part of foreign language education in the next years (O'Dowd, 2021).

Among the most frequently mentioned objectives of tandem interaction are cultural learning, developing learners' intercultural competence, and improving fluency and discursive competences in the second language (see, e.g., Hilton, 2020; Lewis, 2020; Vinagre & Corral Esteban, 2020; Ware & Pérez Cañado, 2007; Woodin, 2018). The intercultural dimension is highly emphasized in tandem literature as this learning context, particularly with the new possibilities offered by technology, can foster "communication across cultures and borders as a form of social action in service of better inclusion and social citizenship" (Tardieu & Horgues, 2020a, p. 4).

Nevertheless, some scholars have highlighted the value of having participants actively reflect on language form in tandem learning (see, e.g., Belz, 2006; Hilton, 2020; Kramersch, 2014; Lee, 2006; O'Dowd, Sauro, & Spector-Cohen, 2020; Sotillo, 2005; Vinagre & Muñoz, 2011; Ware & O'Dowd, 2008; Ware & Pérez Cañado, 2007) and to incorporate a focus-on-form component within this low-structured learning environment. For example, Vinagre and Muñoz (2011) claimed that

introducing a focus on form in these exchanges is crucial, especially for those students with intermediate or higher levels of competence in the foreign language. Students who learn a foreign language, especially at university, are expected not only to communicate with native speakers but also to develop a command of the language that will allow them to use it correctly in their future working environments. (p. 72)

The majority of the studies reporting on tandem programs which include a focus-on-form component or analyzing tandem interactions from the perspective of learners' attention to form explored the use and the effects of peer corrective feedback as a pedagogical tool for the development of language proficiency. Most of them examined asynchronous (email) or synchronous (real-time chat) text-based exchanges (Bower & Kawaguchi, 2011; Dussias, 2006; Kessler, 2009; Giguère & Parks, 2018; Lee, 2008, 2011; Muñoz, 2008; O'Rourke, 2005; Sauro, 2009; Sotillo, 2006; Vinagre & Lera, 2008; Vinagre & Maíllo, 2007; Vinagre & Muñoz, 2011; Ware & O'Dowd, 2008). Fewer scholars, at least to my knowledge, observed spoken tandem interactions with an eye to learners' attention to form (Akiyama, 2014,

2016; Debras, Horgues & Scheuer, 2015; Nuzzo & Cortés Velásquez, 2021; Saito & Akiyama, 2017; Saito, Suzuki, Oyama, & Akiyama, 2019; Scheuer & Horgues, 2020). This is not surprising, as video-based environments have not been investigated as extensively as other types of CMC, and only recently have SLA researchers begun to explore how L2 learners develop their linguistic knowledge and skills through interacting with NS interlocutors using video-conferencing (Saito et al., 2019, p. 550).

These studies, mainly observational and descriptive in nature, reported somewhat conflicting findings with respect to peer feedback provided by native speakers within tandem settings.

The quality of NSs participants' corrective feedback appears often questionable for a number of reasons. Firstly, learners are differently equipped to provide accurate feedback, depending on both individual background and different educational traditions (Bower & Kawaguchi, 2011; Scheuer & Horgues, 2020). Generally speaking, the feedback provided by students tends to be limited in scope or accuracy (Lee, 2011), and considerably less complete than what a teacher could provide (Ware & O'Dowd, 2008). Secondly, it has been observed that, with growing familiarity between the partners, NS participants become increasingly tolerant of their partners' erroneous productions and prioritize smooth communication (Brammerts & Calvert, 2003; Helmling, 2002; Scheuer & Horgues, 2020). In some cases, tandem participants do not even include peer feedback into their interactions unless given explicit directions to do so by the instructor (Ware & O'Dowd, 2008). This is not surprising, as tandem participants may feel they are not expected to behave as language teachers. Rather, they may feel they are expected to act as an empathizing peer, and to maintain a friendly and smooth relationship that is not as face-threatening as a teacher-student interaction (Horgues & Scheuer, 2015). Therefore, tandem learning organizers should be able to present peer-to-peer corrective feedback as a natural component of tandem interactions and to support participants in their role of feedback providers. In fact, as other autonomous learning environments, tandem may lower anxiety and function as a space where learners can practice their target languages without feeling overly self-conscious about their speech production (Akiyama, 2016, p. 2).

However, learners involved in these types of exchanges usually appreciate and even ask for NSs' corrective feedback (Akiyama, 2016), as they feel that they do not improve their language competence unless some form of feedback has been provided by their partners (Vinagre & Muñoz, 2011). Corrective feedback is often solicited by the non-native participant during conversation, either explicitly or implicitly, with various forms of appeal to the native speaker, both verbal and non-verbal (Debras, Horgues, & Scheuer, 2015; Scheuer & Horgues, 2020). This suggests that tandem is perceived by participants as a learning context which

naturally involves corrective feedback. As pointed out by Little, Ushioda, Appel, Moran, O'Rourke, & Schwienhorst, (1999, p. 39), "error correction is the central overtly pedagogical element of a tandem partnership".

3. Focus on form

As we have seen, focus on form usually coincides with corrective feedback in tandem literature. This is not without a reason, as corrective feedback is the most common instantiation of focus on form. In addition, it is worth remembering that the term focus on form has often been misused and distorted since when Michael Long coined it more than thirty years ago. However, the reactive nature of Long's focus on form does not imply necessarily that it is triggered by a learner's error, as will be discussed in the following paragraphs, and tandem learning could benefit from a wider variety of pedagogical procedures pertaining to the category of focus on form.

Michael Long proposed the construct of focus on form (FonF) to refer to different pedagogical procedures that can be used to briefly and overtly draw learners' attention to linguistic code features in context, during communicative activities (Long, 1988, 1991; Long & Crookes, 1992; Long & Robinson, 1998). A key characteristic of Long's FonF is that of being reactive in nature, with two important consequences. On one hand, this reactive approach to treating linguistic elements increases "the likelihood that attention to code features will be synchronized with the learner's internal syllabus, developmental stage, and processing ability" (Long, 2015, p. 27). In other words, FonF instructional intervention has good chances to be developmentally appropriate and aligned with learners' linguistic and cognitive capacities, because "a student's attempt to produce a form is not always, but often, an indication of his or her developmental readiness to acquire it" (Long, 2015, p. 28). On the other hand, in this reactive mode "the learner's underlying psychological state is more likely to be optimal" (Long, 2015, p. 27) as the switch in attention is a "response to a difficulty [...] due to a learner's current incomplete stage of L2 development" (Long, 2015, p. 317). The instructional intervention arrives at a moment when the learner is highly motivated to notice and possibly learn a formal feature that is needed to convey a particular meaning. Furthermore, when the FonF treatment takes the form of negative feedback, the learner "hears the correct form in close juxtaposition to his or her own, facilitating cognitive comparison" (Long, 2015, p. 27).

The reactive nature of FonF distinguishes it substantially from *focus on forms* (FonFs), a teaching approach in which the forms for treatment are selected *a priori* by the instructor based on an external synthetic syllabus. According to Long (Doughty & Williams, 1998; Long & Robinson, 1998), FonF is a "third option" as

opposed to FonFs, which he considers in contrast with what we know about how second language acquisition works, and *focus on meaning* (FoM), which research suggests to be inadequate for rate of acquisition and level of ultimate attainment, at least with adult learners (Doughty, 2003).

Given its reactive nature, FonF needs a trigger. FonF intervention is usually triggered by a communicative problem, either receptive or productive: perceived troubles with comprehension or production may naturally cause an attentional shift from meaning to form through negotiation (Long, 1996). However, communicative difficulties are not the only trigger for FonF, which can also be used to extend “a learner’s repertoire as opportunities arise, e.g., by a teacher reformulating and extending already acceptable learner speech or writing” (Long, 2015, p. 28). A teacher can draw a learner’s attention to an alternative way of formulating what he or she has just said perfectly correctly, not because communication is failing, but because the alternative is a better (e.g., more native-like, or pragmatically more adequate) option. This may happen also with another student, or with an “interlocutor outside the classroom” (Long, 2015, p. 317), though it is obvious that the classroom, and teacher-student interaction specifically, is a privileged context for these FonF episodes.

Long (2009) distinguished between methodological principles (MP) and pedagogic procedures (PP), with the former referring to instructional design features founded upon research findings, and the latter to the potentially infinite range of options for implementing the principles. As a methodological principle – one of the ten proposed for Task-Based Language Teaching (Long, 2015, p. 301) –, FonF can be implemented with a variety of pedagogic procedures. Given that variations in implementation respond to needs at the local level, choices among PPs must be made according to local conditions.

4. Interaction enhancement

In the usual classroom setting, it is the teacher who decides whether and how to apply FonF techniques during a communicative activity which she or he has planned and put out. In the very particular context of tandem learning, the teacher’s role is shared between the instructor organizing the program and the current NS in the partnership. The former plans the activities and instructs the participants, whereas the latter is in charge of managing the attentional shifts from meaning to form while interacting with the NNS partner. As the NS ‘teacher’ in the partnership may not spontaneously initiate a FonF episode unless communication is impeded, organizers of tandem programs have the task of increasing the opportunities for participants to focus on form by indirectly manipulating the interaction.

Muranoi (2000, p. 617) used the term ‘interaction enhancement’ to refer to “a treatment that guides learners to focus on form by providing interactional modifications and leads learners to produce modified output within a problem-solving task”. In his study, interaction enhancement was operationalized as an instructional technique in which the teacher provided corrective feedback while performing an interactive task with students in turn. In the context of tandem learning, I suggest using the term interaction enhancement to identify a process of external manipulation of the NS-NNS peer interaction carried out by the language instructor with the purpose of pushing the activation of FonF procedures.

Tandem interaction enhancement operates on two levels, namely participants’ training and tasks design. The first goal of the training is to make participants aware that FonF is a crucial component of their activity when they assume the teaching role as NSs of the partner’s target language. Once this crucial point has been clarified, tandem learners need to be taught what FonF is, and what its reactive nature implies. Finally, instructors should show how FonF procedures can be implemented in interaction not only when communication breakdowns occur, but also when the partner’s speech or writing can be better reformulated or extended, although already acceptable. Furthermore, teachers can use the training phase to help learners explore language reference tools which can help them in case of doubt about formal aspects of their L1.

The second component of tandem interaction enhancement is task design. Tasks for tandem programs should be carefully selected so that they can both respond to the specific needs of the learners involved and provide participants with numerous opportunities for focusing on form. To this end, tasks may include activities in which the language expertise of the current NS in the partnership is particularly relevant and interesting for task completion because the partners need to receptively or productively manipulate the language being spoken to reach their outcome. For example, learners can be asked to collaboratively produce written texts or video presentations or some other “products” entailing the effective and accurate use of the target language(s), or to rely on specific sources (written and/or audio/video material with a certain degree of difficulty for the NNS) in order to find the information needed to complete the task or to reach a shared decision on some issue. These activities, and others of a similar nature, can provide the NS with occasions to initiate language-related episodes by drawing his or her partner’s attention to an alternative (e.g., more native-like, or pragmatically or socio-linguistically more adequate) option for what he or she has just said, as well as by addressing the partner’s misunderstandings and uncertainties and providing corrective feedback. Nevertheless, the main focus of interaction remains on communication and task completion.

Of course, tandem participants cannot be expected to implement FonF procedures with the same level of correctness and confidence of a professional teacher, even after the training. This should not discourage tandem organizers, as a percentage of incorrect or naïve information on the target language is unlikely to negatively affect the general positive impact of FonF on the NNSs' interlanguage development. On the other hand, the effort required of the NSs can have valuable side effects on their metalinguistic awareness of the first language (cf. La Russa & Nuzzo, 2021). Furthermore, it should be remembered that many professional teachers are not native speakers of the language they teach. A NS tandem partner and a NNS trained teacher would perfectly complement each other in learners' instructed second language acquisition.

5. Empirical evidence and further research

Future research will be needed to empirically test the applicability and effectiveness of tandem interaction enhancement as outlined here – that is, a set of procedures aimed at systematically manipulating tandem learning at different levels to promote focus on form. However, some of the previously mentioned studies on computer-mediated spoken NS-NNS interactions have already gone in this direction and shed light on the effects of some kind of external manipulation on tandem learning. In those studies, which looked at tandem learning from an ISLA perspective, interactions were partially manipulated by researchers who selected the tasks to be completed by the participants and trained them on how to provide corrective feedback. Saito and Akiyama (2017) investigated the impact of video-based conversation on longitudinal L2 development, and Saito et al. (2019) further expanded the research focus to include the relationship between the impact of interaction and learners' different experience and proficiency levels. Akiyama (2016) examined the mediating effects of beliefs on the provision and potential noticing of corrective feedback, whereas Nuzzo and Cortés Velásquez (2021) observed how different ways of receiving and discussing peer written corrective feedback, namely direct WCF plus oral metalinguistic explanation and indirect WCF plus oral prompt, affected learners' opportunities to negotiate form.

A number of useful suggestions for systematically implementing interaction enhancement in the context of tandem learning emerged from these studies. Saito and Akiyama (2017), for example, found that that the experimental (= tandem) group significantly improved in global, temporal, and grammatical qualities of their L2 speech. However, the findings suggest the need for long programs of language exchange, longer than one academic semester. Furthermore, the authors call for the

implementation of “intentional focus on form (e.g., explicit instruction) for learners to attend to and practice nonsalient and difficult features” (p. 69). Although it is not completely clear how exactly “intentional focus on form” should be operationalized, it can be assumed that the idea goes in the same direction as the construct proposed in the present chapter, that is a focus on form which is not limited to corrective feedback.

The findings of Akiyama’s (2016) study suggest that participants’ beliefs play a role in successful repair. The author emphasizes that learners “often long for FonF but need to carry out communicative tasks in a limited amount time” and that “it is difficult for many participants [...] even when they are trained to do so” (p. 17). To help learners engage in FonF while placing primary focus on meaning it is recommended that participants engage in guided reflection after each tandem session and be supported by form-focused practices in class.

In the study by Nuzzo and Cortés Velásquez (2021), the results suggest that a more implicit type of peer written corrective feedback is more effective in pushing the NNS participants to actively engage in negotiation of form. It also confirmed that the role of language tutor assumed by the NSs in tandem learning needs support from the instructor.

6. Conclusion

Tandem learning offers second language learners an interesting opportunity to meaningfully interact with native speakers of their target language, and the recent development of web-based communication technology has expanded its potential applications. Although the most frequently mentioned objective of tandem learning is the improvement of learners’ intercultural and communicative competences, several scholars have claimed that attention to form and accuracy should not be disregarded in this low-structured learning environment. Some studies have reported on the use of peer corrective feedback both in asynchronous or synchronous text-based e-tandem exchanges and in spoken tandem interactions, either face to face or through videocall. Despite these attempts, it seems that a more general view of how attention to form should be promoted in tandem contexts is still lacking. A stronger dialogue between the research areas of tandem learning and instructed second language acquisition might help in this direction. Therefore, I have tried to show that tandem configuration, with the two partners assuming in turn the role of the teacher, could benefit from the systematic implementation of focus-on-form pedagogical procedures including, but not limited to, corrective feedback. To this end, a process of interaction enhancement is needed, consisting in both learner training and task design.

References

- Akiyama, Y. (2014). Using Skype to focus on form in Japanese telecollaboration: Lexical categories as a new task variable. In S. Li & P. Swanson (Eds.), *Engaging language learners through technology integration: Theory, applications, and outcomes* (pp. 181–209). IGI Global. <https://doi.org/10.4018/978-1-4666-6174-5.ch009>
- Akiyama, Y. (2016). Learner beliefs and corrective feedback in telecollaboration: A longitudinal investigation. *System*, 64, 58–73. <https://doi.org/10.1016/j.system.2016.12.007>
- Belz, J. (2006). At the intersection of telecollaboration, learner corpus analysis, and L2 pragmatics: Considerations for language program direction. In J. Belz & S. Thorne (Eds.), *Internet-mediated intercultural foreign language education* (pp. 207–246). Heinle & Heinle.
- Bower, J. & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning & Technology*, 15(1), 41–71. <https://doi.org/10125/44237>
- Brammerts, H. (1996). Language learning in tandem using the internet. In M. Warschauer (Ed.), *Telecollaboration in foreign language learning* (pp. 121–130). University of Hawaii.
- Brammerts, H., & Calvert, M. (2003). Learning by communicating in tandem. In T. Lewis & L. Walker (Eds.), *Autonomous language learning in tandem* (pp. 45–59). Academy Electronic Publications.
- Debras, C., Horgues, C., & Scheuer, S. (2015). The multimodality of corrective feedback in tandem interactions. In S. Molina-Plaza (Ed.), *Multimodal communication in the 21st Century: Professional and academic challenges. Procedia: Social and Behavioural Sciences*, 212, 16–22. <https://doi.org/10.1016/j.sbspro.2015.11.292>
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 256–310). Blackwell. <https://doi.org/10.1002/9780470756492.ch10>
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). Cambridge University Press.
- Dussias, P. E. (2006). Morphological development in Spanish-American telecollaboration. In J. Belz & S. Thorne (Eds.), *Internet-mediated intercultural foreign language education* (pp. 121–146). Heinle & Heinle.
- Giguère, C., & Parks, S. (2018). Child-to-child interaction and corrective feedback during eTandem ESL–FSL chat exchanges. *Language Learning & Technology*, 22(3), 176–192. <https://doi.org/10125/44663>
- Helmling, B. (2002). *L'apprentissage autonome des langues en tandem* [Autonomous language learning in tandem]. Didier.
- Hilton, H. E. (2020). A psycholinguistic perspective on “tandem learning” in the foreign language classroom. In C. Tardieu & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 46–59). Routledge.
- Horgues, C., & Scheuer, S. (2015). Why some things are better done in tandem? In J. A. Mompean & J. Fouz-Gonzalez (Eds.), *Investigating English pronunciation: Current trends and directions* (pp. 47–82). Palgrave. https://doi.org/10.1057/9781137509437_3
- Hylthenstam, K., & Abrahamsson, N. (2022). The ‘nativeness paradigm’ in CPH research. In A. G. Benati & J. W. Schwieter (Eds.), *Second language acquisition theory: The legacy of professor Michael Long*. John Benjamins. (this volume).

- Kessler, G. (2009). Student-initiated attention to form in wiki-based collaborative writing. *Language Learning & Technology*, 13(1), 79–95. <https://doi.org/10.125/44169>
- Kramersch, C. (2009). *The multilingual subject*. Oxford University Press.
- Kramersch, C. (2014). Teaching foreign languages in an era of globalization. Introduction. *The Modern Language Journal*, 98(1): 296–311. <https://doi.org/10.1111/j.1540-4781.2014.12057.x>
- La Russa, F., & Nuzzo, E. (2021). Peer corrective feedback as an opportunity for metalinguistic reflection in tandem telecollaboration. *EuroAmerican Journal of Applied Linguistics and Languages*, 8(2), 87–104. <https://doi.org/10.21283/2376905X.14.241>
- Lee, Lina. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning & Technology*, 8(1), 83–100. <https://doi.org/10.10125/25231>
- Lee, L. (2006). A study of native and nonnative speakers' feedback and responses in Spanish-American networked collaborative interaction. In J. Belz & S. Thorne (Eds.), *Internet-mediated intercultural foreign language education* (pp. 147–176). Heinle & Heinle.
- Lee, L. (2008). Focus-on-form through collaborative scaffolding in expert-to-novice online interaction. *Language Learning & Technology*, 12(3), 53–72. <https://doi.org/10.10125/44155>
- Lee, L. (2011). Focus-on-form through peer feedback in a Spanish – American telecollaborative exchange. *Language Awareness*, 20(4), 343–57. <https://doi.org/10.1080/09658416.2011.592589>
- Lewis, T. (2020). From tandem learning to e-tandem learning: How languages are learnt in tandem exchanges. In S. Gola, M. Pierrard, E. Tops, & D. Van Raemdonck (Eds.), *Enseigner et apprendre les langues au XXIe siècle. Méthodes alternatives et nouveaux dispositifs d'accompagnement*. Peter Lang.
- Little, M., & Brammerts, H. (1996). *A guide to language learning in tandem via the Internet* (CLCS Occasional Paper 46). Trinity College, Center for Communication and Language Studies.
- Little, D., Ushioda, E., Appel, M. C., Moran, J., O'Rourke, B., & Schwienhorst, K. (1999). *Evaluating tandem language learning by e-mail: Report on a bilateral project* (CLCS Occasional Paper 55, pp. 1–54). Trinity College, Centre for Language and Communication Studies.
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329. <https://doi.org/10.1017/S0261444818000125>
- Long, M. H. (1988). Instructed interlanguage development. In L. M. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 115–141). HarperCollins.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, C. Kramersch, & R. Ginsburg (Eds.), *Foreign language research in a cross-cultural perspective* (pp. 39–52). John Benjamins. <https://doi.org/10.1075/sibll.2.07lon>
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Emerald.
- Long, M. H. (2009). Methodological principles in language teaching. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 373–394). Blackwell. <https://doi.org/10.1002/9781444315783.ch21>
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based language teaching. *TESOL Quarterly*, 26(1), 27–56. Reprinted in K. Van den Branden, M. Bygate, & J. M. Norris (Eds.), *Task Based Language Teaching: A reader* (pp. 57–81). John Benjamins. <https://doi.org/10.1075/tblt.1.02van>

- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15–41). Cambridge University Press.
- Muñoz, B. (2008). Análisis y corrección de errores en un intercambio tandem español-alemán a través del correo electrónico (Unpublished doctoral dissertation). Universidad Nebrija, Madrid.
- Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50(4), 617–673. <https://doi.org/10.1111/0023-8333.00142>
- Nuzzo, E., & Cortés Velásquez, D. (2021). Minding the gap: A small-scale study on negotiation of form in telecollaborative tasks. *Instructed Second Language Acquisition*, 5(2), 232–257. <https://doi.org/10.1558/isla.19812>
- O’Dowd, R. (2021). Virtual exchange: Moving forward into the next decade. *Computer Assisted Language Learning*, 34(3), 209–224. <https://doi.org/10.1080/09588221.2021.1902201>
- O’Dowd, R., Sauro, S., & Spector-Cohen, E. (2020). The role of pedagogical mentoring in virtual exchange. *TESOL Quarterly*, 54(1), 146–72. <https://doi.org/10.1002/tesq.543>
- O’Rourke, B. (2005). Form-focused interaction in online tandem learning. *CALICO Journal*, 22(3), 433–466. <https://doi.org/10.1558/cj.v22i3.433-466>
- Reymond, C., & Tardieu, C. (2001). *Guide tandem pour l'apprentissage des langues en binôme dans le secondaire*. IUFM et CRDP de l’académie de Haute-Normandie/CNDP, Coll. Chemin faisant.
- Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, 67(1), 43–74. <https://doi.org/10.1111/lang.12184>
- Saito, K., Suzuki, S., Oyama, T., & Akiyama, Y. (2021). How does longitudinal interaction promote second language speech learning? Roles of learner experience and proficiency levels. *Second Language Research*, 37(4), 547–571. <https://doi.org/10.1177/0267658319884981>
- Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, 13(1), 96–120. <https://doi.org/10.125/44170>
- Scheuer, S., & Horgues, C. (2020). Corrective feedback in English/French spoken tandem interactions. In C. Tardieu & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 146–160). Heinle & Heinle.
- Sotillo, M. S. (2005). Corrective feedback via instant messenger learning activities in NS-NNS and NNS-NNS dyads. *CALICO Journal*, 22(3), 467–496. <https://doi.org/10.1558/cj.v22i3.467-496>
- Sotillo, S. M. (2006). Using instant messaging for collaborative learning: A case study. *Journal of Online Education*, 2(3). Retrieved on 9 August 2021 from <https://www.learntechlib.org/p/104277/>
- Tardieu, C., & Horgues, C. (2020a). Introduction. In C. Tardieu & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 1–12). Heinle & Heinle.
- Tardieu, C., & Horgues, C. (2020b). Conclusion. In C. Tardieu, & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 270–280). Heinle & Heinle.
- Telles, J. A., & Vassallo, M. L. (2006). Foreign language learning in-tandem: Teletandem as an alternative proposal in CALLT. *The ESPecialist*, 27(2), 189–212.
- Vinagre, M., & Corral Esteban, A. (2020). Tracing the development of intercultural competence in telecollaborative interaction: An analysis of evaluative language in eTandem exchanges. In C. Tardieu & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 95–113). Heinle & Heinle.

- Vinagre, M., & Lera, M. (2008). The role of error correction in online exchanges. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced language acquisition and learning* (pp. 326–341). IGI Global. <https://doi.org/10.4018/978-1-59904-895-6.ch019>
- Vinagre, M., & Maíllo, C. (2007). Focus on form in on-line projects: Linguistic development in e-mail tandem exchanges. In C. Perrián (Ed.), *Revisiting language learning resources* (pp. 91–112). Cambridge Scholars.
- Vinagre, M., & Muñoz, B. (2011). Computer-mediated corrective feedback and language accuracy in telecollaborative exchanges. *Language Learning & Technology*, 15(1), 72–103.
- Ware, P. D., & Pérez-Cañado, L. (2007). Grammar and feedback: Turning to language form in telecollaboration. In R. O’Dowd (Ed.), *Online intercultural exchange: An introduction for foreign language teachers* (pp. 107–126). Multilingual Matters. <https://doi.org/10.21832/9781847690104-008>
- Ware, P. D., & O’Dowd, R. (2008). Peer feedback on language form in telecollaboration. *Language Learning & Technology*, 12(1), 43–63. Retrieved on 9 August 2021 from <http://llt.msu.edu/vol12num1/pdf/wareodowd.pdf>
- Woodin, J. (2018). *Interculturality, interaction and language learning: Insights from tandem partnerships*. Routledge. <https://doi.org/10.4324/9781315640525>
- Woodin, Jane. (2020). From a cultural to an intercultural approach. Tandem learning and the intercultural speaker. In C. Tardieu & C. Horgues (Eds.), *Redefining tandem language and culture learning in higher education* (pp. 30–46). Heinle & Heinle.

Immediate versus delayed oral negative feedback

A comparison of psycholinguistic advantages

Gisela Granena and Yucel Yilmaz

Universitat Oberta de Catalunya / Indiana University

Long's Interaction Hypothesis (1981, 1983) provides a theoretical framework for understanding how interaction can contribute to second language (L2) acquisition. Specifically, negative feedback provided during interaction is viewed as a feature that can facilitate L2 development by drawing learners' attention to problematic forms (Long, 1996). Long's influential hypotheses were originally formulated in the context of face-to-face conversational interaction and, therefore, negative feedback was understood as feedback that is provided immediately to the learner in oral conversation. However, negative feedback cannot always be provided immediately to L2 learners. In distance language learning settings where the L2 is learned online and where communication between learner and instructor may take place asynchronously, feedback cannot be provided during a communicative task and it is provided at a later time, for example, at the end of a teaching unit.

Introduction

This chapter begins with a brief introduction to one of Long's main theoretical contributions to the Second Language Acquisition (SLA) field, the Interaction Hypothesis (Long, 1981, 1983a, 1983b, 1996). The next section discusses the notion of feedback timing and reviews relevant research. We then compare the affordances of immediate and delayed negative feedback from a psycholinguistic perspective and discuss the potential benefits of delayed negative feedback drawing on Long's own thoughts on the issue as informally shared with us in a series of personal communications. The chapter concludes with some tentative arguments for the utility of delayed feedback and some directions for future research.

The interaction hypothesis

SLA research on linguistic input to learners and input comprehensibility began in the 1970's. Challenging the idea that grammatical knowledge needs to be developed first in order to be put to use during conversation later, Hatch (1978) argued that knowledge of L2 structures could develop through conversation. Motivated by Hatch's (1978) claims, the focus of SLA research broadened. SLA researchers turned their attention to the structure of conversational interaction and explored the possibility of a causal link between conversational interaction and L2 acquisition. One of these researchers, Mike Long, captured these relationships in the form of a hypothesis, the Interaction Hypothesis (Long, 1981, 1983a, 1983b) and provided a framework for understanding how interaction could facilitate L2 development. There are multiple ways in which interaction, and particularly modified interaction, can facilitate learning. One of these ways is by providing learners with direct or indirect information about what is ungrammatical in the L2 (i.e., negative evidence), for example, and primarily, via negative feedback. The updated version of the Interaction Hypothesis (Long, 1996) highlighted the role of negative feedback and its potential to draw learners' attention to problematic language forms: "negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology and language-specific syntax, and essential for learning certain specifiable L1–L2 contrasts" (Long, 1996, p. 414).

Long (1996) described negative feedback as one of the "attention-focusing devices" (p. 429) that can cause learners to focus on form (Long, 1988, 1991), together with other forms of negative evidence, such as grammar rules or communication breakdowns followed by repairs containing positive evidence. According to Long's theoretical claim, as captured by the idea of *focus on form* (FonF), for acquisition to occur, learners need to attend to language as object while carrying out a meaning-oriented activity. Contrary to *focus on forms*, in FonF, meaning and communication are brought to the forefront and attention to language as object takes place in context. According to Long, the type of negative feedback that can be of special utility to orient the learner to both form and meaning is implicit negative feedback, or recasts:

a reformulation of all or part of a learner's immediately preceding utterance in which one or more non-target-like (phonological, lexical, grammatical, etc.) items are replaced by the corresponding target language versions, and where, throughout, the interlocutors' focus is on meaning, not language as object.

(Long, to appear)

This is the type of negative feedback that Long considered the most advantageous from a psycholinguistic perspective. His claim was based on the type of negative

evidence found in native and non-native conversations, which typically lacks a metalinguistic focus and overt error correction, but which includes positive evidence regarding the target-like form. Long listed a number of potential psycholinguistic advantages of implicit negative feedback in the form of recasts (see Long, to appear, for the most recent contribution). Recasts provide needed information about the target form in context. The interlocutors share a common attentional focus. The learner is vested in the exchange. The learner is probably motivated and attending. The learner has prior comprehension of at least part of the interlocutor's response, which can free up attentional resources that can be allocated to the form of the response, facilitating form-function mapping. Finally, recasts are immediately contingent on the learner's non-target-like form, which means that that the incorrect and the correct form are juxtaposed, facilitating cognitive comparison.

Long's arguments provided a theoretical framework for understanding the utility of negative feedback and particularly of recasts. Since then, two and a half decades ago, negative feedback has been one of the most vibrant areas of research in SLA. Plonsky and Brown (2015), for example, identified 18 unique meta-analyses synthesizing feedback studies. Empirical research has focused on the effectiveness of negative feedback from various perspectives. From the investigation of the effectiveness of negative feedback the research agenda has gradually broadened to the investigation of factors that moderate the effectiveness of feedback. The next section reviews the literature on these factors.

Factors moderating the effectiveness of negative feedback

Some of the factors that moderate the effectiveness of feedback are internal to the learner, such as learner attitudes (Sheen, 2007), language analytical ability (Yilmaz, 2013), implicit and explicit language aptitude (Yilmaz & Granena, 2021), or working memory (Goo, 2016). Some other factors are contextual and external to the learner. These factors are related to how a particular type of instruction is delivered, or the nature of the linguistic feature targeted by the instruction. A wide array of factors under this category has been investigated, including type of feedback (e.g., Ellis et al. 2006), exposure condition (e.g., Yilmaz, 2016), tasks characteristics (e.g., Revesz & Han, 2006), type of linguistic target (e.g., Yilmaz & Yuksel, 2011), and communication mode (e.g., Yilmaz, 2012).

Feedback timing is another learner-external factor that has only recently started to attract the attention of L2 researchers. Researchers have operationalized feedback timing as immediate versus delayed feedback. Feedback is considered immediate if interlocutors provide feedback to learners' non-target-like utterances in the turn immediately following the learners' turn. Feedback is considered delayed, however,

if it is provided at the end of the task, the end of the lesson, or even several days after the task or lesson.

It is important to investigate the role of timing in the effectiveness of feedback because language teachers may be unable to provide immediate feedback for various reasons. For example, in a traditional face-to-face classroom, teachers may opt out of providing negative feedback during a student presentation, fearing that learners may be discouraged to produce output after the feedback or that their motivation to carry out the task may be negatively affected by the feedback. Some scholars (e.g., Harmer, 2007) have recommended using delayed feedback in activities such as oral presentations that require learners to use language communicatively and fluently.

In some online teaching contexts, providing immediate feedback is not even an option because these online language programs do not offer synchronous teacher-student sessions. These courses usually target a high number of students and, even though they may still include tasks involving synchronous interaction between pairs of learners, synchronous sessions between instructors and students may not be possible given the volume of students. In such contexts, delayed feedback is the only possible form of feedback instructors can provide. Therefore, it is crucial for instructors and program coordinators working in such programs to know whether delayed feedback can be a viable alternative to immediate feedback.

In addition, feedback timing research can be a venue for testing the theoretical claims made in the FonF literature (e.g., Doughty, 2001) in relation to the connection between feedback immediacy and the effectiveness of various cognitive processes (e.g., cognitive comparison). The central claim of FonF is that learners' attention should be briefly drawn to formal linguistic features during a communicative or meaning-based activity if an intervention is justified through an observation of a linguistic need (Long, 1991; Long & Robinson, 1998). Negative feedback, mainly non-obtrusive types such as recasts, has been considered an ideal means of implementing FonF. Doughty (2001) has claimed that one of the cognitive mechanisms through which learners take advantage of recasts is cognitive comparison, which involves a mental comparison between intention (the meaning the learner wants to express), output (what the learner actually produces), and input carrying the feedback in working memory. Doughty has argued that it is more likely for this cognitive comparison to be successful if the feedback is provided within a cognitive window of opportunity as the temporal proximity between the error and feedback will allow the learner to keep all necessary components active in working memory. She specifically predicted that "with regard to the timing of the information to be compared, the most efficient means to promoting cognitive comparison would seem to be provision of immediately contingent recasts" (p. 253).

Several studies have looked at the role of timing in the effectiveness of feedback in various contexts. Shintani and Aubrey (2016) focused on the timing of

feedback provided during or after web-based writing tasks. The target form was the hypothetical conditional structure in English. The feedback took the form of immediate and delayed reformulations of learners' errors using the comment box function of Google Docs. The feedback procedure in each timing condition also involved marking the errors on the original document. The results showed an advantage for immediate feedback in the delayed posttest. Other feedback timing studies focused on oral interaction. For example, Quinn's (2014) study followed a pretest-posttest-delayed posttest design with three groups (immediate feedback, delayed feedback, control) and targeted the English passive construction. Learners in all groups carried out three interactional tasks with the researcher and received a hybrid form of feedback including both a prompt and a reformulation. The immediate group's errors were treated immediately after they occurred. The delayed group's errors were treated at the end of each task. Oral production, auditory grammaticality judgment, and error correction tests used to measure learners' improvement did not reveal any statistically significant differences between the immediate and delayed groups.

Another study that investigated the timing of feedback in the context of oral interaction was Li et al. (2016). The study targeted the English past passive. The feedback provided to the immediate group included prompts (i.e., feedback that pushes learners to repair their non-target-like utterances) and reformulation. The delayed group received feedback at the end of the treatment. The feedback they received included an explicit prompt ("You said 'the driver wanted to run away, but he stopped by a policeman.' Can you say it correctly?"), and a reformulation (e.g., "he was stopped"). Li et al.'s (2016) results showed an advantage for the immediate feedback group over the delayed feedback group in one of the outcome measures, a grammaticality judgment test, but not in the elicited imitation task.

Two other studies (Arroyo & Yilmaz, 2018; Henderson, 2020) provided delayed feedback on errors that arose during text-based synchronous computer-mediated communication. Arroyo and Yilmaz (2018) focused on the acquisition of Spanish noun-adjective gender agreement in an experimental study where learners interacted with a native-speaker interlocutor via text chat using Skype. The feedback the learners received in each feedback timing condition involved partial reformulations which were provided via the chat feature of Skype in the immediate-feedback condition and via a written document compiling errors and reformulations in the delayed-feedback condition. The study showed that the immediate-feedback group outperformed the delayed-feedback group in one of the measures (an oral production test), but not in the other measure (a grammaticality judgement test). Henderson (2020) investigated the effect of feedback timing both in face-to-face and text contexts. The acquisitional target of the study was the Spanish subjunctive. The feedback, which included error repetition followed by a recast, was provided

during or after two story-retell tasks that were carried out via Skype. Oral production and multiple-choice receptive tests administered at three different times (pretest, immediate posttest, delayed posttests) to assess linguistic improvement showed an advantage for immediate feedback regardless of the communication mode (face-to-face or text chat).

Except for one study (Quinn, 2014), the studies reviewed above show that immediate feedback is more effective than delayed feedback (e.g., Arroyo & Yilmaz, 2018; Li et al., 2016; Henderson, 2020; Shintani & Aubrey, 2016). It is important to note that these studies, however, investigated face-to-face oral feedback (Li et al., 2016), written feedback, via web-based collaborative writing (Shintani & Aubrey, 2016) or text chat (Arroyo & Yilmaz, 2018), or both face-to-face oral feedback and written feedback via text chat (Henderson, 2020). Research in other modalities is needed to determine the extent to which these results can be generalized. With the aim of addressing this gap, Canals et al. (2021) focused on one such modality, video-based chat, which is used extensively in distance language learning environments. Video-based chat, which is a form of computer-mediated communication, allows for the provision of immediate feedback when the interaction is held synchronously, but also for the provision of video-recorded oral feedback delivered by the instructor asynchronously. The study investigated the relative effectiveness of immediate and delayed negative feedback in this modality on the acquisition of -ing/-ed participial adjectives by fifty-two English-as-a-foreign-language Spanish learners. Learners met with an experimenter online in a one-on-one video-conferencing session to carry out a communicative task. Learners' errors that arose during a communicative task were treated according to their group assignment. The immediate-feedback group received explicit corrections during the task. The delayed-feedback group did not receive any feedback during the task but their task performance was video recorded, and 24 hours later they received an edited video recording of the interaction with feedback instances inserted after their errors. Feedback instances were inserted as video overlays, that is, as a video-on-video effect, immediately after target language errors. The same experimenter recorded herself providing the correction and this appeared in a small window on the screen. The study also included control participants that carried out the treatment tasks without receiving any feedback. The effect of feedback timing on linguistic improvement was measured using an oral production task and a written untimed grammaticality judgment task. There was a considerable difference between the immediate-feedback group and the delayed-feedback and control group in favor of the immediate-feedback group in the grammaticality judgment task in terms of effect sizes. However, these differences were not statistically significant, which casts doubt on the generalizability of the difference.

In the oral production task, both groups statistically outperformed the control group but there was no statistical difference between the two timing conditions, indicating that feedback was effective regardless of its timing. Overall, the results of Canals et al. (2021) for delayed feedback were promising in that they showed that delayed feedback can be at least as effective as immediate feedback in leading to linguistic improvement.

Arguments for the utility of delayed feedback

On July 23, 2020, we contacted Mike Long to ask for his thoughts on the novel approach to the provision of delayed oral feedback we were working on (Canals et al., 2020, 2021). The theoretical and empirical literature on delayed feedback did not seem to support a delayed approach to the provision of feedback from a psycholinguistic perspective. Theoretically, the provision of delayed negative feedback out of context in the form of isolated error-correction pairs did not align with Long's (1988, 1991) principle of FonF. Even if provided in the form of recasts, these reformulations did not share the same psycholinguistic advantages that Long (1996) attributed to recasts over models (i.e., instances of items or structures in ambient input, not triggered by learners' own production). Empirically, the evidence showing that delayed feedback facilitates L2 development was not strong either. Several interactional studies (Arroyo & Yilmaz, 2018; Li et al., 2016; Henderson, 2020; Shintani & Aubrey, 2016) showed that immediate feedback was more effective than delayed feedback. In addition, delayed feedback groups did not consistently outperformed no-feedback control groups, suggesting no significant effects of delayed feedback on L2 development (Li et al., 2016; Quinn, 2014). However, these studies (e.g., Arroyo & Yilmaz, 2018; Li et al., 2016) had provided delayed feedback in the form of error-correction pairs, a procedure that decontextualizes delayed feedback from the broader communicative context (i.e., the task).

The novel technique we investigated in Canals et al. (2021) allowed us to give oral feedback in the context of the interaction where the learner's errors took place. Although still unavoidably delayed in time, we viewed our technique as more comparable to the way immediate feedback is provided and, therefore, as potentially more able to provide a more optimal psycholinguistic environment for L2 learning (Doughty & Long, 2003). However, we lacked a theoretical explanation for its effectiveness and this is why we contacted Mike Long. Regarding the question of whether the technique we used to provide delayed feedback would constitute FonF, Long commented the following:

As you know, the claimed advantages of FonF — especially if delivered in the form of a recast — include the fact that the learner receives the correct version (i) while still focused on meaning, and (ii) soon enough (seconds later) to be able to do what Saxton calls ‘cognitive comparison’ of the deviant and reformulated versions. Receiving negative feedback 10 minutes or two days later, on the other hand, would lose the second of those advantages, and possibly the first, too, depending on how genuinely “re-immersed” in the content of the original conversation you could make the learner.

So at first sight, the procedure you describe sounds like it should be less effective than immediate feedback, due to having lost the benefit of (ii), and potentially also of (i). As I understand it, the delayed provision of video comments inserted in the recording of the original conversation would still constitute FonF. Even though it makes me nervous, FonF can be explicit (more often with beginners than at later levels); the key point is that it must be reactive, which your system is.

(M. Long, personal communication, July 23, 2020)

M. Long (personal communication, July 23, 2020) further compared the learner’s psychological state in focus on forms, FonF, and when receiving delayed negative feedback. He created the following figure (see Table 1), where *D* means a feature that was present in delayed feedback, *D+* a feature that was present in delayed feedback and that was potentially better than FonF, and *D++* a feature that was present in delayed feedback and that was even better than that.

Table 1. The learner’s psychological state in focus on forms, FonF, and when receiving delayed negative feedback (*D* = OK; *D+* = potentially better than FonF; *D++* = potentially even better than that) (M. Long, personal communication, July 23, 2020)

Focus on forms	Focus on form (delayed negative feedback)
decontextualized	contextualized <i>D</i>
proactive	reactive <i>D</i>
learner unmotivated	motivated/vested?
learner often not attending	attending <i>D+</i>
learner having to process form and meaning simultaneously	processing form only, meaning/function understood <i>D</i>
depleted learner attentional resources	freed-up attentional resources <i>D+</i>
non-contingent	contingent <i>D+</i>
target form often unlearnable	target form more likely to be learnable <i>D</i>
cognitive comparison of deviant form and input difficult and unlikely	cognitive comparison of deviant form and input easy and likely <i>D+</i>
noticing and form-function mapping less likely	noticing and form-function mapping more likely <i>D++</i>

As Table 1 shows, Long considered that delayed negative feedback (at least as provided in Canals et al., 2021) conveys needed information about the target language in context, allows the learner to process form because the learner has prior comprehension of the message, and it is reactive (i.e., occurs in response to an ill-formed utterance), which makes the target form more likely to be learnable because it builds on the learner's preexisting (inaccurate or incomplete) knowledge. There are other features of delayed negative feedback that Long viewed as potentially better than FonF because they might lighten the cognitive processing load. These features are that the learner is attending, that attentional resources are freed up and can be allocated to form and form-function mapping, and that the feedback is contingent on deviant learner input, making the cognitive comparison of deviant form and input easier and more likely. Finally, there is one feature in Long's view that makes delayed negative feedback even better than FonF and this is noticing of the interlocutor's response and form-function mapping, which are more likely to occur. Long considered it hard for the learner not to notice negative feedback with our technique because during the playback of the original conversation the learner was attending and under no pressure.

However, one of the psychological advantages of immediate negative feedback was not shared by delayed feedback. This was the fact that, in immediate feedback, the learner receives the correct version while motivated and vested in the exchange. When engaged in meaningful communication, learners are potentially interested in seeing the effect of their utterance on the interlocutor. Although our technique "re-immerses" the learner in the content of the original conversation where the error occurred, the learner is not engaged in meaningful communication when they receive the delayed feedback. This is why Long concluded that we could motivate delayed negative feedback by claiming that it is a way of delivering FonF, but that, compared to immediate (and implicit) negative feedback, delayed feedback was an explicit type of feedback, even if provided in the form of recasts. In other words, the fact that learners can be more focused on language form than on meaning and communication can facilitate that delayed feedback is noticed and, therefore, processed by the learner, but the learning process will be explicit, regardless of the type of feedback, whether more implicit or explicit.

The scenario this creates is an interesting one from the point of view of the cognitive processes that play a role during the provision of feedback. If the type of learning assumed to be associated with the processing of delayed feedback is explicit, this means that the cognitive processes under implicit and explicit *delayed* feedback conditions can also be assumed to be explicit. If so, will implicit and explicit delayed feedback conditions be equally effective? Will they contribute to the same type of L2 knowledge? Will the role of cognitive individual differences

in the two conditions be the same? We are currently investigating these questions as part of a new project, but further research on feedback timing is needed. In the next section, we turn to future research directions in the area of feedback timing.

Future research directions

Feedback timing is an area to which SLA researchers have turned their attention relatively recently and, therefore, with many topics and questions in need of investigation. The following are some of the issues we consider worth exploring.

Type of cognitive processing

Research is needed that investigates the nature of the processing under delayed feedback conditions. The labels *explicit* and *implicit* in the feedback literature have been applied to qualify the type of information presented with the feedback and not the type of cognitive processing that the feedback leads to. In theory, any feedback type could lead to explicit or implicit processing. Feedback has been considered explicit when it provides some kind of metalinguistic information either in the form of accuracy-related statements or in the form of references to metalinguistic rules or terminology. Ideally, explicitness of this sort should be controlled for in an experiment in which immediate feedback and delayed feedback are compared in regard to their effectiveness. However, there is no logical necessity for the explicitness of feedback to correlate with the degree of explicitness of the cognitive processing that it leads to. Two equally explicit feedback types with regard to their informational content may lead to more or less explicit processing depending on task-related factors. Several methods can be used to determine the extent to which different feedback timing conditions lead to explicit or implicit processing. One of these methods involves dissociating the cognitive processes that are engaged by relying on the type of knowledge gauged by the outcome measure used. If there is evidence of gained knowledge on a measure conducive to the retrieval of explicit knowledge, one can infer that the processing was also explicit, based on the assumption that consciously (or explicitly) learned knowledge remains explicit (DeKeyser, 2003). The previous literature has shown that tasks such as elicited imitation, word monitoring, and oral narrative tasks are more likely to tap into implicit knowledge (Ellis, 2005; Vafaei et al., 2017; but see Suzuki & DeKeyser, 2015, for findings showing that elicited imitation is more likely to assess automatized explicit knowledge), whereas tasks such as untimed written grammaticality judgement test tend to tap into explicit knowledge (Ellis, 2005). Research findings showing that a delayed feedback condition is more effective than a no-feedback control group on an explicit measure,

but not on an implicit measure, would provide support to the interpretation that delayed feedback leads only to explicit knowledge, which can be assumed to be the result of explicit processing. Another method that can be used to determine the type of processing that takes place under different timing conditions is the use of stimulated recall. Researchers who would like to use this technique can show a video-recording of the interaction between the learners and their interlocutors after the feedback stage is over and ask learners to report (Mackey & Gass, 2005) what they thought when they performed the task including feedback instances. If the processing was more explicit in the delayed condition, they would be expected to report more noticing of the target form, more noticing of the difference between error and feedback, and more instances of conscious attempts to test hypotheses about how the target form works. A last method that can be used to indirectly shed light on the nature of the cognitive processes under each timing condition is the aptitude-treatment interaction (ATI) paradigm. The ATI paradigm aims to examine how learning outcomes depend on a match, or mismatch, between learners' specific cognitive individual differences and the treatment they receive. The investigation of the relationship between cognitive individual differences and treatments could provide indirect evidence about the learning under different instructional conditions. A positive association between a cognitive individual difference that involves explicit cognitive processes (such as language analytic ability) and learning outcomes under delayed feedback conditions would mean that the cognitive processes engaged by that type of feedback are facilitated by that particular cognitive individual difference and that, as a result, the type of processing involved is the same. This reasoning hinges on the theoretical and empirical support (Granena, 2013, 2019; Yilmaz & Granena, 2021) for the dissociation between cognitive abilities that are purported to be relevant to learn intentionally through reasoning, deliberate hypothesis testing, and memorization and other cognitive abilities that are claimed to be related to cognitive processes labeled in the literature as automatic, associative, nonconscious, and unintentional. For example, Granena (2019) and Yilmaz and Granena (2021) showed that tests hypothesized to measure explicit cognitive abilities were correlated with each other and uncorrelated with tests hypothesized to measure implicit cognitive abilities.

Source of error attribution

There is a fundamental piece of information, currently not available to feedback researchers, which is crucial to determining the type of processing that is involved in delayed feedback. Corrective feedback researchers assume that when an error has been made at one point in time and feedback relevant to that specific error is provided at a later time, the link between the error and feedback is clear to the learner.

This may not be the case. The procedures followed so far to give delayed feedback (e.g., providing an error-feedback list, asking learners to re-make the same error that they made previously during interaction, or in the case of Canals et al., 2021, re-immersing learners in the original situation where the error took place) have always reminded learners of their error. This re-presentation of the error relieves learners from the necessity of bringing their initial error to their memory using their own mental resources. In other words, the learner does not have to retrieve the memory of their error in order to receive and, maybe, take advantage of delayed feedback. This means that there is the possibility for the learner to take advantage of the so-called *delayed* feedback instance based on the error at the delayed stage itself. It is possible, therefore, that having made an error at an earlier time may not play any role in the process of taking advantage of delayed feedback. If that is the case, however, the use of the label *delayed* would not be justified because, from the learner's perspective, there may be no prior instance to be remembered or to refer back to. It is important for future research, therefore, to provide evidence that the way delayed feedback is actually processed is in line with the way it is assumed to be processed in the feedback literature. Otherwise, there is a risk of developing theoretical explanations for the wrong empirical phenomenon or of drawing on the wrong theoretical principles to explain the effectiveness or ineffectiveness of the so-called delayed feedback.

Relative effectiveness of delayed feedback

Research is needed that explores how to optimize the conditions under which delayed feedback is delivered. This can involve investigating different modalities in feedback timing conditions (e.g., text-based vs. video-based), as well as different feedback types and potential interactions between these factors. The Canals et al. (2021) study is an example of an attempt to improve the conditions in which delayed oral feedback is delivered. Another study that tried to improve delayed oral feedback by introducing teacher-initiated student-correction was Quinn (2014). Quinn (2021) argued that this was an attempt to reengage the learner in meaningful communication at the delayed stage and that this could potentially facilitate the development of implicit L2 knowledge. Drawing on the notions of reactivation and reconsolidation, Quinn argued that delayed feedback can reactivate an inaccurate long-term mental representation in procedural memory and that the correction can contribute to update and reconsolidate the inaccurate representation with new (and accurate) information (Else et al., 2018; Lee et al., 2017). Indeed, as we saw earlier in this chapter, a psycholinguistic disadvantage of delayed feedback is the fact that the learner is not vested in a communicative exchange at the time delayed feedback is provided. However, any attempts to compel meaningful use of the language as

part of delayed feedback (see Quinn, 2014, for an example) run the risk of changing the nature of the feedback condition from delayed to immediate and are likely to create experimental and practical challenges. Finding a way for the learner to receive delayed feedback while still focused on meaning and vested in a communicative exchange is no doubt desirable from a psycholinguistic standpoint, but it may be as desirable as impossible.

Conclusion

When it comes to computer- or technology- mediated language learning programs, one typically (and wrongly) thinks of computer-assisted language learning (CALL) materials and learner-computer interactions. In fact, the bulk of research on corrective feedback in the area of CALL has focused on systems and applications that identify, process, and communicate errors to learners, or, in other words, computer-generated corrective feedback (see Heift et al., 2021 for a review). Computer- or technology-mediated corrective feedback, however, does not necessarily, or exclusively, have to be *computer-generated*. Language programs that are delivered online typically offer CALL supplementary course materials that provide computer-generated feedback, but they can also include computer-mediated interactions in which L2 learners, instructors, and/or speakers interact with each other via the computer. In some of these programs, learners interact with each other synchronously (i.e., in real time) via video-conferencing, but instructor's feedback on their performance is provided asynchronously, with delays that can be a week or longer.

In one such context, where delayed feedback was a practical necessity, we conducted a research project on a novel approach to the provision of delayed oral feedback (Canals et al., 2020, 2021) and we asked Mike Long to share his thoughts on our approach. Long (M. Long, personal communication, July 23, 2020) compared the probable psycholinguistic advantages of immediate (implicit) negative feedback and delayed negative feedback and tentatively concluded that delayed feedback as provided in Canals et al. (2021) was a way of delivering FonF, albeit a very explicit procedure. The crucial difference between both timing conditions was that the learner was motivated and vested in the communicative exchange in the immediate-feedback condition but not in the delayed-feedback condition. Based on this difference, Long argued, delayed feedback should be expected to be less effective than immediate feedback. Delayed feedback would be more likely to lead to explicit processing, and, eventually, to explicit knowledge, a type of knowledge the value of which, at least when it comes to its availability for use in spontaneous production, is debated in the SLA literature (Ellis, 2005).

We suggested investigating these claims in future research through methodologies that are commonly used in studies aiming to determine whether implicit language learning is possible. We also proposed that in order to gain a more complete theoretical understanding about how delayed feedback works, future research should investigate the validity of the tacit assumption that learners can cognitively link the negative information provided in the delayed feedback with the original error they made during interaction. Finally, we recommended future research to be conducted to determine the factors that can optimize the effectiveness of delayed feedback. Practically, this research has important pedagogical implications and can inform about those conditions under which delayed feedback can be a good alternative to immediate feedback and, theoretically, it can help shed more light on the cognitive mechanisms involved in the processing of different types of feedback.

Funding

This work was supported by a research grant by the Spanish Ministry of Science and Innovation in the Knowledge Generation program (#PGC2018-093865-A-I00).

References

- Arroyo, D. C., & Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning*, 68, 942–972.
<https://doi.org/10.1111/lang.12300>
- Canals, L., Granena, G., Yilmaz, Y., & Malicka, A. (2020). Second language learners' and teachers' perceptions of delayed immediate corrective feedback in an asynchronous online setting: An exploratory study. *TESL Canada Journal*, 37, 181–209.
<https://doi.org/10.18806/tesl.v37i2.1336>
- Canals, L., Granena, G., Yilmaz, Y., & Malicka, A. (2021). The relative effectiveness of immediate and delayed corrective feedback in video-based computer-mediated communication. *Language Teaching Research*. Advanced online publication.
<https://doi.org/10.1177/13621688211052793>
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Blackwell.
<https://doi.org/10.1002/9780470756492.ch11>
- Doughty, C. J. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139524780.010>
- Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language, Learning & Technology*, 7, 50–80.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
<https://doi.org/10.1017/S0272263105050096>

- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368. <https://doi.org/10.1017/S0272263106060141>
- Elsej, J. W. B., Van Ast, V. A., & Kindt, M. (2018). Human memory reconsolidation: A guiding framework and critical review of the evidence. *Psychological Bulletin*, 144, 797–848. <https://doi.org/10.1037/bul0000152>
- Goo, J. (2016). Corrective feedback and working memory capacity: A replication. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 279–302). John Benjamins. <https://doi.org/10.1075/bpa.3.13g00>
- Granena, G. (2013). Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 105–129). John Benjamins. <https://doi.org/10.1075/llt.35.04gra>
- Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency: Links between LLAMA and Hi-LAB. *Studies in Second Language Acquisition*, 41(2), 313–336. <https://doi.org/10.1017/S0272263118000256>
- Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Pearson.
- Hatch, E. M. (1978). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 401–435). Newbury House.
- Heift, T., Nguyen, P., & Hegelheimer, V. (2021). Technology-mediated corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language learning and teaching* (pp. 226–250). Cambridge University Press. <https://doi.org/10.1017/9781108589789.012>
- Henderson, C. (2020). Perfect timing? Exploring the effects of immediate and delayed corrective feedback, communication mode, and working memory on the acquisition of Spanish as a foreign language (Unpublished doctoral dissertation). Indiana University.
- Lee, J. L. C., Nader, K., & Schiller, D. (2017). An update on memory reconsolidation updating. *Trends in Cognitive Sciences*, 21, 531–545. <https://doi.org/10.1016/j.tics.2017.04.006>
- Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *The Modern Language Journal*, 100, 276–295. <https://doi.org/10.1111/modl.12315>
- Long, M. H. (1981). Input, interaction and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences*, 379, 259–278. <https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>
- Long, M. H. (1983a). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126–141. <https://doi.org/10.1093/applin/4.2.126>
- Long, M. H. (1983b). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5, 177–193. <https://doi.org/10.1017/S0272263100004848>
- Long, M. H. (1988). Instructed interlanguage development. In L. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 115–141). Newbury House.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. *Foreign Language Research in Cross-Cultural Perspective*, 2, 39–52. <https://doi.org/10.1075/sibil.2.07lon>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bahtia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Long, M. H. (to appear). The psycholinguistics of L2 interaction. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics*. Routledge.

- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 15–41). Cambridge University Press.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates.
- Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31, 267–278. <https://doi.org/10.1177/0267658314536436>
- Quinn, P. (2014). Delayed versus immediate corrective feedback on orally produced passive errors in English (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.
- Quinn, P. (2021). The timing of corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language learning and teaching* (pp. 322–340). Cambridge University Press.
- Révész, A., & Han, Z. (2006). Task content familiarity, task type, and efficacy of recasts. *Language Awareness*, 15, 160–179. <https://doi.org/10.2167/la401.o>
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–322). Oxford University Press.
- Shintani, N., & Aubrey, S. (2016). The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *The Modern Language Journal*, 100, 296–319. <https://doi.org/10.1111/modl.12317>
- Suzuki, Y., & DeKeyser, R. (2015). Does elicited imitation measure implicit knowledge? Evidence from the word-monitoring task. *Language Learning*, 65, 860–895. <https://doi.org/10.1111/lang.12138>
- Vafae, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, 39, 59–95. <https://doi.org/10.1017/S0272263115000455>
- Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, 62, 1134–1169. <https://doi.org/10.1111/j.1467-9922.2012.00726.x>
- Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, 34, 344–368. <https://doi.org/10.1093/applin/ams044>
- Yilmaz, Y. (2016). The role of exposure condition in the effectiveness of explicit correction. *Studies in Second Language Acquisition*, 38, 65–96. <https://doi.org/10.1017/S0272263115000212>
- Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research*, 15, 457–477. <https://doi.org/10.1177/1362168811412873>
- Yilmaz, Y., & Granena, G. (2021). Implicitness and explicitness in cognitive abilities and corrective feedback. A double dissociation? *Studies in Second Language Acquisition*, 43(3), 523–550. <https://doi.org/10.1017/S0272263120000601>

CHAPTER 8

A calculus for L1 transfer

William O'Grady

University of Hawai'i at Mānoa

The phenomenon of transfer plays a prominent role, either explicitly or implicitly, in most approaches to SLA, including Focus on Form and the Interaction Hypothesis – to name two lines of inquiry that lay at the heart of Mike Long's scholarship. The central thesis of this chapter is that transfer is best seen as a processing-driven strategy: L2 learners transfer operations from their L1 to the L2, unless those operations are more costly in the L2 than in the L1. Much of the chapter is devoted to the application of this idea to phenomena that arise when adult native speakers of English, Japanese, Korean, Spanish and Catalan go about acquiring a second or third language, producing transfer effects that have been difficult to accommodate in traditional approaches to cross-linguistic influence.

1. Introduction

Mike Long's contributions to the study of second language acquisition are well documented and leave a legacy that will influence the field for decades to come. As Gass, Plonsky, and Huntley (this volume) note in their bibliometric analysis of Mike's influence, his interests encompassed several major lines of research – input, interaction, needs analysis, task-based language teaching, maturational constraints and ultimate attainment. I will focus here on transfer, a topic that interacts closely with those interests and is of broad relevance to the field of L2 research in general.

I will begin in the next section by outlining the approach to transfer that I have in mind and situating it within the larger explanatory framework that has come to be known as emergentism (e.g., MacWhinney & O'Grady, 2015). In the subsequent five sections, I will review the results of several transfer-related studies which, taken together, point to a new perspective on this phenomenon. The chapter ends with some concluding remarks.

2. A brief overview

Well before the study of second language acquisition became a branch of linguistics, scholars were aware of transfer and its impact on the performance of adult L2 learners. An insightful observation along these lines was made more than a century ago by Harold Palmer, an early pioneer in the field:

[t]he adult will perceive dangerous analogies which will lead him astray. The child, whose reasoning faculties are comparatively underdeveloped, will not fall into these traps. (Palmer, 1917, p. 48)

Transfer continues to be a major topic of research in the field of second language acquisition. Indeed, it has become the primary concern of several contemporary approaches to SLA, including highly influential work within generative grammar (for succinct reviews, see Slabakova et al., 2020, p. 13ff and Schwartz & Sprouse, 2021a). I will begin by outlining the assumptions about language and learning that shape the approach to transfer adopted in this chapter.

2.1 Linguistic emergentism

The central tenet of linguistic emergentism is that language is a complex system in the technical sense, which means that its properties are shaped by the interaction of more basic forces. These forces appear to consist largely of processing pressures that seek to minimize the cost of the operations that are required for speech and comprehension (O'Grady, 2005, 2015, 2021; Hawkins, 2004, 2014). One such pressure, on which I focus in this chapter, involves an immediacy advantage.¹

Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

The advantages of Immediacy are easy to see since operations that have this property avoid two potential obstacles to efficiency – delays that increase the burden on working memory and the need to revise a previously assigned interpretation. A simple example of an immediacy effect can be seen in the typology of *wh* questions in a language such as English.

Who did you meet _ yesterday?

1. This insight is widely recognized in the processing literature, where it goes by a number of different names including Maximize Online Processing (Hawkins, 2004, 2015) and the Efficiency Requirement (O'Grady, 2005), among others.

Here, the empty position after the verb must be associated with the utterance-initial *wh* word in order for the sentence to be interpreted as a direct object question. Although this requires a certain amount of extra processing (e.g., Hofmeister & Sag, 2010, p. 380), its cost is minimized by the fact that the *wh* word is available as soon as the presence of a gap becomes evident.

Who did you meet _ ...?


The gap can be interpreted immediately.

Consistent with Immediacy, there is therefore no delay in interpreting the gap once it has been discovered.²

Interestingly, there are no spoken languages in which *wh* words systematically occur to the right of the gap (Petronio & Lillo-Martin, 1997, p. 18; Hawkins, 2014, p. 30).

* _ read the book who?

This prohibition is arguably an Immediacy effect since patterns with sentence-final *wh* words force an obvious delay in the interpretation of the gap.

* _ read the book who?


Immediate interpretation of the gap is not possible.

I do not propose that Immediacy embodies an absolute requirement; in fact, it is well known that exceptions are possible (e.g., Hawkins, 2014, p. 215ff). The key claim is simply that instances of non-compliance add to processing cost, with consequences that can be discerned in phenomena ranging from typology to language acquisition, including the learning of a second language.

2.2 Transfer and processing cost

The leading idea that I will outline in this chapter is that transfer takes place in order to minimize processing cost. I'll formulate this idea as follows, slightly revising the principle proposed by O'Grady (2010, 2012, 2013). (By 'operation' I mean any procedure involved in forming or interpreting a sentence: arranging words in a particular order, resolving a filler-gap dependency, finding an antecedent for a pronoun, and so on.)

2. This generalization is sometimes referred to as the Active Filler Hypothesis, so named because the 'filler' (the *wh* word) actively seeks a gap with which it can be associated (Clifton & Frazier, 1989, p. 292 & p. 297; Wagers & Phillips, 2009, pp. 396–397, and many others).

The Transfer Calculus

L2 learners transfer operations from their L1 – unless those operations are more costly in the L2 than in the L1.

The Transfer Calculus embodies two simple claims.

- i. Transfer takes place for a reason, which is to minimize processing cost by carrying familiar operations in the L1 over to the L2, where feasible.
- ii. Transfer does not occur if it would have the effect of increasing processing cost in the L2 – by violating Immediacy, for example.

On this view, the existence of transfer is a basic fact of second language acquisition, but should not be treated as either the default state of affairs or as an exception. Rather, its occurrence is determined by a fundamental calculus: transfer can occur if there is no resulting increase in processing cost, but not otherwise.

Before evaluating these claims, it is important to recognize that previous research has also recognized a connection between transfer and processing cost – a natural line of inquiry, as Filipović & Hawkins (2013) note.

Learners of a second language (L2) [should] prefer to minimize processing effort when using the grammatical and lexical properties of the L2, just as native speakers do.

An early and still very influential processing-based approach to transfer is the Competition Model pioneered by Bates and MacWhinney (1987). A key component of their theory involves 'cue strength,' which is shaped by the character of the input to which learners are exposed. In the case of English, word order is the most frequently used and most reliable clue for subjecthood, in Japanese case marking has priority, in Arabic animacy is a stronger cue than case marking (MacWhinney, 2008, p. 355), and so on. MacWhinney (2005, p. 57) proposes that, by virtue of their entrenchment, cue rankings for the L1 are initially carried over to the L2, creating transfer effects such as the well-known oversensitivity to word order manifested by native speakers of English learning Japanese.

Another well-known processing-based approach to second language acquisition is the Developmentally Moderated Transfer Hypothesis put forward by Pienemann et al. (2005a; b) as part of what has come to be called Processability Theory.

Transfer is constrained by processability, in particular by the capacity of the L2 learner's language processor... (Pienemann, 2015, p. 136)

On this view, acquisition takes place in phrases before clauses, and in simple clauses before subordinate clauses, reflecting a sort of processing-cost hierarchy: phrase > clause > embedded clause.

Still another approach to transfer, more widely used in studies of L2 phonology than syntax, focuses on the possible role of typological markedness – a notion that is now often taken to be grounded in processing cost (Hawkins, 2004, 2014; O’Grady, 2021). Eckman (2010) summarizes the key insight underlying the markedness theory as follows:

... markedness in second language learning is a measure of relative difficulty and ease of transferability. The less marked a structure is, the easier it is to learn, and the less marked a construction is, the more likely it is to be transferred from the native to the target language. (Eckman, 2010, p. 490)

A classic example of syntactic markedness involves the NP Accessibility Hierarchy proposed by Keenan and Comrie (1977), which predicts (for example) that any language that allows indirect object relative clauses will also allow subject and direct object relative clauses.

Subject > Direct Object > Indirect Object > ...

Subject RC: the student [that _ introduced Harry to Jane]

Direct object RC: the student [that Harry introduced _ to Jane]

Indirect object RC: the student [that Harry introduced Jane to _]

Doughty (1991) provided particularly striking evidence for the relevance of the Accessibility Hierarchy to second language acquisition. Working within a focus-on-form paradigm, she showed that instruction on a marked pattern (such as indirect object relative clauses) suffices to improve learners’ ability to produce less marked patterns (subject and direct object relative clauses). Later work extended these findings to the study of transfer effects (e.g. Hawkins, 2007; Algady, 2013; Xu, 2014; Alroudhan & Ibn, 2016; Alosaimi, 2021).

In the rest of this chapter, I will pursue and extend the processing-based line of inquiry by examining the relevance of the Transfer Calculus to a series of phenomena that have thus far received relatively little attention in the literature. I will begin by considering a curious pattern involving verb deletion.

3. Verb deletion in English and Japanese

A well-known phenomenon in the syntax of ellipsis involves the pattern illustrated below, in which the verb in the second conjunct of the English sentence can be dropped – an instance of a phenomenon often called ‘gapping’ (e.g., Hwang, 2020).

Gapping in English:

I read a book and [Sue ___ a magazine].
 ↑ ↑
 verb gap

The English pattern stands in sharp contrast to its counterpart in SOV languages such as Japanese (and Korean), in which the verb in the first conjunct is dropped.

Gapping in Japanese:

[Watashi-wa	hon-o	___]	[Sue-wa	magajin-o	yon-da].
I-TOP	book-ACC	___	Sue-TOP	magazine-ACC	read-PST
		↑			↑
		<i>gap</i>			<i>verb</i>

The two gapping patterns differ from each other in a fundamental way with regard to the principle of Immediacy, repeated here.

Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

Whereas the Verb–Gap pattern of English is compatible with this preference, its Gap–Verb counterpart in Japanese is not.

Verb–Gap (English)

I read a book and [Sue ___ a magazine].

└──────────────────────────┬↑

The gap can be interpreted immediately.

Gap–Verb (Japanese)

[Watashi-wa hon-o ___] [Sue-wa magajin-o **yon-da**].

I-TOP book-ACC ___ Sue-TOP magazine-ACC read-PST

└ ↘ ?

*Immediate interpretation
of the gap is not possible.*

These differences are highly relevant to the metric on which the Transfer Calculus is based.

The Transfer Calculus

L2 learners transfer operations from their L1 – unless those operations are more costly in the L2 than in the L1.

The Transfer Calculus makes two predictions about gapping, one for English speakers learning Japanese and the other for Japanese speakers learning English. Let us consider each in turn.

3.1 English-speaking learners of Japanese


The prediction that the Transfer Calculus makes for English-speaking learners of Japanese can be formulated as follows:

PREDICTION 1 (VERB DELETION IN JAPANESE)

English-speaking learners should prefer the Verb–Gap pattern of their L1 over the Gap–Verb pattern in Japanese.


The basis for this prediction is not simply that English happens to make use of the Verb–Gap pattern and that its properties are automatically carried over to the second language. It is that the Verb–Gap pattern incurs a lower cost than its Japanese counterpart since the missing verb can be immediately identified at the point where the gap is encountered by simple reference to the first conjunct.

The Verb–Gap pattern:


[NP V NP] and [NP __ NP]

The gap can be interpreted immediately.

By transferring this pattern to Japanese, learners reduce the cost of gapping in their second language, as depicted below.

After transfer to Japanese:

[NP NP V] [NP NP __]

The gap can be interpreted immediately.

Cost-wise, this is an obvious improvement over the Japanese pattern, in which interpretation of the missing verb has to be delayed, in violation of Immediacy.

[NP NP __] [NP NP V]

 ↳?

Immediate interpretation of the gap is not possible.

Data from an experiment conducted by O'Grady (1999) is relevant to the assessment of this scenario.

Participants

75 English-speaking college students studying Japanese as an L2
 (20 second semester, 32 third semester, 23 fourth semester)

Method and materials

- Subjects were asked to judge the acceptability of gapping patterns by rating them on the following scale.

1 2 3 4 5 I don't
Bad Good know

- There were 5 tokens of the Verb–Gap pattern, and 5 tokens of the Gap–Verb pattern.³

Sample Gap–Verb pattern (acceptable in Japanese):

I was asked what Suzuki drank and what Tanaka drank. Here is what I think.

[Suzuki-san-wa biiru-o ___] [Tanaka-san-wa sooda-o **non-da** to] omo-u.
Suzuki-HON-TOP beer-ACC ___ Tanaka-HON-TOP soda-ACC drink-PST that think-PRS

Sample Verb –Gap pattern (unacceptable in Japanese):

I was asked what Suzuki drank and what Tanaka drank. Here is what I think.

*[Suzuki-san-wa biiru-o **non-de**] [Tanaka-san-wa sooda-o ___ to] omo-u.
Suzuki-HON-TOP beer-ACC drink-CON Tanaka-HON-TOP soda-ACC ___ that think-PRS

Results

As predicted, English-speaking learners show a preference for the low-cost Verb–Gap pattern even though it is unacceptable in Japanese. As reported below, that pattern received a mean rating of 3.09 (out of 5) compared to just 2.36 for the Gap–Verb pattern. The difference is statistically significant ($p = .000$).

Gap-Verb pattern (✓ in Japanese)		RATING
[Watashi-wa hon-o ___]	[Sue-wa magajin-o yom-da].	2.36
I-TOP book-ACC ___	Sue-TOP magazine-ACC read-PST	
Verb-Gap pattern (* in Japanese)		
[Watashi-wa hon-o yom-i]	[Sue-wa magajin-o ___].	3.09
I-TOP book-ACC read-CON	Sue-TOP magazine-ACC ___	

Now let us consider the reverse situation involving the acquisition of English by Japanese speakers, for which the Transfer Calculus makes a quite different prediction.

3. HON = honorific, TOP = topic, ACC = accusative, PST = past, CON = connective. Because the JSL learners were not yet fully fluent in Japanese, the context was presented in English; however, the test sentences were presented in Japanese in the standard kana-kanji script.

3.2 Japanese-speaking learners of English

If the Transfer Calculus is right, Japanese speakers should be reluctant to transfer the Gap–V pattern of their first language to English because of its extra cost compared to the V–Gap option offered by the L2.

The Gap-Verb pattern of Japanese transferred to English:

The Verb-Gap pattern in English:

[NP __ NP] and [NP V NP]

↳ ?

Immediate interpretation of the gap is not possible.

The Verb-Gap pattern in English:

[NP V NP] and [NP __ NP]

└──────────────────┘ ↑

The gap can be interpreted immediately.

The Transfer Calculus therefore makes the following prediction.

PREDICTION 2 (VERB DELETION IN ENGLISH)

Japanese-speaking learners should prefer the Verb–Gap pattern of English over the Gap–Verb pattern of their first language.

Data from the afore-mentioned 1999 study sheds light on this issue as well.

Participants

34 Japanese native speakers learning English as an L2
(22 first-year students, 12 third-year students)

Method and materials

- A written acceptability judgment task similar to the one in the first experiment was employed.⁴
- There were 5 tokens of the Verb–Gap pattern, and 5 tokens of the Gap–Verb pattern, as exemplified below.

4. Because the Japanese learners were not yet fully fluent in English, the context was presented in Japanese.

Sample V–Gap pattern (acceptable in English):

Peter asked me what Susan ate and what Harvey ate.

I said that [Susan **ate** a hamburger] and [Harvey ___ a sandwich].

Sample Gap–V pattern (unacceptable in English):

Peter asked me what Susan ate and what Harvey ate.

*I said that [Susan ___ a hamburger] and [Harvey **ate** a sandwich].

Results

As predicted, Japanese-speaking L2 learners show a preference for the low-cost Verb–Gap pattern of English (mean rating 2.33) over the Gap–Verb pattern of their first language (1.75). The difference is statistically significant ($p < .0022$).

Verb–Gap pattern (✓ in English)	RATING
[I read a book] and [Sue ___ a magazine].	2.33

Gap–verb pattern (* in English)	
[I ___ a book] and [Sue read a magazine].	1.75


3.3 Implications

To review, there is an evident asymmetry in our results:

- Native speakers of English prefer the Verb–Gap pattern of their first language when assessing gapping in Japanese.
- Native speakers of Japanese are reluctant to accept the Gap–Verb pattern of their first language when assessing gapping in English.

This asymmetry makes perfect sense from a cost-based perspective. The preference by English speakers for the Verb–Gap pattern in Japanese facilitates the processing of ellipsis in that language, just as it does in English.

The V–Gap pattern transferred to Japanese:

[NP NP V] ... [NP NP ___]


The gap can be interpreted immediately.

In contrast, adopting the Gap–Verb pattern of Japanese in English has the opposite effect, precluding immediate interpretation of the gap.

The Gap–Verb pattern transferred to English:

[NP __ NP] ... [NP V NP]

↳ ?

Immediate interpretation of the gap is not possible.

In sum, the predictions of the Transfer Calculus appear to be borne out in the case of gapping: the likelihood of transfer is driven by a sensitivity to processing cost. A similar effect can be discerned in the very different phenomenon to which we turn next.

4. The interpretation of quantified NPs in English and Korean

A striking difference between English and Korean involves the preferred interpretation of sentences in which negation interacts with a universal quantifier in the direct object position, creating the relationship known as ‘scope.’

English pattern:

The boy didn’t catch **all the fish**. (= ‘The boy caught only some of the fish.’)

Korean pattern:

Sonyen-i **motun punge-lul** an cap-ass-ta.

boy-NOM all fish-ACC not catch-PST-DECL

‘The boy didn’t catch all the fish.’ (= ‘All the fish went uncaught.’)

The two languages differ with respect to the interpretation of these patterns, as described in detail by O’Grady (2013). Once again, the key variable involves Immediacy.

Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

The English preference for the negated (‘not all’) interpretation of the QNP, in which the boy caught just some fish, is fully consistent with Immediacy. At the point at which the processor comes upon the quantified NP (QNP) *all the fish*, the previously encountered negative (*not*) is available – allowing a negated interpretation to be derived without delay.

The negated interpretation of the QNP (English):

The boy didn't catch all the fish.



Negated interpretation of the QNP is immediately available.



CAUGHT NOT CAUGHT

In contrast, we find a very different situation in Korean.

The non-negated interpretation of the QNP (Korean):

Sonyen-i **motun punge-lul** an cap-ass-ta.

boy-NOM all fish-ACC not catch-PST-DECL



The QNP is interpreted before the negative is encountered.

As illustrated here, the QNP *motun punge-lul* 'all the fish' is encountered *before* it can fall under the influence of the negative. In contrast to its English counterpart, the sentence should therefore describe a situation in which all the fish went uncaught. Several experimental studies have confirmed that this interpretation is in fact strongly preferred by native speakers of Korean (Han et al. 2007; Lee, 2009; O'Grady et al. 2009, 2011).

The difference between the English and Korean patterns lays the groundwork for a further test of the Transfer Calculus, repeated here.

The Transfer Calculus

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

Two predictions can be made – one for English-speaking learners of Korean and the other for Korean learners of English. Let us consider each in turn.

4.1 English-speaking learners of Korean

A first prediction involves the interpretation of QNPs in Korean by native speakers of English.

PREDICTION 3 (QNPs IN KOREAN)

English-speaking learners should prefer the non-negated interpretation of the QNP that is also preferred by native speakers of Korean. They should reject the negated interpretation that is dominant in their first language, English.

The basis for this prediction lies in Immediacy. In contrast to English (see the preceding page), the QNP occurs prior to the negative in Korean.

Interpretation of the QNP in Korean:

The boy **all the fish** not caught.

↑

The QNP is interpreted before the negative is encountered; there is no immediate opportunity to derive the negated interpretation.

A negated ‘not all’ reading in this case would call for reinterpretation of the QNP, contra the principle of Immediacy. This in turn would require a more costly interpretive procedure than the one employed in the L1 (English). Transfer is thus ruled out. Put simply, if the Transfer Calculus is right, English speakers are better off interpreting the Korean sentence as if they were Koreans.

The prediction is confirmed by the results of an experiment conducted by O’Grady (2013).

Participants

10 native speakers of English in their late teens or early twenties who were enrolled in a 4th-semester course in Korean as a second language at an American university. (None were heritage learners of Korean.)

Method and materials

Participants were asked to match a test sentence with an appropriate contextual passage, one of which was compatible with a negated interpretation of the QNP and the other with a non-negated interpretation. There were ten test items in all. Here is a sample.

Tom-i **motun chayk-ul** an ilk-ess-ta.

Tom-NOM all book-ACC not read

Context 1 (negated interpretation):

I gave Tom all the books that he was supposed to read, but he read only some of them.

Context 2 (non-negated interpretation):

I gave Tom all the books that he was supposed to read, but he didn’t read any of them.

Participants who prefer the negated interpretation of the QNP should match the sample test sentence with the context on the left, in which not all the books were read. In contrast, those who favor a non-negated interpretation should select the context on the right, in which all the books go unread.

Results

Native speakers of English were strongly adverse to the negated interpretation of the QNP in Korean, favoring the less costly non-negated reading 100% of the time. As predicted, there are no signs of transfer of any sort.

4.2 Korean-speaking learners of English

If the theory of cost-based transfer is right, Korean-speaking learners of L2 English should prefer the non-negated interpretation of the QNP that is also favored in Korean.

PREDICTION 4 (QNPS IN ENGLISH)

Korean-speaking learners should prefer the non-negated interpretation of the QNP that is dominant in their L1 over the negated interpretation that is dominant in English.

This prediction follows from the central tenet of the Transfer Calculus, which is that learners should carry the procedures from their L1 over to the L2 as long as there is no increase in processing cost.

The Transfer Calculus

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

Crucially, the procedure that yields a non-negated interpretation for the QNP in Korean can be applied to English with no increase in cost.

How the QNP is interpreted by Koreans in their L1:

The boy **all the fish** not caught.

↑

The QNP is interpreted here with no regard for the negative.

How the QNP is interpreted by Koreans in L2 English:

The boy didn't catch **all the fish**.

↑

The QNP is interpreted here with no regard for the negative.

Even though *not* precedes the QNP in English and is therefore *available* to derive the negated interpretation, nothing requires that this actually happen. Indeed, children learning English as a first language commonly ignore the negative when interpreting the QNP (Musolino et al., 2000; Musolino & Lidz, 2006). The Transfer Calculus predicts that Korean-speaking learners of English will adopt a similar course of action, at least initially. Put simply, they should interpret English as if it were Korean.

Key evidence comes from an experiment conducted by O'Grady, Kwak, Lee, & Lee (2011).

Participants

42 native speakers of Korean at a university in Seoul (intermediate to high-intermediate learners of English)

Method and materials

- Truth Value Judgment task
- 4 test items in which a negated interpretation of the QNP was favored and 4 in which a non-negated reading should be preferred.

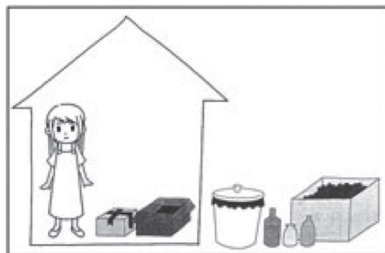


Figure 1. Sample test item
The girl didn't throw out all the boxes.

If the L2 learners adopt the negated ‘not all’ interpretation of the QNP, they should accept the sentence in Figure 1 as a true description of the depicted situation (since not all of the boxes were thrown out). On the other hand, if they adopt the non-negated interpretation (according to which all boxes remain inside), they should reject the sentence as false.

Results

The Korean learners of English showed a strong preference for the non-negated reading of the QNP, rejecting sentences such as the one above as false 72% of the time. When tested on similar items in their native Korean, they rejected the sentence as false 79% of the time. The similarity of the scores in the two conditions strongly suggests that the same interpretive mechanism is at work – consistent with the predicted transfer effect.⁵

4.3 Implications

In sum, as was the case with verb deletion, we see a major and intriguing asymmetry in the workings of transfer in the sentences that involve the interaction of a negative and a universal quantifier.

5. Native speakers of English accept such sentences as *true* 100% of the time (Musolino et al., 2000; Musolino & Lidz, 2006).

- Native English speakers avoid transfer when learning Korean: as we have seen, they favor the non-negated reading of the QNP rather than the ‘not all’ reading typical of English. As explained in Section 4.1, this happens because the ‘not all’ interpretation cannot be derived in Korean without violating Immediacy, making it more costly in the L2 than in the L1.

Interpretation of the QNP in Korean:

The boy **all the fish** not caught.

↑

The QNP is interpreted before the negative is encountered; there is no immediate opportunity to derive the negated interpretation.

- In contrast, native Korean speakers carry over the non-negated interpretation preferred in their native language when interpreting English sentences – a result that suggests transfer.

How can this be? The answer is in fact quite straightforward. Both developmental trajectories are perfectly aligned with the metric underlying the Transfer Calculus, reflecting the simple fact that learners seek to minimize processing cost in their L2.

As we will see next, a different scopal phenomenon – this time in Japanese – offers still further support for a theory of cost-related transfer.

5. The interpretation of indefinite NPs in English and Japanese

The two previous sections of this chapter have dealt with what might be called ‘bi-directional transfer’ since they focus on the comparison of mirror-image situations: the acquisition of language x by speakers of language y and the acquisition of language y by speakers of language x . In this section, I examine the more usual context for the study of transfer, which involves the acquisition of a single particular second language by native speakers of another particular language.

5.1 The phenomenon

The phenomenon on which I will focus here involves the interpretation of indefinite pronouns in patterns that also contain a universal quantifier, following the lead of Marsden (2005). The following English sentence is a case in point.

Someone stroked	every cat.
↑	↑
indefinite pronoun	universal quantifier

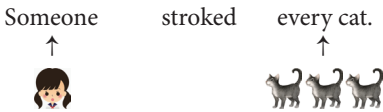
The dominant interpretation of this sentence is that a particular person stroked a number of cats. This reading is derived by interpreting each NP as it is encountered, consistent with Immediacy.

Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

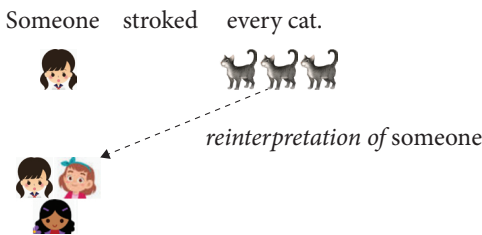
According to this scenario, the processor first encounters *someone*, which it takes to pick out a single person; later, it comes upon *every cat*, which it takes to denote multiple felines. Each NP is interpreted as soon as it is encountered, and neither influences the interpretation of the other. I will call this the ‘specific interpretation’ since *someone* is taken to refer to a single particular individual.

The specific interpretation of an indefinite NP:



A second and less common interpretation is also possible: different people are involved in stroking the cats. As depicted below, this reading involves reinterpretation of the indefinite NP upon encountering the QNP *every cat*, so that there is (for instance) a different girl for each cat. This is often called the ‘distributed interpretation.’⁶

The distributed interpretation of an indefinite NP:



By the standards of Immediacy, the distributed interpretation should be more costly since it requires reinterpretation of the subject NP. This seems to be correct: as the literature on quantifiers notes, even native speakers find it difficult to derive the distributed reading.

6. Drawing on a different terminology, Marsden (2005) uses the term ‘forward scope’ for the specific interpretation and ‘inverse scope’ for the distributed interpretation.

[the distributed interpretation] is a marked option: It is often very hard to obtain and it requires a strong discourse motivation. (Reinhart, 1997, p. 370)

Moreover, there is evidence that (all other things being equal) the distributed interpretation takes longer to process (Anderson, 2004) and is actually prohibited in many languages. (Keenan, 1974, p. 301–302, 1976, p. 319; Hawkins, 2004, p. 17)

Japanese is one of the many languages that do not permit a distributed interpretation in the patterns under consideration.

Dareka-ga dono neko-mo nade-ta.
someone-NOM every cat-FOC stroke-PST

'Someone stroked every cat.'

(sole interpretation = 'One specific person stroked all the cats.')

The contrast between English and Japanese raises an obvious question: how do native speakers of English go about interpreting Japanese sentences like the one above?

5.2 The interpretation of *dareka* by native speakers of English

On the view I propose, the answer for the question at hand should follow from the Transfer Calculus, repeated here.

The Transfer Calculus

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

The specific interpretation of an indefinite NP should therefore be relatively easy in Japanese, just as it is in English, since Immediacy is satisfied in this case in both languages. As illustrated below, the specific interpretation in Japanese involves picking out a single girl upon encountering *dareka-ga* 'someone' and a set of cats upon encountering *dono neko-mo* 'every cat' – with neither NP affecting the interpretation of the other.

The specific interpretation of an indefinite NP in Japanese:

(compatible with Immediacy, as in English):

Dareka-ga dono neko-mo nade-ta.
someone-NOM every cat-FOC stroke-PST



'Someone stroked every cat.' (= 'One specific person stroked all the cats.')

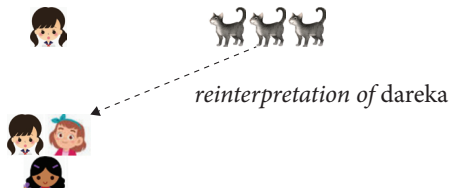
A distributed interpretation – were it allowed in Japanese – would be more costly than the specific interpretation, but no more costly than the distributed reading in

English. In fact, the relevant operations would be essentially identical to those used in the earlier English example.

The distributed interpretation of an indefinite NP in Japanese:

(incompatible with Immediacy, as in English):

Dareka-ga dono neko-mo nade-ta.
 someone-NOM every cat-FOC stroke-PST



Given these facts, the Transfer Calculus predicts that the procedures required for both the specific interpretation and the distributed interpretation should be transferable from English to Japanese.

PREDICTION 5 (INDEFINITE NPs IN JAPANESE)

English-speaking learners of Japanese should permit both the specific and distributed interpretation of indefinite NPs in Japanese.

Marsden (2005) conducted an experiment to investigate how native speakers of English interpret *dareka* in the sort of sentences that we are considering.

Participants

- 18 intermediate English-speaking learners of Japanese as a second language
- 9 advanced English-speaking learners of Japanese as a second language

Method and materials

Participants were asked to indicate how well a particular sentence matched one or the other of the pictures below by ranking it on a scale of 0 to 3. (A score of 1.5 or higher was taken to indicate acceptance of the relevant interpretation.)



Dareka-ga dono neko-mo nade-ta.
 someone-NOM every cat-FOC stroke-PST

Figure 2. Sample pictures from Marsden's study (p. 123)

Results

The table below summarizes Marsden's results (p. 195) for English-speaking learners of Japanese.

Table 1. Mean ratings (on a scale of 0 to 3)

Participants	Specific	Distributed
Intermediate	2.74	1.61
Advanced	2.82	1.38
Native speakers of Japanese	2.26	0.69

5.3 Implications

Two findings support the prediction of the Transfer Calculus:

- i. The L2 learners assign very high ratings to the specific interpretation of *dareka*. Since that reading complies with Immediacy and therefore has a low cost in both languages, the underlying interpretive operation should be carried over from English to Japanese. This seems to be the case.
- ii. The L2 learners assign moderately high ratings to the distributed interpretation of *dareka*, especially compared to the rating of 0.69 that it received from native speakers of Japanese. Although this interpretation is costly in English, there is (as noted previously) no reason to think that it would be inherently more costly in Japanese if it were allowed there. Transfer should therefore take place, as it indeed does.

There is nonetheless a potential mystery here. The ratings that native speakers of English assign to the specific interpretation in Japanese are very close to the one that they assign to the comparable interpretation in English (2.68); see Table 2.

Table 2. Mean ratings for the specific interpretation (Marsden, p. 195)

Language and participants	Rating
Intermediate L2 learners judging Japanese	2.74
Advanced L2 learners judging Japanese	2.82
Native speakers of English judging English	2.68

In contrast, the ratings that they assign to the distributed interpretation in Japanese (1.61 and 1.38) are substantially lower than the one that they give for that interpretation in English (1.96); see Table 3.

Table 3. Mean ratings for the distributed interpretation (Marsden, p. 195)

Language and participants	Rating
Intermediate L2 learners judging Japanese	1.61
Advanced L2 learners judging Japanese	1.38
Native speakers of English judging English	1.96

Why is this the case?

5.4 The Weakness Corollary

As I see it, the puzzle is best subsumed under a more general question, which can be formulated as follows:

Why are costly operations more difficult to transfer than their less costly counterparts? The answer, I believe, lies in a supplementary principle that, like the Transfer Calculus, is motivated by processing pressures.⁷

The Weakness Corollary

Weak L1 routines become even more feeble when carried over to an L2.

This idea fits well with the widely accepted finding that processing in a second language is slower and less efficient than in one's first language.

... what do studies of the aging brain reveal about L2 acquisition and processing? From the cognitive literature, we learn that the associative memory and incremental learning elements of language learning are steadily compromised by age, as are the working memory and processing speed components of language processing and production. (Birdsong, 2006, p. 34)

For all three factors (working memory capacity, decoding ability, and speed of processing) there is evidence that late L2 learners have poorer abilities than native speakers. (McDonald, 2006, p. 385)

If this is right, then the transferred version of the procedure that English-speaking learners use for the distributed interpretation should be doubly weak. It is costly in the learners' first language, and the computational resources required to implement it in the second language are compromised.

7. I use the term 'corollary' here in the non-technical sense of the Merriam-Webster dictionary: 'something that naturally follows; something that incidentally or naturally accompanies' another proposition.

Taken together, these considerations lead to the following prediction by the Weakness Corollary.

PREDICTION 6 (DISTRIBUTED INTERPRETATION OF INDEFINITE NPs IN JAPANESE)
 English-speaking learners should rate the distributed interpretation of Japanese *dareka* less favorably than the distributed interpretation of English *someone*.

Marsden's results appear to confirm this prediction, as the data in Table 1–3 show: there is a strong tendency to accept the specific interpretation of the indefinite pronoun and a weaker tendency to permit the distributed reading.⁸

In sum, the developmental facts that surround the interpretation of indefinite NPs in Japanese fit well with the Transfer Calculus.

- English-speaking learners welcome the specific interpretation of *dareka* in Japanese, since that reading involves the same low-cost interpretive procedure that is used in English; there is no additional cost when it is employed in Japanese.
- English-speaking learners show low-to-marginal rates of acceptance for the distributed reading in Japanese, for which their first language provides a costly and little-used interpretive operation. Although that procedure is transferred to Japanese, it is less accessible in the second language, consistent with the Weakness Corollary.

As we will see next, the Weakness Corollary has a role to play in a quite different phenomenon – this one involving the acquisition of Mandarin by native speakers of Japanese.

6. The interpretation of null arguments in Japanese and Chinese

A well-documented feature of many languages is the possibility of null arguments, especially in subject position. Although unpronounced subjects were once thought to be essentially alike in all languages that allowed them, later work uncovered striking differences that open the door to intriguing lines of inquiry in the field of second language research. One such investigation is particularly relevant to the view of transfer that I have been developing.

8. A third effect may also be in play. Because L2 learners never encounter the distributed interpretation in Japanese (where it is banned), the underlying operation will never be activated – raising the possibility that it might wane to the point of inaccessibility. Marsden's data suggests that something like this may in fact happen. Whereas learners with intermediate proficiency in Japanese tended to permit the distributed interpretation at least some of the time, six of the nine speakers with advanced proficiency 'consistently rejected' it (p. 208). Significantly, members of that latter group had spent substantially more time in Japan, thereby increasing the period during which the distributed derivation went unactivated.

6.1 The phenomenon

Monou (2013) reports on a little-known fact that arises in contexts such as the following in Japanese.

San-nin-no keisatsu-ga Sato-san-no ie-ni kita.
 three-CL-GEN police-SUBJ Sato-HON-GEN house-to came
 ‘Three police officers came to Ms Sato’s house.’

pro Yamada-san-no ie-ni-mo kita.
 Yamada-HON-GEN house-to-also came
 ‘[They] also came to Ms Yamada’s house.’

On one interpretation, the null subject pronoun (represented as *pro*) in the second sentence refers to the set of police officers to which reference is made in the first sentence. I will refer to this as the ‘definite’ interpretation, which is illustrated in the figure below by depicting the same set of police officers twice – once at Ms Sato’s house and once at Ms Yamada’s house.

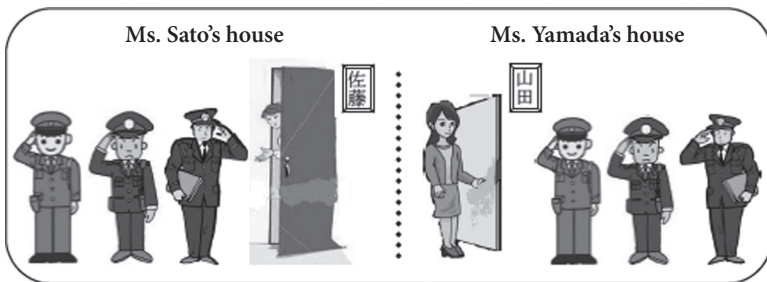


Figure 3. The definite interpretation as depicted by Monou

A second reading, which I will call the ‘indefinite’ interpretation, is more demanding since it requires the construction of a new set of police officers, as depicted below.

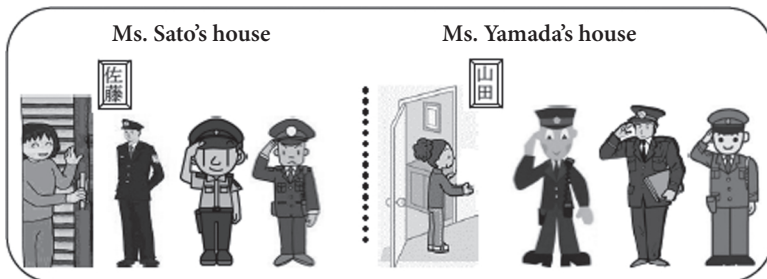


Figure 4. The indefinite interpretation as depicted by Monou

In contrast to the definite interpretation, this second reading does not satisfy Immediacy since the reference of the null pronoun cannot be determined by simply calling up the set of police officers alluded to in the first sentence of the test item. An entirely new set of officers must be conjured.

The definite interpretation (the referents of *pro* are immediately available):

pro Yamada-san-no ie-ni-mo kita.
pro Yamada-HON-GEN house-to-also came

immediate interpretation ↙
*by reference to the previously
 mentioned set of police officers*

The indefinite interpretation (the referents of *pro* are not immediately available):

pro Yamada-san-no ie-ni-mo kita.
pro Yamada-HON-GEN house-to-also came
 ?↙

*Immediate interpretation
 is not possible; a new set of
 police officers must be constructed.*

We can therefore conclude that the indefinite interpretation is more costly – an assessment for which there is also typological evidence. As Monou notes (p. 96), some null-subject languages allow only the definite interpretation, but no language allows only the indefinite interpretation.

One language that allows just the definite interpretation is Mandarin, as illustrated in the following example from Monou (p. 9).

Sān-gè jǐngchá lái-le Zuǒtáng jiā.
 three-CL police officer come-ASP Ms Sato's house
 'Three police officers came to Ms Sato's house.'

pro yě lái-le Shāntián jiā.
 also come-ASP Ms Yamada's house
 'They also came to Ms Yamada's house.' [definite interpretation only]

A question now arises: what can be expected to happen when Japanese speakers learn Mandarin as a second language?

6.2 The interpretation of Chinese *pro* by native speakers of Japanese

Given the Transfer Calculus and the Weakness Corollary, two predictions can be made about how native speakers of Japanese will go about interpreting *pro* in Mandarin.

PREDICTION 7 (DEFINITE INTERPRETATION OF *PRO* IN MANDARIN)

Consistent with the Transfer Calculus, Japanese-speaking learners should accept the definite interpretation of *pro* in Mandarin since the relevant interpretive procedure can be carried over from Japanese at no additional cost.

PREDICTION 8 (INDEFINITE INTERPRETATION OF *PRO* IN MANDARIN)

Consistent with the Weakness Corollary, Japanese-speaking learners should rate the more costly indefinite interpretation of *pro* in Mandarin substantially lower than in their first language.

We are able to test these predictions with the help of data from an experiment conducted by Monou for a different purpose.

Participants

22 undergraduate students at a university in Tokyo, all of whom were native speakers of Japanese. On average, they had been studying Mandarin at their university for 11 months. A pre-test revealed that all 22 participants accepted both the definite and indefinite interpretations of *pro* in their first language.

Method and materials

- Truth Value Judgment Task
- 5 sets of test items such as the one illustrated above, each of which was matched with pictures to test its compatibility with a definite interpretation and an indefinite interpretation.

Results

Monou (p. 36) reports a sharp contrast in the rate at which the two interpretations were accepted by the Japanese-speaking learners of Mandarin.

Table 4. Acceptance rates for the two interpretations

Definite interpretation	Indefinite interpretation
(same 3 police officers)	(different set of 3 police officers)
86.4%	35.5%

Not only is the acceptance rate much lower for the indefinite interpretation, 10 of Monou's 22 participants rejected the truth of all the Mandarin test items that occurred in contexts calling for the indefinite reading (p. 41).

6.3 Implications

Monou's results are expected if the Transfer Calculus interacts with the Weakness Corollary in the way that I have proposed, creating the following developmental scenario.

- Consistent with the Transfer Calculus, the low-cost and frequently instantiated procedure associated with the definite interpretation is carried over to Mandarin since it is no more costly in that language.
- Consistent with the Weakness Corollary, the high-cost and infrequently activated procedure associated with the indefinite interpretation is transferred to the second language in a very weak form, where it is highly susceptible to demise.

This is exactly what Monou's findings appear to show.

7. A note on third language acquisition

In recent years, the study of transfer has moved in a new and challenging direction to include the acquisition of third languages (e.g., Rothman, González, & Puig-Mayenco, 2019; Puig-Mayenco, González, & Rothman, 2020; Schwartz & Sprouse, 2021a; b). The key question in this endeavor turns on which of two earlier acquired languages has the greater influence on the L3 – a matter to which I will now turn.

7.1 The phenomenon

I will focus here on an intriguing contrast involving negative concord – a phenomenon that requires a negative pronoun to be accompanied by a sentential negative, as in the following example from Catalan.⁹

9. In fact, the situation appears to be somewhat more complex. Based on experimental data, Déprez et al. (2015) report that one 'population' of Catalan speakers treats the sentential negative in this pattern as optional.

Pre-verbal negative pronoun with a sentential negative in Catalan:

Ningú no beu cafè.
 nobody not drink coffee
 ‘Nobody drinks coffee.’

Despite the presence of two negatives, the sentence has a ‘single-negative’ interpretation: nobody drinks coffee.

In contrast, the comparable pattern in Spanish has a ‘double-negative’ interpretation that essentially cancels negation.

Pre-verbal negative pronoun with a sentential negative in Spanish:

Nadie no bebe café.
 nobody not drink coffee
 ‘There is nobody that does not drink coffee.’ (= ‘Everyone drinks coffee.’)

A question that now arises has to do with how bilingual speakers of Catalan and Spanish will behave when learning English as a third language. In particular, which pattern of negation will they transfer to English – the one from Catalan or the one from Spanish? Rothman (2015) suggests that the key variable is the ‘structural proximity’ of the three languages. According to his Typological Proximity Model, transfer comes from the previously learned language that is typologically most similar to the L3 (see, e.g., Rothman et al., 2019, p. 154ff).

Puig-Mayenco and Rothman (2020) argue that Catalan is typologically more like English than is Spanish, based on certain phonological and phonotactic similarities identified by Puig-Mayenco and Marsden (2018, p. 508). This leads to the following prediction:

PREDICTION OF THE TYPOLOGICAL PROXIMITY MODEL

Catalan-Spanish bilinguals should transfer features of Catalan rather than Spanish to English, regardless of which language they learned first.

7.2 The acquisition of English negation by Catalan and Spanish bilinguals

Puig-Mayenco & Rothman tested their prediction with the help of an experiment involving two groups of Catalan-Spanish bilinguals who had just completed a two-month English-language course specifically designed for beginners.

Participants

- 35 L1-Catalan, L2-Spanish bilinguals (mean age = 54.4)
- 24 L1-Spanish, L2-Catalan bilinguals (mean age = 52.1)

Method and materials

In order to determine which of the two languages has more influence on the acquisition of English as a third language, Puig-Mayenco and Rothman tested the two groups of learners on unfamiliar sentences such as the following.

Nobody doesn't drink coffee.

The task involved choosing between a picture in which no one is drinking coffee and a picture in which there is no one who is *not* drinking coffee.

The English sentence permits only a double-negative interpretation ('There's nobody that doesn't drink coffee'), but this is irrelevant. What matters for the purposes of the experiment is what the learners *think* that the sentence means.

Results

Table 5 presents Puig-Mayenco & Rothman's results.

Table 5. Preferred interpretation of the English sentences

Participants	Single negative	Double negative
Interpretation ^a	Interpretation ^b	
L1-Catalan, L2-Spanish	90%	10%
L1-Spanish, L2-Catalan	78%	22%

a 'Nobody drinks coffee.'

b 'There's nobody that doesn't drink coffee.'

Crucially, both the L1-Spanish speakers and the L1-Catalan speakers prefer the single-negative interpretation ('Nobody drinks coffee') – the reading that is associated with the corresponding Catalan sentence rather than its Spanish counterpart.


7.3 Implications

Puig-Mayenco and Rothman interpret their result as evidence for transfer from Catalan, in support of their Typological Proximity Model. But their conclusion is not without problems, since a precise metric for typological proximity remains to be developed. Interestingly, Immediacy offers an alternative explanation for why the single-negative interpretation is preferred.


Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

The key observation is that only the single-negative reading of the English sentence allows an immediate and stable interpretation of the negative pronoun *nobody*.¹⁰ (I use the symbol \emptyset to indicate a null set.)

Nobody doesn't drink coffee.
 \emptyset 
 ↑
fully interpreted here

In contrast, the double-negative reading is derived from the interaction of *nobody* with *not*, which cancels negation to give the interpretation in which everyone drinks coffee.

Nobody doesn't drink coffee.
 \emptyset 
 ↙
*reinterpretation
 of the subject*

As illustrated here, derivation of the double-negative reading does not comply with Immediacy since the interpretation of the entire sentence has to be restructured after encountering the second negative. (Instead of meaning ‘Nobody drinks coffee,’ it has to be interpreted as ‘Everybody is a coffee-drinker.’)

On the view just outlined, the preference for Catalan-based transfer simply reflects the fact that the single-negative interpretation is the less costly option. The key intuition can be captured by adopting the following extension to the Transfer Calculus.

The L3 corollary

In the case of competing ‘source languages,’ the best candidate for transfer is the operation that incurs the least cost in the L3, all other things being equal.

If this idea is on the right track, third-language acquisition is subject to the same processing pressures that shape transfer in the more widely studied L2 context – obviously a desirable state of affairs.

10. Given its low processing cost, one might ask why the sentence does not have a single-negative interpretation in English. The answer is that English uses an even simpler form to express that meaning: *Nobody drinks coffee*.

8. Concluding remarks

Transfer has long occupied a prominent place in the study of second language acquisition. Indeed, it is often treated as the single most important tool for addressing the foundational puzzle of why second language learners make the particular types of errors that they do.

Crucially, though, transfer itself needs to be explained. We need to understand why it exists and why it operates in the way that it does. These priorities have been the primary concern of this chapter. The idea that I have proposed seeks to incorporate the study of transfer into a larger emergentist program devoted to processing-based explanations for the many puzzles that arise in the study of syntax and development.

If this line of inquiry is on the right track, then it makes sense to explore the possibility that transfer too is shaped by processing pressures. I have illustrated how this might work by focusing on the role of Immediacy in the operations that produce speech and derive interpretations. As I have tried to demonstrate, this approach offers an explanation for a wide range of developmental tendencies in second language acquisition.

- Why native speakers of English prefer the Verb–Gap pattern of their first language in learning Japanese.
- Why native speakers of Japanese do not prefer the Gap–Verb pattern of their first language in learning English.
- Why native English speakers reject the negated interpretation of QNPs that is preferred in their L1 when learning Korean.
- Why native Korean speakers prefer the non-negated interpretation of QNPs that is preferred in their L1 when learning English.
- Why native English speakers are reluctant to accept the distributed interpretation of indefinite NPs in Japanese.
- Why native Japanese speakers are reluctant to accept an indefinite interpretation for *pro* in Mandarin.
- Why bilingual speakers of Catalan and Spanish are more likely to transfer a feature of Catalan rather than Spanish to a particular pattern of negation in L3 English.

All these things happen for the same reason: as predicted by the Transfer Calculus and the two related corollaries, learners draw on the resources of a previously acquired language to reduce the cost of the operations required to learn and use a new language.

The result is the complex web of effects and non-effects that we have been considering – a small fraction of the larger reality of second language acquisition.

References

- Algady, D. (2013). The acquisition of relative clauses: How do adult second language learners of Arabic do it? (Unpublished doctoral dissertation). University of Wisconsin–Milwaukee.
- Alosaimi, N. (2021). An empirical study of resumptive pronouns in Arabs' acquisition of English restrictive relative clauses. *International Journal of English Language and Linguistics Research*, 9(2), 8–31.
- Alroudhan, H. E., & Ibn Saud, A. M. (2016). The acquisition of English restrictive relative clauses by Arab adults EFL learners. *Advances in Language and Literary Studies*, 7(1), 33–53. <https://doi.org/10.7575/aiac.all.v.7.n1p.33>
- Anderson, C. (2004). The structure and real-time comprehension of quantifier scope ambiguity (Unpublished doctoral dissertation). Northwestern University.
- Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Lawrence Erlbaum Associates.
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56(s1), 9–49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Kluwer. https://doi.org/10.1007/978-94-009-2729-2_8
- Dépres, V., Tubau, S., Cheylus, A., & Espinal, M. T. (2015). Double negation in a negative concord language: An experimental investigation. *Lingua*, 163, 75–107. <https://doi.org/10.1016/j.lingua.2015.05.012>
- Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 431–469. <https://doi.org/10.1017/S0272263100010287>
- Eckman, F. (2010). Linguistic typology and second language acquisition. In J. J. Sung (Ed.), *The Oxford handbook of linguistic typology* (pp. 485–494). Oxford University Press.
- Filopović, L., & Hawkins, J. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics*, 51, 145–176. <https://doi.org/10.1515/ling-2013-0005>
- Gass, S., Plonsky, L., & Huntley, L. (2022). Taking the Long view: A bibliometric analysis. In A. G. Benati & J. Schwieter (Eds.), *Second language acquisition theory: The legacy of professor Michael Long*.
- Han, C., Lidz, J., & Musolino, J. (2007). V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry*, 38, 1–48. <https://doi.org/10.1162/ling.2007.38.1.1>
- Hawkins, J. (2004). *Efficiency and complexity in grammar*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199252695.001.0001>
- Hawkins, J. (2007). Acquisition of relative clauses in relation to language universal. *Studies in Second Language Acquisition*, 29, 337–344. <https://doi.org/10.1017/S0272263107070192>
- Hawkins, J. (2014). *Cross-linguistic variation and efficiency*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199664993.001.0001>
- Hofmeister, P., & Sag, I. (2010). Cognitive constraints and island effects. *Language*, 86, 366–415. <https://doi.org/10.1353/lan.o.0223>
- Hwang, H. (2020). A contrast between VP-ellipsis and gapping in English: L1 acquisition, L2 acquisition, and L2 processing (Unpublished doctoral dissertation). Department of Second Language Studies, University of Hawai'i at Mānoa.

- Keenan, E. (1974). The functional principle: Generalizing the notion of 'subject of'. In M. W. La Galy, R. A. Fox, & A. Bruck (Eds.), *Proceedings of the 10th Regional Meeting of the Chicago Linguistic Society* (pp. 298–309). CLS.
- Keenan, E. (1976). Towards a universal definition of 'subject'. In C. Li (Ed.), *Subject and topic* (pp. 303–33). Academic Press.
- Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry*, 8, 63–100.
- Lee, S. (2009). Interpreting scope ambiguity in first and second language processing: Universal quantifiers and negation (Unpublished doctoral dissertation). University of Hawai'i at Mānoa.
- MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.
- MacWhinney, B. (2008). A unified model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–370). Routledge.
- MacWhinney, B. & O'Grady, W. (Eds.). (2015). *The handbook of language emergence*. Wiley-Blackwell. <https://doi.org/10.1002/9781118346136>
- Marsden, H. (2005). Quantifier scope in non-native Japanese: A comparative interlanguage study of Chinese-, English- and Korean-speaking learners (Unpublished doctoral dissertation). Durham University.
- McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 44, 381–401. <https://doi.org/10.1016/j.jml.2006.06.006>
- Monou, T. (2013). Restrictive second language development paths: Evidence from subject ellipsis constructions in Japanese and Mandarin (Unpublished doctoral dissertation). Graduate School of Human Relations, Keio University.
- Musolino, J. & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics*, 44, 817–852. <https://doi.org/10.1515/LING.2006.026>
- Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, 38, 1–32. <https://doi.org/10.1515/ling.38.1.1>
- O'Grady, W. (1999). Gapping and coordination in second language acquisition. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 141–157). John Benjamins. <https://doi.org/10.1075/lald.20.11ogr>
- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410612571>
- O'Grady, W. (2010). Language acquisition without an acquisition device. Invited plenary talk to the Second Language Research Forum. October 2010, University of Maryland.
- O'Grady, W. (2012). Language acquisition without an acquisition device. *Language Teaching*, 45, 116–30. <https://doi.org/10.1017/S026144481000056X>
- O'Grady, W. (2013). The illusion of language acquisition. *Linguistic Approaches to Bilingualism*, 3, 253–85. <https://doi.org/10.1075/lab.3.3.01ogr>
- O'Grady, W. (2015). Anaphora and the case for emergentism. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 100–122). Wiley-Blackwell. <https://doi.org/10.1002/9781118346136.ch4>
- O'Grady, W. (2021). *Natural syntax: An emergentist primer* (2nd ed.). Retrieved on 8 February 2022 from <http://ling.hawaii.edu/wp-content/uploads/OGrady-Natural-Syntax-2ndEd-2021.pdf>

- O'Grady, W., Kwak, H.-Y., Lee, O., & Lee, M. (2011). An emergentist perspective on partial language acquisition. *Studies in Second Language Acquisition*, 33, 223–245. <https://doi.org/10.1017/S0272263110000744>
- O'Grady, W., Lee, M., & Kwak, H.-Y. (2009). Emergentism and second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 69–88). Emerald Press.
- Palmer, H. (1917). *The scientific study & teaching of languages*. George Harrap & Co.
- Petronio, K., & Lillo-Martin, D. (1997). Wh-movement and the position of Spec-CP: Evidence from American Sign Language. *Language*, 73, 18–57. <https://doi.org/10.2307/416592>
- Pienemann, M. (2015). An outline of Processability Theory and its relationship to other approaches to SLA. *Language Learning*, 65, 123–51. <https://doi.org/10.1111/lang.12095>
- Pienemann, M., Di Biase, B., & Kawaguchi, S. (2005a). Extending Processability Theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of Processability Theory* (pp. 199–251). John Benjamins. <https://doi.org/10.1075/sibil.30.09pie>
- Pienemann, M., Di Biase, B., Kawaguchi, S., & Hakansson, G. (2005b). Processing constraints on L1 transfer. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 128–153). Oxford University Press.
- Puig-Mayenco, E., & Rothman, J. (2020). Low proficiency does not mean *ab initio*: A methodological footnote for linguistic transfer studies. *Language Acquisition*, 27, 217–226. <https://doi.org/10.1080/10489223.2019.1677668>
- Puig-Mayenco, E. & Marsden, H. (2018). Polarity-item *anything* in L3 English: Where does transfer come from when the L1 is Catalan and the L2 is Spanish? *Second Language Research*, 34, 487–515. <https://doi.org/10.1177/0267658317747926>
- Puig-Mayenco, E., Alonso González, J., & Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research*, 36, 31–64. <https://doi.org/10.1177/0267658318809147>
- Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy*, 20, 335–397. <https://doi.org/10.1023/A:1005349801431>
- Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition*, 18, 179–90. <https://doi.org/10.1017/S136672891300059X>
- Rothman, J., González Alonso, J., & Puig-Mayenco, E. (2019). *Third language acquisition and linguistic transfer*. Cambridge University Press. <https://doi.org/10.1017/9781316014660>
- Schwartz, B., & Sprouse, R. (2021a). Making models, making predictions. *Linguistic Approaches to Bilingualism*, 11, 116–129. <https://doi.org/10.1075/lab.00032.sch>
- Schwartz, B., & Sprouse, R. (2021b). The Full Transfer/Full Access model of L3 cognitive states. *Linguistic Approaches to Bilingualism*, 11, 1–29. <https://doi.org/10.1075/lab.20055.sch>
- Slabakova, R., Leal, T., Dudley, A. & Stack, M. (2020). *Generative second language acquisition*. Cambridge University Press. <https://doi.org/10.1017/9781108762380>
- Wagers, M., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, 45, 395–433. <https://doi.org/10.1017/S0022226709005726>
- Xu, Yi. (2014). Evidence of the accessibility hierarchy in relative clauses in Chinese as a second language. *Language and Linguistics*, 15, 435–464.

Age of acquisition in second language thinking¹

Emanuel Bylund^{1,2} and Panos Athanasopoulos^{2,3}

¹Stockholm University / ²Stellenbosch University / ³Lancaster University

Debates around age effects in L2 learning have mostly centred around the domains of phonology, morphosyntax, and lexis. In this chapter, we expand the scope to conceptual and cognitive restructuring, much in line with Long's (1990, 2013) discussion of the limits of age effects. Drawing on an extensive review of the extant literature, the current chapter examines the role that age of L2 acquisition plays for phenomena of thinking-for-speaking and linguistic relativity. The chapter closes with methodological suggestions for future research.

Introduction

In the language sciences, there are some topics that have instigated heated debates, spanning decades and generating rich bodies of empirical research and conceptual analysis. One such controversial topic, particularly within the field of second language acquisition (SLA), is the role that age of acquisition (AoA) plays in the learning of a new language. Here, differences in the endstate of second language (L2) proficiency between child and adult learners have led some scholars to posit that if L2 acquisition does not start before a certain age, the attainment of natively-like proficiency will be compromised (e.g., Abrahamsson & Hyltenstam, 2009; DeKeyser, 2000; Flege, 1999; Munro & Mann, 2005). The scope and ultimate cause of such AoA differences have been fiercely debated, with some scholars positing a biologically based sensitive period for L2 acquisition (e.g., Johnson & Newport, 1989; Lenneberg, 1967), while others consider such hypothesis an “unfalsifiable embarrassment” (Singleton & Leśniewska, 2021). Other approaches have proposed a gradual decline of language learning abilities across the life span, not necessarily governed by neurobiological changes (e.g., Hakuta et al., 2003).

Another controversial topic in linguistics, dating back centuries if not even millennia, is the principle of linguistic relativity, also known as the Sapir-Whorf

1. This research was carried out with grant RJ SAB20-0050 (EB).

hypothesis. This notion holds that the grammatical and lexical categories specific to a given language influence the way in which speakers of that language categorize reality: ultimately, it posits that speakers of different languages think differently. The idea of linguistic relativity has been described as the “bête noire” of linguistics (Lakoff, 1987, p. 304), and the views surrounding it are as strong as they are varied. For instance, while Dennett (2017, p. 304) holds that “language infects and inflects our thoughts at every level”, Pullum (1988, p. 38) views linguistic relativity as “utterly boring”, and Pinker (1994, p. 52), in turn, considers it a “conventional absurdity”. Inherent in the theory complex of linguistic relativity is also the idea that learning a new language might change an individual’s thought patterns. However, the degree to which this may happen is heavily disputed, with scholars endeavouring to understand which factors of the bilingual experience may contribute to cognitive restructuring – if any is observed in the first place. While most studies focus on variables such as L2 proficiency, frequency of use, or length of immersion, other factors, such as AoA, have been afforded much less rigorous treatment and scrutiny in empirical investigations.

The aim of the current chapter is to bring together the topics of linguistic relativity and age of L2 acquisition. While both of these topics have received substantial attention separately, they are seldom discussed together. The specific question we will be addressing may be summarised as follows: Does the age at which you acquire a second language modulate the extent to which you acquire its thought patterns? In his several publications on child-adult differences in L2 ultimate attainment (e.g., Long, 1990, 1993, 2005, 2007), Mike Long often discussed whether AoA affects language proficiency across the board, or whether different subsystems of language exhibit differential susceptibility to age. Unlike much traditional research on AoA effects in SLA, Long went beyond the classic focus of phonology and syntax, positing – and later on, empirically demonstrating – that other less researched subsystems, such as lexis, may be affected by age of acquisition (e.g., Granena & Long, 2013). In his seminal 1990 publication on maturational constraints, Long also drew attention to discourse and pragmatics as a potential domain for AoA effects, citing preliminary findings on conversational strategies including topic choice and sequencing (Scarcella, 1983). Coincidentally, in the 1990s, research on the relationship between language and thought was just taking off (Gumperz & Levinson, 1996; Hunt & Agnoli, 1991; Lucy, 1992a, 1992b), spawning the beginning of a Whorfian renaissance that is still ongoing. However, systematic research on a larger scale on language and thought in L2 learners entered the SLA agenda later, and it is only more recently that research has generated empirical evidence on questions concerning the role of age in the acquisition of thought patterns in a L2. Inspired by Mike Long’s line of inquiry into the scope of age effects, this chapter seeks to extend the discussion on AoA to the domain of conceptual and cognitive processing. In doing

so, it is our humble intention to honour Long's legacy of critical and progressive discussions around age effects in SLA.

In an attempt to keep our treatment of age in L2 thinking as broad as possible, we have chosen a definition of the terms “thought patterns” and “thinking” that includes behaviour of both verbal (e.g., information structuring) and non-verbal (e.g., categorization, recognition memory) nature. Seeing that the empirical evidence on AoA in L2 thinking is still limited, this wider scope will allow us to approach the matter from a more inclusive viewpoint. This terminological broadening notwithstanding, it is important to recognize that modern research on language and thought places central importance on distinguishing between verbal and non-verbal behaviour (e.g., Casasanto, 2008; Levinson, 1997; Lucy, 1992b). Therefore, while we include both verbal and non-verbal evidence in our discussion, the two will not be conflated. In keeping with preferred nomenclature, when necessary we will use the term “conceptual” to refer to thought processes that occur in relation to speaking (e.g., Levelt, 1989; Slobin, 1996) and the terms “cognitive” and “perceptual” to refer to processing that occurs in the absence of overt speech (e.g., Bylund & Athanasopoulos, 2014). Moreover, we will use the theory-neutral terms “AoA effects” and “age effects” to refer to correlations between the age at which L2 learning starts and conceptual and cognitive behaviour. A different question, to which we will return after reviewing the available evidence, concerns the mechanisms that underlie such AoA effects. Finally, because the focus of the chapter is on *L2* thinking, we will delimit our review and discussion to evidence that concerns thought processes in an L2 context/L2 language mode. This means that, unless relevant to a central argument or interpretation, we will not review evidence of restructuring of L1 thinking patterns in bilinguals (instead, the reader is referred to Bylund, 2009; Bylund & Athanasopoulos, 2014).

The chapter is organized in the following way. We first provide a brief overview of research on language and thought, highlighting the conceptual and methodological frameworks currently guiding this line of investigation. Against this background, we review findings relevant to the role of AoA in L2 thinking. We then turn to delineate an interpretation of the effects (or lack thereof) of AoA on L2 thought patterns, and close the chapter by formulating recommendations for future research on the topic.

Language and thought

Modern research on language and thought has tended to focus on two different manifestations of the effects of linguistic categories on mental processes. These are commonly subsumed under the terms linguistic relativity and thinking-for-speaking.

Linguistic relativity

The term was coined by Benjamin Lee Whorf, who referred to it as the ‘linguistic relativity principle’, and essentially attempted to formalise the idea that language affects thinking, in predictable ways (an idea that Whorf himself developed from his Linguistics teacher Edward Sapir). Since different languages have different concepts, according to Whorf, speakers of different languages think, or interpret, the world differently. Whorf provided several statements that could be taken as definitions of the principle, varying in the degree to which differences in thinking between populations can be attributed to the way their languages encoded reality, from absolute determinism, where speakers are “at the mercy” of their native language (Sapir, quoted in Whorf, 1956, p. 134), to more temperate formulations where language acts as an attention directing mechanism, or a spotlight, to perceptual phenomena that may otherwise be processed similarly at a deeper level (“...language, for all its kingly role, is in some sense a superficial embroidery upon deeper processes of consciousness...”, Whorf, 1956, p. 239). The definition that, according to these authors, has been the cornerstone of modern, workable empirical investigations holds that “users of markedly different grammars are pointed by their grammars toward different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world.” (Whorf, 1956, p. 221).

This definition, and particularly the formulation ‘different evaluations of externally similar acts of observation’, dovetails nicely, if coincidentally, with later descriptions of basic human cognitive processes like categorisation. Indeed, the cognitive process of categorisation has become the locus classicus of empirical investigations of linguistic relativity, because on the one hand it is an essential element of human cognition (Harnad, 1987), and on the other hand, similarity is the basis of categorisation (Nosofsky, 1986). Therefore, judging the similarity between stimuli offers a readily testable hypothesis of the Whorfian question: stimuli encoded differently in different languages will tend to be judged as less similar by speakers of those languages, revealing differences in a fundamental aspect of human cognition, namely categorisation.

Using variations of this paradigm, research for example shows that speakers of languages with a ‘grue’ term to refer to green and blue will judge blue and green stimuli as more similar to each other than speakers of languages that encode the terms (Kay & Kempton, 1984; Roberson et al., 2005). In the domain of motion, speakers of languages that encode manner on the main verb (e.g. English: *The boy ran into the house*) will learn to categorise novel alien creatures faster and more accurately based on the perceived similarity of their manner of motion. Speakers of

languages that encode path on the main verb and manner in a satellite (e.g. Spanish: *The boy entered the house running*) learn to categorise novel alien creatures faster and more accurately based on the perceived similarity of their path trajectories (Kersten et al., 2010).

Beyond categorisation, more recent studies employ paradigms that tap into more implicit processes, such as psychophysical and visual perception. So Swedish speakers will tend to experience the duration it takes for a line to grow as shorter or longer depending on the physical growth of the line (longer lines tend to be perceived as taking longer to grow than shorter lines, even if the stimulus duration between the two different lines is identical), because duration is usually expressed via spatiotemporal metaphors denoting distance (e.g. long/short time). Spanish speakers on the other hand will tend to experience the duration it takes for a container to be filled with liquid as longer or shorter depending on the amount the container fills up to. Fuller containers are estimated to have taken longer to fill than less full containers, even when the duration of the filling was identical. This conforms to the prevalent spatiotemporal metaphors for duration in that language, which employ quantity terms (much/little time) (Bylund & Athanasopoulos, 2017). Regarding visual perception, based on a lexical distinction that carves the blue area of colour space into two using distinct terms to refer to light and dark blue, Greek speakers will perceive differences in light and dark blue as greater very early in the visual processing stream, and outside of conscious awareness, than English speakers, for whom the same lexical distinction is not present (Maier & Rahman, 2018; Thierry et al., 2009).

Thinking-for-speaking

The hypothesis of thinking-for-speaking (Slobin, 1996, 2003) differs from the linguistic relativity principle in that it zooms in on the thought processes that occur in relation to the production and comprehension of speech. Instead of positing a general influence of linguistic categories on cognitive and perceptual processing, Slobin suggests that when in the process of preparing content for speech, the speaker typically picks those features of the situation to-be-conveyed that are readily encodable in their language. From this follows, Slobin argues, that thinking takes on a language-specific form when occurring in relation to speech production. For instance, a speaker of a language that has the grammatical category of evidentiality (e.g., Turkish) has to encode the information source (e.g., witnessed first-hand, second-hand source, or hearsay) for most of their statements. Similarly, a speaker of a language in which animacy is a central grammatical category (e.g., Basque) has to morphologically mark the living status of referents. Most research on TFS has

focused on the linguistic packaging of motion. The starting point here is Talmy's typology of manner and path, as per the crosslinguistic differences between English and Spanish outlined in Example (1).

- (1) a. Non-boundary crossing motion (no change in location of motion):
The dog ran towards the garage
El perro corrió hacia el garaje
 'The dog ran towards the garage'
- b. Boundary-crossing motion (change in location of motion):
The dog ran into the garage
El perro entró en el garaje (corriendo)
 'The dog entered the garage (running)'

In so-called "verb-framed languages", such as Spanish, the manner of motion (running) cannot be encoded in the main verb if the event is of a boundary-crossing nature, as in (1b). Instead, the main verb is a path verb (enter) that conveys the change of location (into), and manner of motion is optionally encoded as an adjunct. Using different types of elicitation techniques, Slobin and colleagues have documented how speakers of different languages exhibit different patterns of selecting and organising information about motion, time, and space (e.g., Berman & Slobin, 1994; Strömquist & Verhoeven, 2004; Verhoeven & Strömquist, 2001).

To account for a range of TFS phenomena, Slobin (2003) suggests that researchers study thinking-before-speaking (e.g., anticipatory effects of attention allocation), thinking-during-speaking (i.e., the original TFS), and thinking-after-speaking (e.g., effects of verbalisation on memory). For instance, consistent with this suggestion, von Stutterheim and colleagues (2012) examined the effects of progressive aspect on event construal. They found that speakers of non-aspect languages were more likely to look at the event endpoints prior to speech, to mention endpoints in their event descriptions (e.g., "a woman walking to a house"), and to remember the endpoints in a subsequent memory test. In addition to testing different temporal phases of TFS, useful evidence can also be generated by studying gesture, under the assumption that gestural behaviour reveals conceptual representation. For instance, saying "that's a long time ago" whilst pointing backwards/shoulder wards, would indicate that the speaker construes the past as spatially being behind them (e.g., Núñez & Sweetser, 2006). Such information on the spatial mapping of time could not be obtained from studying the utterance alone. In short, relying on different data types to understand the thought processes that occur in relation to speech is important for the TFS enterprise, as it will allow researchers to draw firmer conclusions about the thought processes at stake in relation to the production and comprehension of speech (Athanasopoulos & Bylund, 2013).

Effects of age of acquisition in L2 thinking

Historically, L2 learning – and multilingualism in general – has enjoyed mixed degrees of visibility on the language and thought research agenda. Whereas early scholars would talk about multilingualism (Whorf, 1956), and even about how learning a new language could be equated with acquiring a new worldview (Humboldt, 1836), the late 1990s and early 2000s saw very little mention – sometimes even an outright disregard – of questions relating to language and thought in multilingual individuals (with a few exceptions, e.g., Ameel et al., 2005; Athanasopoulos, 2006; Boroditsky, 2001; Pavlenko, 1999; von Stutterheim, 2003). However, since the 2010s, there has been a steady increase of studies looking at phenomena of thinking-for-speaking and linguistic relativity from the perspective of L2 acquisition (e.g., see contributions in Athanasopoulos et al., 2016; Cook & Bassetti, 2011; Pavlenko, 2011). In what follows, we will review the available evidence relevant to AoA specifically.

Hohenstein et al.'s (2006) study on motion event construal was one of the first attempts to look systematically at age effects in thinking for speaking. They examined L1 Spanish – L2 English bilinguals residing in the US, with different ages of L2 English acquisition. Using videoclips of motion events with varied manner and path components, Hohenstein and associates elicited motion descriptions from two groups of bilinguals: an early group, who had acquired English before or at age 5, and a late group, who had started their English learning after age 12. The findings showed pronounced differences between the groups, such that the early group aligned with L1 English monolinguals in their use of manner and path verbs, whereas the later group used significantly fewer manner verbs and modifiers, and more path verbs in their descriptions. Potential effects of length of L2 exposure were not addressed.

A somewhat similar finding was reported by Pavlenko and Volynsky (2015), who examined motion event construal in adult Russian-English bilinguals. The age of arrival in the US of these participants ranged from 0 to 29 years of age, and length of residence was between 2 and 28 years. Russian and English are both satellite languages according to Talmy's typology and thus more closely related in the way they package information about manner and path than Spanish and English (however, Pavlenko and Volynsky make a compelling case for considerable intra-typological variability). As shown by the authors, going from L1 Russian to L2 English affords the learner with ample opportunities for positive transfer, and most learners did indeed converge with the L1 English speaker group. Some AoA effects were nonetheless found for motion verb diversity, but these are tentative due to what the authors claim to be a strong relationship between AoA and length of residence. In the absence of statistical analyses that probe the strength of that

relationship and/or partial out the confounding effects of length of residence, the reported AoA effects are difficult to interpret.

An alternative approach was taken in Bylund (2011a), where event construal was studied with regards to granularity and temporal linking. Using a silent film, Bylund elicited retellings from L1 Spanish – L2 Swedish adult bilinguals, whose AoA ranged from 1 to 19 years. Unlike previous studies, this investigation implemented strict participant selection criteria. Only individuals who had been perceived of as L1 speakers of Swedish by a panel of native listener judges were selected for further testing. This technique, a feature of Long's (Long, 1993, 2005) so-called nativelikeness paradigm, allowed for an inclusion of L2 speakers with highly advanced L2 proficiency (for further details on participant screening and selection, see Abrahamsson & Hyltenstam, 2009). The results showed that the L2 speakers, regardless of AoA, used the same temporal linking strategies as the L1 comparison group, inserting anaphoric adverbials (e.g., "then") to link the events together in narrative. In contrast, in terms of event granularity the L2 speakers were likely to segment the events in a more fine-grained way than the Swedish L1 speakers. Crucially, this pattern applied to the L2 group as a whole, and did not vary as a function of AoA. In the absence of AoA effects, one possibility is that the idiosyncratic event segmentation strategies were an artefact of bilingualism. Indeed, Bylund's (2011b) study of early bilinguals tested in both L1 Spanish and L2 Swedish revealed that while this group kept the language-specific patterns of temporal linking apart (thus aligning themselves with native speakers of each language), they exhibited the same degree of event granularity in both languages, regardless of AoA.

Focusing on thinking-after-speaking effects in the domain of motion, Lai, Garrido-Rodríguez and Narasimhan (2014) investigated whether verbal encoding of motion events prior to categorization would influence cognitive preferences among Spanish-English bilinguals, who had acquired Spanish from birth and English between 0 and 15 years of age. Using the same cut-off as Hohenstein et al. (2006), Lai et al. divided their participants into early (AoA < 6) and late (AoA ≥ 6) bilinguals. A triads-matching task was used, in which the target scene could be matched either with a manner alternate or a path alternate (for details, see Montero-Melis et al., 2017). The results showed that the early acquirers aligned with the L1 English monolingual group, exhibiting the same degree of manner preference in their categorisation. The late acquirers, too, exhibited a general preference for manner alternates, but to a significantly lesser extent than their early peers. While potentially confounding effects of length of residence was not controlled for, the overall findings suggest that early AoA was more likely to lead to nativelike categorisation preferences than late AoA.

In a test of linguistic relativity, Kersten et al. (2010) asked Spanish-English bilinguals to choose whether to categorize (in the absence of prior linguistic priming)

novel stimuli on the basis of the manner in which they moved, or the path with they traversed. In this category learning task, participants were assigned into either the manner (where classifying on the basis of similar manner was the correct option and classifying on the basis of same path was the wrong option) or the path condition (where the reverse of the manner condition was the case). Participants had to discover which the correct basis of categorisation was, by means of a supervised classification task that directed participants towards a path based or a manner based classification, based on feedback. Results from monolinguals had shown that English speakers were more successful in the manner condition than Spanish speakers, with no significant differences between the groups in the path condition.

Kersten et al. (2010) implemented their supervised classification task in bilinguals in two different language contexts, Spanish (L1 context) and English (L2 context), in a full between subjects design. The researchers examined AoA effects on category discrimination, by performing a median split on their bilingual sample. The median AoA turned out to be 5 years. Consequently, participants with an AoA on or before 5 years of age were considered to be early learners of English. Conversely, participants with an AoA after 5 years of age were considered to be late learners of English. Kersten et al. (2010) found that bilinguals in the manner condition (the pattern preferentially encoded in the L2) tested in a Spanish (L1) language context presented with a significant main effect of AoA, such that the late bilingual group (as per the above criteria) were less successful learning the correct classification pattern than the early bilingual group. AoA had no effect on category learning in the path condition and, interestingly, no effect on performance in an English language context. In a correlational analysis, where AoA was treated as a continuous rather than a categorical variable, the performance of bilinguals tested in the manner condition in a Spanish language context was negatively correlated with their AoA. These results reveal a better ultimate learning outcome in bilinguals who were exposed to English at an earlier age, although the fact that such an effect was not found in the English language context is in need of further explanation.

Malt and Sloman (2003) studied the linguistic categorization of common household objects in a group of English L2 speakers with varying L1 backgrounds and acquisition trajectories. Some effects of AoA were documented, as earlier learners were more likely to converge with L1 English speakers in their object construal patterns. However, AoA turned out to be highly correlated with length of residence in the L2 environment, and the effects of AoA disappeared in a multiple regression analysis with length of residence as co-predictor. Importantly, however, two different types of AoA measures were used: age of (first) exposure, and age of arrival (in the US). While relatively similar results were obtained for both measures, the fact that there were different operationalizations of AoA suggests that age of first exposure did not always coincide with age of immersion (i.e., a person could have been

exposed to English in school and only later in life moved to the US). As a result, it is possible that neither age factor represented the type of predictor variable that is ideal for testing AoA effects, that is, when age of first exposure coincides with age of immersion. For this reason, it is difficult to know whether the lack of robust AoA effects were due only to an overshadowing effect of length of residence, or whether the nature of the age factors tested also contributed to this result.

In the domain of time, Boroditsky (2001) examined Mandarin-English bilinguals with an AoA ranging from 3 to 13 years (M 9.4 years, SD 3.3 years). The researcher also looked at potential effects of length of exposure. All participants had at least 10 years of exposure to English (M 14.0 years, SD 2.3 years). Results from a temporal cognition task prompting participants to make decisions on events presented either in a vertical temporal sequence (present only in Mandarin) or in a horizontal one (present in both English and Mandarin) showed that those Mandarin speakers who started learning English later in life were biased to think about time vertically, following the pattern exclusively encoded in their L1. In other words, the later a bilingual individual started speaking English, the greater their vertical bias was in the temporal judgment task. Interestingly, length of exposure to the L2 English had no effect on the observed vertical bias. In a series of partial correlations, the strong correlation between AoA and vertical bias remained even after controlling for the effect of length of exposure, while there was no significant correlation between vertical bias and length of L2 exposure with and without taking into account AoA. Boroditsky (2001) likened these findings to those reported in the classic study of Johnson and Newport (1989), where acquisition of L2 grammatical proficiency correlated with AoA but not length of L2 exposure, but the researcher stopped short of proposing the existence of a critical or sensitive period for cognitive restructuring in bilingualism. Given the limited age range of the participants in Boroditsky's (2001), and the fact that the majority of the participants were pre-pubescent learners, it would have been indeed premature to posit any chronological marker as a cut-off point.

In another experiment, Boroditsky (2001) briefly trained native English speakers to talk about time using vertical terms. The trained participants showed very similar patterns of temporal cognition to the Mandarin speakers, that is, a vertical bias, which was absent in English speaking participants who hadn't received such training. This finding on the one hand strengthens the interpretation of previous findings that language per se, rather than some undefined cultural variable, is driving the cross-linguistic cognitive differences between populations, but on the other it shows that internalising new ways to talk about time in the L2 can in fact alter the way one interprets temporal events, at least in the short term, even in adulthood. The question then is, how pervasive are these changes in temporal cognition in the

longer term? And does AoA modulate the longevity of these cognitive changes, rather the degree to which they can occur?

Athanasopoulos and Bylund (under review) examined time estimation in Spanish-Swedish bilinguals. Previous research had shown that when estimating the duration of events, native speakers of languages where time is usually talked about as if it were long or short (e.g., English and Swedish), tend to perceive physically stretched out lines as also being stretched out in time. Conversely, speakers of languages where time is mainly represented as big or small (e.g., Greek and Spanish), are instead misled by the amount with which virtual container is filled (Bylund & Athanasopoulos, 2017; Casasanto et al., 2004). Furthermore, in a laboratory training paradigm, Casasanto (2008) trained English native speakers to use quantity-based terminology in a gap filling task (e.g. a sneeze is bigger/smaller than a vacation) and found that participants did indeed learn to think about duration like Greek or Spanish speakers do, misled by the quantity or volume of filling container when estimating the time it took for the container to fill.

Short-term training paradigms do show that it is possible for adults to internalise a new cognitive pattern, however they still leave the long-term effects of learning unaddressed. Athanasopoulos and Bylund (under review) aimed to fill that gap, by looking at how Spanish-Swedish bilinguals estimate duration of growing lines (congruent to distance-based terminology in the L2) and filling containers (congruent to quantity-based terminology in the L1), in a fully between subjects design. AoA ranges varied substantially in both participant groups, specifically 1–28 years ($M = 9.6$ years) in the growing lines experiment, and 1–31 years ($M = 9.1$ years) in the filling containers experiment.

The results of the study showed no significant correlation between AoA and time estimation in the lines experiment (the L2 pattern). Instead, frequency of daily usage of the L2 reliably predicted the degree of shift towards the L2 pattern, that is, the more an individual used their L2 on a daily basis, the more likely they were to be misled by how long the line grew when estimating the duration it took to grow. Interestingly, when performance was examined in the containers experiment (the L1 pattern), AoA was found to significantly predict the degree to which bilinguals followed the L1 pattern of time estimation. Specifically, the earlier bilinguals started learning their L2, the less likely they were to be misled by how much a container had been filled when estimating the duration it took to be filled.

As becomes clear from the studies reviewed above, the evidence of AoA effects in L2 thinking is rather mixed. In addition, some studies do not provide explicit information about their participants' learning situation (i.e., whether immersed in the L2 settings or mainly through formal classroom teaching), and some do not address the potential confound of length of residence/exposure. While this may be

understandable because some studies were not designed to test AoA specifically, it nevertheless poses a limitation on the robustness of the evidence available to date: Keeping the learning situation constant and controlling for potential confounding variables such as length of exposure are fundamentals in SLA research methodology on AoA effects, and without adherence to them there is a serious risk that the evidence is compromised in terms of their interpretational value to the age question. Table 1 lists information about the reviewed studies and the robustness of the produced evidence. Here, evidentiary robustness ranges from High to Low. If a study keeps learning situation constant, and moreover controls for length of exposure, we have assigned the evidentiary degree High. If a study fails to take account of one of these, the degree Medium is conferred. Studies that do not take account of neither learning situation nor length of exposure are conferred the degree Low. As becomes clear in Table 1, most of the evidence of AoA effects on L2 thinking is on the Medium to Low side.

Table 1. Empirical evidence of AoA effects on L2 thought

Study	Domain	Evidence type	L2 learning situation	Length of exposure controlled	AoA effects	Robustness of evidence
Athanasopoulos & Bylund, <i>subm</i>	Time	Non-verbal	Immersed	Yes	No	High
Boroditsky, 2001	Time	Non-verbal	Immersed?	Yes	Yes	Medium
Bylund, 2011a	Events	Verbal	Immersed	Yes	No	High
Hohenstein et al., 2006	Motion	Verbal	Immersed?	No	Yes	Low
Kersten et al., 2010	Motion	Non-verbal	Immersed?	No	Yes	Low
Lai et al., 2014	Motion	Non-verbal	Immersed?	No	Yes	Low
Malt & Sloman, 2003	Objects	Verbal	Mixed?	Yes	No	Medium
Pavlenko & Volynsky, 2015	Motion	Verbal	Immersed	No	Yes	Medium

Interpreting age effects in L2 thinking

At a first glance, the review would seem to suggest that there may be an age effect with regards to L2 thinking about manner and path of motion, seeing that most studies that investigated this domain did report an AoA effect of sorts. However, a problematic dimension of the evidence, as stated above, is that length of exposure has not been controlled for in these studies. This potential confound allows us to bring in the notion of entrenchment as a mechanism that might be applicable to

effects of both AoA and length of exposure. Entrenchment refers to an emergent result of network specialisation, the ultimate effect of which is that early acquired representations interfere with the learning of novel representations later in life. (Zevin, 2012). Specifically, drawing on principles of Hebbian learning, it is assumed that once a network has been specialised to categorise stimuli in a certain way, it achieves an attractor state (MacWhinney, 2016). Such states lead to efficient processing, but they also stabilise the networks to the extent that these become difficult to change and insensitive to new information (Brooks & Tomasello, 1999). The notion of entrenchment has been applied to analyse the development of phonetic, morphological, lexical, and semantic representation in L2 acquisition (e.g., Flege, 1999; Hernandez et al., 2005; Li et al., 2004; Monner et al., 2013; Munnich & Landau, 2010).

An inherent assumption of entrenchment accounts is that it is the learning itself (i.e., the formation and subsequent strengthening of language-specific categories) that reduces the plasticity of the system and, with this, the ability to form novel categories with equal representational strength. In this sense, entrenchment accounts differ from maturational accounts (e.g., Johnson & Newport, 1989; Lenneberg, 1967; Mayberry et al., 2011), which instead typically postulate a reduction of plasticity due to biologically scheduled maturation. Entrenchment accounts also differ in that they ascribe less importance to timing; timing matters in the sense of order, such that representations that are acquired first have primacy in representational space compared to those that are acquired later (MacWhinney, 2004). It is, however, also assumed that representational space can be modulated as a function of extensive changes in experience, such as prolonged exposure.

In the case of motion events, the categories of manner and path will exhibit different degrees of entrenchment depending on language-specific weighting (e.g., manner will occupy a smaller representational space in speakers of verb-framed languages than in speakers of satellite languages). These biases will then, as shown by the studies reviewed above, have a predictable influence on the way in which L2 learners conceptualise and categorise motion depending on the typology of their L1 (and other previously acquired languages). Because of its emphasis on amount of exposure (as opposed to timing), an entrenchment approach to the current evidence on L2 thinking about manner and path of motion can accommodate the fact that the observed effects could be due to either AoA, length of exposure, or both.

An explanation based on entrenchment could possibly be offered for Boroditsky's (2001) finding on the categorisation of temporal sequences in English L2 speakers with L1 Mandarin Chinese: the later the learners acquired English, the more entrenched the vertical time schemas would be. However, what is noteworthy about this finding is that no effects of length of exposure were documented. This

would seem to suggest that the vertical bias among the later learners was not simply an effect of amount of exposure at any point in time, but that timing of exposure was the determining factor.

An important question that emerges from the review is why no effects of AoA obtained with the construal of events and objects, or with duration perception. These perceptual domains and their relevant linguistic constructions vary considerably: object terms and duration metaphors have, similar to the packaging of manner and path, a grammaticality element to them. For instance, slotting manner information into the main verb in Spanish is simply not grammatical for boundary-crossing motion, nor is calling a cup “a glass” in English). In contrast, the organisation of event information reflects patterns of preference that are neither grammatically nor lexically mandated. There is, in other words, no obvious common denominator for these domains and behaviours.

Towards a better understanding of AoA in L2 thinking

In view of the discussions above, we propose that future investigations into AoA effects in L2 thinking will benefit from observing the following suggestions:²

- Differentiating between immersed learning and foreign language learning
As research shows, AoA effects virtually disappear in formal learning situations (e.g., Muñoz, 2006). Therefore, avoiding to mix L2 speakers with different learning backgrounds in a study design will enhance the robustness of any evidence for or against AoA effects in L2 thinking.
- Using a wide range of AoAs
Including AoAs that range from very early childhood (e.g., 2 years of age) well into adulthood (e.g., 45 years of age) will allow for greater precision in determining the temporal and geometric properties of any documented AoA effects (e.g., off-set ages and discontinuities). Needless to say, potential confounds of length of exposure need to be controlled for here.
- Using simultaneous bilinguals for comparison purposes
In order to ascertain that any documented differences between L2 speakers and L1 speakers are not an artefact of functional bilingualism, some studies have introduced simultaneous bilinguals as a comparison group, instead of or

2. Admittedly, the list of suggestions could be made longer. It would, for instance, be possible to, in line with Long's (1993, 2005) nativelikeness paradigm, apply a strict selection of potentially nativelike L2 speakers. However, this methodological choice will ultimately boil down to the research question.

in addition to L1 functional monolinguals (e.g., Bylund et al., 2019; Verissimo et al., 2018). Because conceptual representation in bilingualism is often different from its monolingual counterpart, it is important to control for this variable. Depending on their research question, studies can implement different designs here. For instance, if the aim is to assess both bilingualism effects and discontinuities in an AoA function, the simultaneous bilingual group would need to be compared to a L2 group with a wide range of AoAs.

– Moving beyond speech alone when investigating AoA effects in thinking-for-speaking

The current evidence on AoA in thinking-for-speaking is largely based on verbal data. To better understand whether the conceptual processes that occur in relation to speech are susceptible to AoA, it is necessary to use methodologies that tap into, for instance, visual attention and gesture (in addition to memory, as done by Lai et al., 2014).

– Promoting an evidence-based interpretation of evidence

As mentioned in the introduction to this chapter, AoA effects and linguistic relativity are controversial topics in linguistics. It has been suggested that the controversy arises, in part, because both topics have implications for the notions of free will and the blank slate (along the lines of “the content of our thoughts should not be determined by language”, and “our ability to learn a new language should not be determined by biology”, see Paradis, 2009; Troyer, 1994). With this, there may have been a tendency to over-interpret evidence (or the absence thereof) to advance an agenda (see discussion in Levinson, 2003). Clearly, sweeping statements that language determines thought, or that differences in thinking between L1 and L2 speakers boil down to a critical period, does little for advancing a nuanced discussion – and enhanced understanding – of the phenomena of linguistic relativity and AoA effects. However, owing to methodological and conceptual developments, these debates now tend to be more constructive and research into AoA effects on L2 thinking will benefit from continuing along this path.

Conclusion

The aim of this chapter has been to bring together the topics of linguistic relativity and AoA in L2 learning in order to expand the scope of current research on AoA effects. An underlying rationale for this enterprise has been that, as argued by Mike Long, by investigating the scope of AoA effects, we will be in a better position to understand the conditions under which L2 learning takes place. A review of the evidence generated to date suggests that AoA does not exert a robust impact on

conceptual and cognitive patterns in the L2: The attainment of nativelike conceptualisation of events and objects, although arguably a late feature in L2 learning (e.g., Schmiedtová et al., 2011; von Stutterheim, 2003), is not subject to AoA effects; for construing and categorising manner and path of motion, the evidence for AoA effects is preliminary; in the domain of temporal cognition, the evidence is mixed, showing seemingly solid AoA effects for categorisation of temporal sequences but not for duration perception.

In addition to considering the suggestions formulated in the preceding section, an important task for future research will be to develop a taxonomy of L2 thinking processes in order to identify possible interactions between perceptual domain (e.g., motion), linguistic structure (e.g., grammatical aspect), behaviour/mental process (e.g., visual attention) and AoA. A fuller picture of the types of behaviours potentially affected by AoA and their linguistic categories will lay the ground for a better understanding of the ways in which putative AoA effects are documented and understood.

References

- Abrahamsson, N., & Hylenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249–306. <https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53(1), 60–80. <https://doi.org/10.1016/j.jml.2005.02.004>
- Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition*, 9(1), 89–96. <https://doi.org/10.1017/S1366728905002397>
- Athanasopoulos, P., & Bylund, E. (2013). The ‘thinking’ in thinking-for-speaking: Where is it? *Language, Interaction, and Acquisition*, 4(1), 91–100. <https://doi.org/10.1075/lia.4.1.05ath>
- Athanasopoulos, P., Bylund, E., & Casasanto, D. (2016). Introduction to the Special Issue: *New and interdisciplinary approaches to linguistic relativity*. *Language Learning*, 66(3), 482–486. <https://doi.org/10.1111/lang.12196>
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum Associates.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers’ conceptions of time. *Cognitive Psychology*, 43(1), 1–22. <https://doi.org/10.1006/cogp.2001.0748>
- Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, 75(4), 720–738. <https://doi.org/10.2307/417731>
- Bylund, E. (2009). Effects of age of L2 acquisition on L1 event conceptualization patterns. *Bilingualism: Language and Cognition*, 12(3), 305–322. <https://doi.org/10.1017/S1366728909990137>
- Bylund, E. (2011a). Ultimate attainment of event segmentation and temporal structuring patterns in speakers of L2 Swedish. *Vigo International Journal of Applied Linguistics*, 8, 28–53.

- Bylund, E. (2011b). Segmentation and temporal structuring of events in early Spanish-Swedish bilinguals. *International Journal of Bilingualism*, 15(1), 56–84.
<https://doi.org/10.1177/1367006910379259>
- Bylund, E., Abrahamsson, N., Hyltenstam, K., & Norrman, G. (2019). Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition*, 182, 45–49.
<https://doi.org/10.1016/j.cognition.2018.08.020>
- Bylund, E., & Athanasopoulos, P. (2014). Linguistic relativity in SLA: Toward a new research program. *Language Learning*, 64(4), 952–985. <https://doi.org/10.1111/lang.12080>
- Bylund, E., & Athanasopoulos, P. (2017). The Whorfian time warp: Representing duration through the language hourglass. *Journal of Experimental Psychology: General*, 146, 911–916.
<https://doi.org/10.1037/xge0000314>
- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning*, 58, 63–79.
<https://doi.org/10.1111/j.1467-9922.2008.00462.x>
- Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., Fotokopolou, O., Pita, R., & Gil, D. (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek and Spanish. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference Cognitive Science Society* (pp. 575–580). Lawrence Erlbaum Associates.
- Cook, V., & Bassetti, B. (Eds.). (2011). *Language and bilingual cognition*. Psychology Press.
<https://doi.org/10.4324/9780203836859>
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533. null.
<https://doi.org/10.1017/S0272263100004022>
- Dennett, D. C. (2017). *Consciousness explained*. Little, Brown.
- Flege, J. E. (1999). Age of learning and second language speech. In *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). Lawrence Erlbaum Associates.
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343.
<https://doi.org/10.1177/0267658312461497>
- Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking linguistic relativity*. Cambridge University Press.
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the Critical-Period Hypothesis for second language acquisition. *Psychological Science*, 14(1), 31–38.
<https://doi.org/10.1111/1467-9280.01415>
- Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 1–52). Cambridge University Press. <http://cogprints.org/1571/>
- Hernandez, A. E., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, 9(5), 220–225. <https://doi.org/10.1016/j.tics.2005.03.003>
- Hohenstein, J., Eisenberg, A., & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish–English bilingual adults. *Bilingualism: Language and Cognition*, 9(3), 249–261. <https://doi.org/10.1017/S1366728906002616>
- Humboldt, W. F. von. (1836). *On language: The diversity of human language-structure and its influence on the mental development of mankind*. Cambridge University Press.

- Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, 98(3), 377–389. <https://doi.org/10.1037/0033-295X.98.3.377>
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Kay, P., & Kempton, W. (1984). What Is the Sapir-Whorf Hypothesis? *American Anthropologist*, 86(1), 65–79. <https://doi.org/10.1525/aa.1984.86.1.02a00050>
- Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, 139(4), 638–653. <https://doi.org/10.1037/a0020507>
- Lai, V. T., Rodriguez, G. G., & Narasimhan, B. (2014). Thinking-for-speaking in early and late bilinguals. *Bilingualism: Language and Cognition*, 17(1), 139–152. <https://doi.org/10.1017/S1366728913000151>
- Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press.
- Levinson, S. C. (1997). Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. *Journal of Linguistic Anthropology*, 7(1), 98–131. <https://doi.org/10.1525/jlin.1997.7.1.98>
- Levinson, S. C. (2003). Language and mind: Let's get the issues straight! In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 25–46). The MIT Press.
- Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, 17(8), 1345–1362. <https://doi.org/10.1016/j.neunet.2004.07.004>
- Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12(03), 251–285. <https://doi.org/10.1017/S0272263100009165>
- Long, M. H. (1993). Second language acquisition as a function of age. In K. Hyltenstam & A. Viberg (Eds.), *Progression and regression in language: Sociocultural, neuropsychological and linguistic perspectives* (pp. 196–221). Cambridge University Press.
- Long, M. H. (2005). *Problems with supposed counter-evidence to the Critical Period Hypothesis*. 43(4), 287–317. <https://doi.org/10.1515/iral.2005.43.4.287>
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. H. (2013). Maturational constraints on child and adult SLA. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 3–42). John Benjamins. <https://benjamins.com/catalog/lilt.35>
- Lucy, J. A. (1992a). *Grammatical categories and cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620713>
- Lucy, J. A. (1992b). *Language diversity and thought*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620843>
- MacWhinney, B. (2004). A unified model of language acquisition. *Department of Psychology*. Retrieved on 9 February 2022 from https://kilthub.cmu.edu/articles/journal_contribution/A_Unified_Model_of_Language_Acquisition/6613271
- MacWhinney, B. (2016). Entrenchment in second-language learning. In H. J. Schmid (Ed.), *Entrenchment and the psychology of language learning. How we reorganize and adapt linguistic knowledge*. De Gruyter Mouton. <https://doi.org/10.1515/9783110341423>
- Maier, M., & Abdel Rahman, R. (2018). Native Language Promotes Access to Visual Consciousness. *Psychological Science*, 29(11), 1757–1772. <https://doi.org/10.1177/0956797618782181>

- Malt, B. C., & Sloman, S. A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition*, 6(1), 47–67.
<https://doi.org/10.1017/S1366728903001020>
- Mayberry, R. I., Chen, J.-K., Witcher, P., & Klein, D. (2011). Age of acquisition effects on the functional organization of language in the adult brain. *Brain and Language*, 119(1), 16–29.
<https://doi.org/10.1016/j.bandl.2011.05.007>
- Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKeyser, R. (2013). A neural network model of the effects of entrenchment and memory development on grammatical gender learning. *Bilingualism: Language and Cognition*, 16(2), 246–265.
<https://doi.org/10.1017/S1366728912000454>
- Montero-Melis, G., Eisenbeiss, S., Narasimhan, B., Ibarretxe-Antuñano, I., Kita, S., Kopecka, A., Lüpke, F., Nikitina, T., Tragel, I., Jaeger, T. F., & Bohnemeyer, J. (2017). Satellite- vs. verb-framing underpredicts nonverbal motion categorization: Insights from a large language sample and simulations. *Cognitive Semantics*, 3(1), 36–61.
<https://doi.org/10.1163/23526416-00301002>
- Munnich, E., & Landau, B. (2010). Developmental decline in the acquisition of spatial language. *Language Learning and Development*, 6(1), 32–59.
<https://doi.org/10.1080/15475440903249979>
- Muñoz, C. (2006). *Age and the rate of foreign language learning*. Multilingual Matters.
<https://doi.org/10.21832/9781853598937>
- Munro, M., & Mann, V. (2005). Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics*, 26(3), 311–341. <https://doi.org/10.1017/S0142716405050198>
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology. General*, 115(1), 39–61.
<https://doi.org/10.1037/0096-3445.115.1.39>
- Núñez, R., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, 30(3), 401–450. https://doi.org/10.1207/s151516709cog0000_62
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.
<https://doi.org/10.1075/sib1.40>
- Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition*, 2(3), 209–230. <https://doi.org/10.1017/S1366728999000322>
- Pavlenko, A. (Ed.). (2011). *Thinking and speaking in two languages*. Multilingual Matters.
<https://doi.org/10.21832/9781847693389>
- Pavlenko, A., & Volynsky, M. (2015). Motion encoding in Russian and English: Moving beyond Talmy’s typology. *The Modern Language Journal*, 99(S1), 32–48.
<https://doi.org/10.1111/modl.12177>
- Pinker, S. (1994). *The language instinct: How the mind creates language*. Penguin Books.
<https://doi.org/10.1037/e412952005-009>
- Pullum, G. (1988). *The great Eskimo vocabulary hoax and other irreverent essays on language*. Equinox.
- Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, 50, 378–411.
<https://doi.org/10.1016/j.cogpsych.2004.10.001>
- Scarcella, R. (1983). Discourse accent in second language performance. In L. Selinker & S. Gass (Eds.), *Language transfer in language learning* (pp. 306–326). Newbury House.

- Schmiedtová, B., Stutterheim, C. von, & Carroll, M. (2011). Language-specific patterns in event construal of advanced second language speakers. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 66–107). Multilingual Matters.
<https://doi.org/10.21832/9781847693389-005>
- Singleton, D., & Leśniewska, J. (2021). The Critical Period Hypothesis for L2 acquisition: An unfalsifiable embarrassment? *Languages*, 6(3), 149. <https://doi.org/10.3390/languages6030149>
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). The MIT Press.
- Strömquist, S., & Verhoeven, L. (2004). *Relating events in narrative, Vol. 2: Typological and contextual perspectives*. Routledge. <https://doi.org/10.4324/9781410609694>
- Stutterheim, C. von. (2003). Linguistic structure and information organisation. The case of very advanced learners. *EuroSLA Yearbook*, 3, 183–206. <https://doi.org/10.1075/eurosla.3.11stu>
- Stutterheim, C. von, Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics*, 50(4), 833–867. <https://doi.org/10.1515/ling-2012-0026>
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11), 4567–4570.
<https://doi.org/10.1073/pnas.0811155106>
- Troyer, G. van. (1994). Linguistic determinism and mutability: The Sapir-Whorf “Hypothesis” and intercultural communication. Retrieved on 9 February 2022 from <https://www.semanticscholar.org/paper/Linguistic-Determinism-and-Mutability%3A-The-and-Troyer/18ebb13b3988e40ec23a946192bbb807e855f9fa>
- Verhoeven, L. T., & Strömquist, S. (2001). *Narrative development in a multilingual context*. John Benjamins. <https://doi.org/10.1075/sibil.23>
- Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition*, 25(3), 315–326. <https://doi.org/10.1080/10489223.2017.1346104>
- Whorf, B. L. (1956). *Language, thought, and reality: Selected writings* (edited by J. B. Carroll). The MIT Press.
- Zevin, J. D. (2012). A sensitive period for shibboleths: The long tail and changing goals of speech perception over the course of development. *Developmental Psychobiology*, 54(6), 632–642. <https://doi.org/10.1002/dev.20611>

Age effects in naturalistic and instructed second language acquisition

Two sides of one coin

Robert M. DeKeyser

University of Maryland

For more than half a century now the idea has been that children are better language learners than adults and that therefore second language learning in school should start early. Ironically, until fairly recently, the SLA literature on age effects, which was used to advocate foreign language in the elementary school, comprised hardly a single classroom study. The more recent literature does include a number of interesting classroom studies that show that children are not necessarily better language learners in the classroom, and actually worse learners for the most part; this literature has now been used to argue against the concept of the critical period. In this article I argue that these two bodies of literature are not contradictory and that age effects cannot be fully understood without reconciling them, showing that children are better at one particular kind of learning. This has profound implications for theories of cognitive development, theories of second language learning, and second language curriculum design.

1. Introduction

The topic of age effects in second language acquisition, and of the “critical period” in particular, has been a controversial one in the last quarter-century. There are various reasons for that, one of which is terminology. Therefore, I want to state clearly at the outset that by age effects we mean variation due to age at any time during the lifespan, while the term “critical period” refers more specifically to the time period (early adolescence) after which learning a (second) language becomes markedly more difficult. The term “critical” does not imply that after that age no language learning is possible anymore; therefore some prefer the term “sensitive” period. I will use the term “critical period” here, simply because it is the most widely used, without implying anything more than what I stated above: a time period after which learning becomes markedly more difficult. The term “critical” does not imply either

that that the reduction in learning capacity happens very abruptly; many studies have shown a gradual decline. Furthermore, terms like “critical” or “markedly decline” do not imply that certain learning processes are completely impossible after a certain age, just that they get markedly weaker, and that this is for reasons of developmental psychology, ultimately rooted in biology, and not to environmental or social-psychological factors (alone). Finally, I want to stress that the concept of a critical period refers to implicit learning, i.e., learning without awareness of (the specifics of) what is learned. A 40-year-old with high aptitude and motivation for second language learning may do better in an intensive form-focused course than a 10-year-old, just like the adult is likely to do better when learning algebra, philosophy, or chemistry. This was already made clear by Eric Lenneberg in his 1967 book *The biology of language*, yet it often seems to be forgotten (see esp. p. 176).

This last point in particular has led to much confusion. The failure to distinguish age effects on implicit versus explicit learning has led to two major mistakes, which I will discuss in more detail below: (a) the *assumption* that the age effects observed in immigrants, whose language learning largely takes place outside of a school context, imply that children should get second language *instruction* at a young age, whatever form that instruction takes; (b) the *assumption* that the failure to find a “younger is better” effect in most studies conducted in a school context can be used as evidence against the concept of a critical period. Findings in either context cannot simply be used to predict outcomes in the other context, because immigrants’ learning is mostly implicit and foreign language learning in the classroom is mostly explicit, even though not necessarily so (if young children are taught a second language through immersion without any focus on form, their learning will be largely implicit, even though it takes place in the classroom; if adult or adolescent immigrants take traditional form-focused courses their learning will be largely explicit).

For decades the vast majority of studies on age effects were conducted with immigrants of varying ages, yet their findings were often used to advocate early classroom teaching. More recently, however, a growing body of research on age effects has accumulated that was conducted in a classroom context, and it usually shows the opposite of “the younger the better.” There are exceptions, though, and those only reinforce the importance of distinguishing implicit from explicit learning in discussions about age effects, whether in the classroom context or among immigrants, as will be explained in Section 3 of this chapter; Section 2 will give a quick overview of the findings on age effects among immigrants, and Section 3 will do the same for the literature on age effects in the classroom. As both bodies of literature are dealt with, there is no space for an exhaustive review here. More detailed literature reviews of the research with immigrants can be found, e.g., in DeKeyser (2012); Long (2013a), and Mayberry & Kluender (2018); for detailed methodological

criticisms of some of the research, see also Long (2007) and DeKeyser (2013). For detailed reviews of the research with young classroom learners, see, e.g., Lambelet & Berthel  (2015); Huang (2016) and Baumert et al. (2020).

2. Research with immigrants

The majority of age effect studies with immigrants concern morphosyntax. These studies are based on grammaticality judgments and generally show large correlations between AoA (age of arrival, often used as a proxy for age of first exposure) and proficiency. More importantly, when different age ranges are looked at separately, various studies show a decline early on, till AoA 16 or 18 or so (e.g., DeKeyser, 2000; DeKeyser et al., 2010; Johnson & Newport, 1989; Hartshorne et al., 2018), followed by decades with little or no further decline till AoA 40 or so, but some find a discontinuity in the age-proficiency function that comes earlier, e.g. at age 12 (e.g. Abrahamsson & Hyltenstam, 2009) or later, e.g. at age 20 (Hakuta et al., 2003).

In the domain of phonology, studies have mostly relied on global ratings (e.g., Asher & Garcia, 1967; Flege et al., 1999; Huang, 2014), and just like the morphosyntax studies, they have shown high correlations between AoA and proficiency, for the lifespan as a whole, or for young learners. Studies on more specific aspects of pronunciation, such as prosody, have yielded similar results (e.g., Huang & Jun, 2011). More recent studies of the effect of AoA on L2 phonology have used a variety of methodologies, including acoustic measures (e.g., St lten, Abrahamsson, & Hyltenstam, 2013, 2015, for effects of AoA on perception and production of voice onset time in Swedish L2). Whatever the methodology, there is little or no evidence for any late learners passing for native on any specific measure (but see Bongaerts et al., 1997).

In the area of semantics, the picture is a bit more complicated, and one has to distinguish between lexical semantics (vocabulary) and phrasal semantics (how grammar conveys meaning). The lexicon is probably the domain that is least affected by AoA; some adult learners even develop a larger vocabulary than many native speakers do. There is, however, evidence that association between words and especially the activation of L1 or L2 words by words in the other language is affected by age of L2 learning (e.g., Silverberg & Samuel, 2004), and that collocations are learned less well by older learners (e.g., Abrahamsson & Hyltenstam, 2009; Granena & Long, 2013). Even for vocabulary in the strictest sense at least one study shows an AoA effect (Spadaro, 2013). For phrasal semantics, on the other hand, the evidence is overwhelming (*pace* Montrul & Slabakova, 2013; Slabakova, 2006) that older learners have a very hard time marking grammatically semantic distinctions that

do not exist in their native language. Articles in English L2, for instance, remain a problem even for very advanced post-puberty learners, and the same can be said for aspectual distinctions (see, e.g., DeKeyser, 2005) and classifiers in various L2s (e.g. Suzuki & DeKeyser, 2017). Even where a semantic distinction clearly exists in L1 but is not expressed grammatically, using the corresponding morphemes in the L2 remains difficult when conditions make use of explicit knowledge hard (see, e.g., Jiang, 2004, 2007, on processing of the English plural by native speakers of Chinese, or Trenkic, 2007 for processing of articles in English L2 by native speakers of Serbian).

Besides these types of behavioral studies, a substantial number of electrophysiological and neuro-imaging studies have accumulated in the last quarter-century, mostly about morphosyntax, but also about phonology (e.g., Pelzl et al., 2019 on perception of lexical tone in Mandarin L2). Some of the earlier studies in this area sometimes did not find effects for AoA, because they confounded AoA with proficiency or with monolingualism versus bilingualism. Among the studies that control for these variables, virtually all show a correlation between AoA and ultimate attainment, in this case in the form of location-specific activation seen with fMRI, or the stimulus-induced positivities and negativities found with ERPs (event-related potentials), with few exceptions, e.g., Perani et al. (1998). It should be noted, of course, that a Type-1 error (not finding differences where they do exist) is likely in this body of literature, because the studies tend to have very specific hypotheses about activation in certain areas at some points in time for certain stimuli, so that other AoA-related differences may go unnoticed.

Some researchers have used more indirect information, from census data (e.g. Chiswick & Miller, 2008; Hakuta et al., 2003). These studies do not show the patterns found in the bodies of literature cited above, and instead show a more gradual decline during most of the lifespan. This is not surprising, however, because the census data only contain extremely coarse assessments of proficiency (participants simply rate their own proficiency on a 4-point scale), and because there probably is a response bias in the sense that older people know they are expected to score lower. For various additional reasons, see Stevens (2004).

Finally, a recent study is in a category of its own that we could call crowd-sourcing: Hartshorne et al. (2018) collected data (with grammaticality judgments and a sentence-picture matching task) from 246,497 monolingual native speakers of English as well as 45,067 immersion learners ($\geq 90\%$ input in English) and 266,701 non-immersion learners ($\leq 10\%$ of input in English). They managed to do this over the internet by framing the data collection as some sort of a quiz with the extra twist that the computer was supposed to guess participants' native language and their dialect of English from their answers. While this internet data collection undoubtedly must have lowered the validity of the data somewhat, it is an amazing

feat. The numbers of participants are of the same order of magnitude as those in the census studies, but the data are much more fine-grained; a distinction is made, for instance, between L1s of different language groups. No individual structures are discussed, however, and no individual L1s. The findings are too rich to be presented in detail here, but the overall conclusion is that there is a breaking point in the AoA-proficiency function around age 17.4. Some have expressed surprise that this breaking point is so high in comparison with most studies, but one has to take into account that the breaking point here is the age at which learning ability declines, so that means that AoA was several years before, given that a few years are necessary for acquisition before the ability for implicit learning declines quickly; in other words a decline in learning ability at age 17 will be reflected in a decline in ultimate attainment for those with an AoA of about 12 to 14. Hyltenstam and Abrahamsson (2003) already drew attention to the fact that AoA precedes the point of strongest decline by several years, but the literature has largely overlooked this point and implicitly equated the AoA for which there is a strong decline with the age at which the decline itself takes place.

Evidence for the *nature* of age effects is both more indirect and more controversial. It comes primarily from research that shows that age effects vary depending on the L2 structure at issue (e.g., DeKeyser, 2000; DeKeyser et al., 2017; Johnson & Newport, 1989) and aptitudes (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; DeKeyser et al., 2010; Granena & Long, 2013; Morgan-Short et al., 2014; Pili-Moss, 2021; Pili-Moss et al., 2020). Comparing age effects for different structures, Johnson and Newport (1989) and DeKeyser (2000) found very similar results for which English L2 structures are most age-sensitive, which DeKeyser interpreted as possibly due to an interaction between AoA and the salience of structures (e.g., the inversion in yes-no questions is quite salient and not age-sensitive, while the inversion in wh-questions is less salient and becomes much harder to learn with increasing AoA). This post-hoc interpretation was confirmed by a study with Hebrew L2 (L1 Russian) designed specifically to look at the role of salience while avoiding some of the confounds in the English L2 literature (DeKeyser et al., 2017). As salience plays a much bigger role in explicit than in implicit learning (Reber et al., 1980), this points to a larger involvement of explicit learning in older learners. Evidence for a bigger role for explicit learning in adults than in children and adolescents also comes from correlations between AoA and aptitude (for explicit learning) as measured by the Modern Language Aptitude Test and similar measures. The Hungarian L1 learners of English L2 in DeKeyser (2000) as well as the Russian L1 learners of Hebrew L2 or English L2 in DeKeyser et al. (2010) show a much stronger effect of this aptitude in adult learners. Comparable findings come from studies on the role of declarative and procedural learning, showing that for children procedural learning plays a strong role from the beginning of the learning process (Pili-Moss,

2021), while for adults it only seems to have a sizeable effect at later stages of learning (Morgan-Short et al., 2014; Pili-Moss et al., 2020).

It appears, therefore, that (aptitude for) explicit /declarative learning tends to play a big role in SLA in adulthood, and less in childhood. The literature is not completely clear-cut on this point, however. On the one hand, Abrahamsson and Hyltenstam (2008) found predictive validity for explicit learning aptitude in child as well as adult arrivals, while Granena and Long (2013), e.g., showed a role for implicit learning aptitude in both children and adults. There may be many reasons for these different findings, e.g., the extent to which either children or adults may have received instruction that focused on form, the extent to which the structures tested are likely to be learned explicitly, and the extent to which the testing format encourages use of explicit knowledge. The same factors may explain some inconsistencies in the literature on classroom foreign language learners, to which we now turn.

3. Research in the classroom

As mentioned above, the literature on age effects in classroom foreign language learners shows virtually the opposite of what has been found in so many studies with immigrants learning a second language: instead of “younger is better,” it is “older is better.” The first study to show this is the BAF (Barcelona Age Factor) Project (see esp. Muñoz, 2006, 2007), which compared children who started EFL classes in Spain at different ages, concentrating on those who started at ages 8 or 11 for a longitudinal study. They were compared at two points in time in the longitudinal study, each time after the same amount of instruction for each group (200 hours at Time 1, 416 at Time 2). There was a parallel cross-sectional study with three groups (starting ages 8, 11, 14) and three points in time for testing, here also with all groups having received the same amount of instruction (200 hours at T1, 416 at T2, 712 at T3). The older the starting age, the better the children/adolescents did on vocabulary and grammar during oral testing. For speech perception and phonetic imitation, as well as fluency in writing, however, the difference was less marked (but still in the same direction). Muñoz makes the distinction between morphosyntactic knowledge (more cognitively demanding) and the phonological knowledge involved in speech perception and imitation (less cognitively demanding). The former was a function of age of learning (the older the better), cognitive development, and L1 proficiency, while the latter was also a function of amount of exposure. Differences between groups would vary a bit with testing time, but at no time did the younger learners surpass the older ones.

Similar findings were obtained in studies by García Mayo and García Lecumberri (2003) in Spain, Liu (2007) in China, Pfenninger (2014a, 2014b) in Switzerland, and Jaekel et al. (2017) and Baumert et al. (2020) in Germany, all with EFL learners. Pfenninger (2014b) investigated 200 EFL learners aged 17–20 divided into four groups: early vs. late starters x CLIL (content and language integrated instruction) methodology vs. mixed methodology. For productive and receptive accuracy the four groups showed very similar scores, while for vocabulary (in reception and production) the older learners did better. For listening, with mixed methodology, late and early groups did equally well, but with CLIL the early group scored higher. Pfenninger (2014a) made a different type of comparison: early learners (starting at age 8–9) were compared to late learners (starting at age 13–14), and all were tested at age 13–14, when the older group had only had 50 hours of instruction and the early group had 440 hours. In spite of the big difference in amount of instruction, an advantage for the earlier starters was only found for a few measures. When all were tested again at age 18–10 (for the early starters this meant after 1170 hours of instruction and for the late starters after 730), no differences were found anymore. These findings are even more surprising from the “younger is better” perspective than Muñoz’, because even with much more cumulative exposure the younger starters did not do better than the older starters. Pfenninger attributes this to the higher literacy level of the older learners (a claim that is supported by both groups’ very high correlations between English and German scores on various tests of literacy).

Liu (2007) compared early and late starters in two different studies, one in Beijing and one in Tunxi, and found similar results to Pfenninger’s. In the Beijing group, the early starters initially did better than the late starters, but this advantage faded away between the 7th and the 9th grade. In the Tunxi group, the advantage disappeared even earlier. Liu attributes the difference between the two locations to the higher qualifications of the teachers in Beijing.

Jaekel et al. (2017) also compared early starters (grade 1, $n = 2,498$) and late starters (grade 3, $n = 2,6320$). Students were tested in grade 5 (after 140/245 hours of instruction) and in grade 7 (after 444/549 hours). In grade 5 the earlier starters did better both in reading and listening, but by grade 7 it was the later starters that did better. In this study the quick loss of the early advantage is even more obvious than in Liu’s.

Baumert et al. (2020) compared the receptive skills of three groups of learners, starting in Year 1 of grade school (ages 6–7), Year 3 (ages 8–9), or Year 5 (age 10). The learners were all tested in Year 9 (age 14, in secondary school). The middle group (ages 8–9) did descriptively better than the oldest, and significantly better than the youngest.

4. Integrating the two bodies of research

The conclusion so far, then, seems to be that in the immersion context (immigrants), starting age is the strongest determinant of ultimate attainment, while in the opposite context (foreign language learning with very limited exposure), cognitive maturity is more important. The very different learning contexts (massive exposure and little focus on form vs. minimal exposure and usually strong focus on form) clearly draw on different aptitudes, which may or may not be compatible with the learners' age. In grade school the aptitude for explicit learning is still very much under development (the difference between age 6 and age 12 being very large on this point; in Piagetian terms early stages of concrete operational thinking vs. early stages of formal operational thinking), so learning in the earliest grades is unlikely to be very effective, especially when instruction is of the form-focused kind. The older students get, the more successful this kind of teaching may become, especially for the individuals with high aptitude. Among adult immigrants, aptitude for explicit learning is not a matter of development anymore, but of individual differences, so aptitude for explicit learning will be a more important predictor of proficiency than further differences in age.

This aptitude-treatment interaction, in the sense of explicit aptitude being more important in a context with minimal input and maximal focus on form and implicit aptitude more important in a context with maximal input and minimal focus on form, predicts good results when adults or adolescents with high aptitude for explicit learning have to function in the typical foreign language classroom with 3 or for hours of instruction per week, usually quite form-focused, or when young children are fully immersed, with little or no focus on form for the first couple of years in grade school. In the other two combinations, i.e., young children in form-focused classrooms with limited input, or adults with large amounts of input but little help with focus on form and limited aptitude for explicit learning, success tends to be quite limited, as we know from the research presented in Sections 2 and 3.

This also helps explain why the role of aptitude for explicit learning as a function of age is not always clear-cut, as we saw in Section 2. Several studies have shown a much larger role for aptitude in adults than in children, but if young children are forced to function in an environment with much more focus on form than input, they are forced to draw on whatever explicit aptitude they have. Conversely, for those adults who do get massive amounts of input, but no help with focus on form, and who perhaps also have low explicit aptitude, the role of implicit aptitude may become bigger.

Further evidence for this comes from the juxtaposition of two recent studies. Roehr-Brackin and Tellier (2019) show that in a rather form-focused foreign language curriculum, even 8–9-year-olds draw on their explicit aptitude, in particular

language-analytic ability, which was the strongest predictor of performance. On average learning was quite limited, however (accuracy on the post-test was barely above 50%), presumably because at that age being forced to draw on explicit aptitude is problematic for all but the most high-aptitude students. This study *shows*, through its correlational data, what we can only *assume* for studies like Muñoz (2006) and so many studies after that: when young children are forced to rely on explicit learning mechanisms because of the classroom context and the teaching methodology, not much is learned and much later starters quickly catch up with them. On the other hand, Lee (2020), comparing the ultimate attainment in high-proficiency Korean speakers of ESL for two groups, those who experienced immersion in English in Kindergarten to those who received more traditional instruction in grade school, found that performance for both groups was at near-chance level on a speeded aural grammaticality test, and that reaction times for both groups were about three times as long as for native speakers, but that the early learners did better than the later learners in rejecting ungrammatical sentences, and their reaction times, while very high compared to native speakers, were lower than for later learners. Here apparently the early learners had benefited from a learning experience that drew more on their implicit learning capacities, and that was clear for this type of outcome test, speeded aural grammaticality judgment. The latter is not a pure measure of implicit knowledge by any means, but certainly encourages use of implicit knowledge more than a non-speeded written test or – to put it differently – makes use of explicit knowledge harder than on a non-speeded test.

This may also explain the results of Baumert et al. (2020), who, as explained above, found that the age group 8–9 did descriptively better than the older group (age 10), and significantly better than the younger one (age 6–7). One possible explanation is that the youngest group could not fully benefit from the form-focused teaching and that the oldest group had not had enough time/exposure; only the middle group had both advantages of being a bit better at handling the formal learning AND of substantial instruction/input. The authors present some evidence for a different explanation, however: that early learners could not really benefit from their previous knowledge, because the lack of adequate articulation of the curricula in primary and secondary school. The same two explanations should also be considered in the research of Pfenninger and Singleton (2016, 2019). Their research already showed a big advantage for late starters at their first point of measurement when they had had 50 hours of instruction in six months and the early starters 440 hours in 5.5 years!¹ Moreover, as Baumert et al. (2020, p. 1100) argue,

1. Anecdotally, the same two explanations seem to apply to my own experience as a student in a FLES context (French L2 in Dutch-speaking Belgium). I started French in third grade (age almost 8), two hours per week, and remember learning frequently used phrases, but finding it

yet another interpretation here may be selective intake to the early and late starting groups (as it may also have been in a variety of age effect studies in both foreign language and bilingual education).

Conclusion and implications

The two bodies of research on age effects, with immigrant second language learners and with foreign language learners in the classroom, are two sides of one coin, and one cannot be fully understood without reference to the other. What the data show is not that because of the critical period phenomena shown in a large body of research with immigrants, the solution to foreign language teaching problems is simply to start early. Nor do the classroom data show there is no critical period because younger is not necessarily better in the classroom. What the data show is that there are different learning processes, and that some of these learning processes work better at some ages and in some contexts. It is the interaction of age and context (amount of input, quality of input, amount of interaction, extent of focus on form, and student aptitude) that determines the degree of success: high for very young learners in a complete immersion context or for high-aptitude adults who focus on form and get lots of practice over a long period of time; low when the interaction between age and context is less fortunate: young learners in form-focused contexts with limited input, or adults without instruction or much aptitude for/experience with second language learning. No context, teaching methodology, or age of learning can be called “the best” regardless of the other two variables and of individual differences in aptitude. I can only agree with Long (2013b) that it would be very interesting from a research point of view (and probably very beneficial from an educational/societal point of view) to start with large amounts of L2 quality exposure early on in school, but that of course assumes the availability of teachers who speak the second language very well in all respects, including accent. Achieving that is very hard in “bilingual” or “multilingual” countries such as Canada, Belgium, or Switzerland, and impossible in countries where a large amount of L2 instruction is concentrated these days, but where relatively few native speakers or very advanced speakers of the target language are available, such as China, Japan, and Korea. For

impossible to understand the difference between definite and indefinite articles in the plural (in fourth grade). When starting seventh grade (first year of secondary school, age almost 12), the transition was not only easy for me, but it did not really pose a problem for the one student in class who was starting from zero either, presumably because just about everything the vast majority of us had learned in grade school was repeated, but just at a faster pace, which was possible both because of our age and, at least for the majority of us, also because of previous learning.

the time being we have to acknowledge that just starting early with the means we have is unlikely to yield very positive results in terms of proficiency. There may be socio-psychological advantages to starting early and making students and parents see a second language as an obvious part of the school curriculum, but as a profession we should be careful not to promise more than we can deliver with the human resources at our disposal, lest it backfires and leads to ever more skepticism about the usefulness of foreign language instruction.

References

- Abrahamsson, N., & Hytlenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30(4), 481–509. <https://doi.org/10.1017/S027226310808073X>
- Abrahamsson, N., & Hytlenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249–306. <https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, 53(5), 334–341. <https://doi.org/10.1111/j.1540-4781.1969.tb04603.x>
- Baumert, J., Fleckenstein, J., Leucht, M., Köller, O., & Möller, J. (2020). The long-term proficiency of early, middle and late starters learning English as a foreign language at school: A narrative review and empirical study. *Language Learning*, 70(4), 1091–1135. <https://doi.org/10.1111/lang.12414>
- Bongaerts, T., Summeren, C. van, Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19(4), 447–465. <https://doi.org/10.1017/S0272263197004026>
- Chiswick, B. R., & Miller, P. W. (2008). A test of the critical period hypothesis for language learning. *Journal of Multilingual and Multicultural Development*, 29(1), 16–29. <https://doi.org/10.2167/jmmd555.0>
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533. <https://doi.org/10.1017/S0272263100004022>
- DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55(s1), 1–25. <https://doi.org/10.1111/j.0023-8333.2005.00294.x>
- DeKeyser, R. M. (2012). Age effects in second language learning. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 442–460). Routledge.
- DeKeyser, R. (2013). Age effects in second language learning: Stepping stones toward better understanding. *Language Learning*, 63(s1), 52–67. <https://doi.org/10.1111/j.1467-9922.2012.00737.x>
- DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31(3), 413–438. <https://doi.org/10.1017/S0142716410000056>
- DeKeyser, R. M., Alfi-Shabtay, I., Ravid, D., & Shi, M. (2017). The role of saliency in the acquisition of Hebrew as a second language: Interaction with age of acquisition. In S. Gass, P. Spinner, & J. Behney (Eds.), *Saliency in second language acquisition* (pp. 131–146). Routledge. <https://doi.org/10.4324/9781315399027-7>

- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41(1), 78–104. <https://doi.org/10.1006/jmla.1999.2638>
- García Mayo, M. del P., & García Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language*. Multilingual Matters. <https://doi.org/10.21832/9781853596407>
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343. <https://doi.org/10.1177/0267658312461497>
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, 14(1), 31–38. <https://doi.org/10.1111/1467-9280.01415>
- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>
- Huang, B. (2014). The effects of age on second language grammar and speech production. *Journal of Psycholinguistic Research*, 43(4), 397–420. <https://doi.org/10.1007/s10936-013-9261-7>
- Huang, B. (2016). A synthesis of empirical research on the linguistic outcomes of early foreign language instruction. *International Journal of Multilingualism*, 13(3), 257–273. <https://doi.org/10.1080/14790718.2015.1066792>
- Huang, B., & Jun, S. (2011). The effect of age on the acquisition of second language prosody. *Language and Speech*, 52(3), 387–414. <https://doi.org/10.1177/0023830911402599>
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturation constraints in second language acquisition. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–588). Blackwell. <https://doi.org/10.1002/9780470756492.ch17>
- Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning*, 67(3), 631–664. <https://doi.org/10.1111/lang.12242>
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25(4), 603–634. <https://doi.org/10.1017/S0142716404001298>
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57(1), 1–33. <https://doi.org/10.1111/j.1467-9922.2007.00397.x>
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Lambelet, A., & Berthele, R. (2015). *Age and foreign language learning in school*. Palgrave MacMillan. <https://doi.org/10.1057/9781137525901>
- Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.
- Lee, E.-K. R. (2020). Age of onset, type of exposure, and ultimate attainment of L2 morpho-syntactic sensitivity. *Studies in Second Language Acquisition*, 42(4), 801–823. <https://doi.org/10.1017/S0272263119000731>
- Liu, J. (2007). Critical period hypothesis retested: The effects of earlier English education in China. In J. Liu (Ed.), *English language teaching in China. New approaches, perspectives, and standards* (pp. 170–191). Continuum.
- Long, M. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. H. (2013a). Maturation constraints on child and adult SLA. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 3–41). John Benjamins. <https://doi.org/10.1075/llt.35.01lon>

- Long, M. H. (2013b). Some implications of research findings on sensitive periods in language learning for educational policy and practice. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 259–271). John Benjamins. <https://doi.org/10.1075/llt.35.10lon>
- Mayberry, R. I., & Kluender, R. A. (2018). Rethinking the critical period for language: New insights into an old question from American Sign Language. *Bilingualism: Language and Cognition*, 21(5), 886–905. <https://doi.org/10.1017/S1366728917000724>
- Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in Second Language Acquisition*, 25(3), 351–398. <https://doi.org/10.1017/S0272263103000159>
- Morgan-Short, K., Faretta-Stutenberg, M., Bill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17(1), 56–72. <https://doi.org/10.1017/S1366728912000715>
- Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 1–40). Multilingual Matters. <https://doi.org/10.21832/9781853598937-003>
- Muñoz, C. (2007). Age differences and their implications for practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 229–255). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667275.014>
- Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. M. (2019). Advanced second language learners' perception of lexical tone contrasts. *Studies in Second Language Acquisition*, 41(1), 59–86. <https://doi.org/10.1017/S0272263117000444>
- Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., ... Mehler, J. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121(10), 1841–1852. <https://doi.org/10.1093/brain/121.10.1841>
- Pfenninger, S. E. (2014a). The literacy factor in the optimal age discussion: A five-year longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 19(3), 217–234. <https://doi.org/10.1080/13670050.2014.972334>
- Pfenninger, S. E. (2014b). The misunderstood variable: Age effects as a function of type of instruction. *Studies in Second Language Learning and Teaching*, 4(3), 529–556. <https://doi.org/10.14746/ssl.2014.4.3.8>
- Pfenninger, S. E. (2014b). The literacy factor in the optimal age discussion: A five-year longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 19(3), 217–234. <https://doi.org/10.1080/13670050.2014.972334>
- Pfenninger, S. E., & Singleton, D. (2016). Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research*, 32(3), 311–345. <https://doi.org/10.1177/0267658315624476>
- Pfenninger, S. E., & Singleton, D. (2019). Starting age overshadowed: The primacy of differential environmental and family support effects on second language attainment in an instructional context. *Language Learning*, 69(s1), 207–234. <https://doi.org/10.1111/lang.12318>
- Pili-Moss, D. (2021). Cognitive predictors of child second language comprehension and syntactic learning. *Language Learning*, 71(3), 907–945. <https://doi.org/10.1111/lang.12454>
- Pili-Moss, D., Brill-Schuetz, K. A., Faretta-Stutenberg, M., & Morgan-Short, K. (2020). Contributions of declarative and procedural memory to accuracy and automatization during second language practice. *Bilingualism: Language and Cognition*, 23(3), 639–651. <https://doi.org/10.1017/S1366728919000543>

- Roehr-Brackin, K., & Tellier, A. (2019). The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition*, 41(5), 1111–1131. <https://doi.org/10.1017/S0272263119000214>
- Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, 51(3), 381–398. <https://doi.org/10.1016/j.jml.2004.05.003>
- Slabakova, R. (2006). Is there a critical period for semantics? *Second Language Research*, 22(2), 1–37. <https://doi.org/10.1191/0267658306sr2700a>
- Spadaro, K. (2013). Maturation constraints on lexical acquisition in a second language. In G. Granena & M. H. Long (Eds.), *Sensitive periods, language aptitudes, and ultimate L2 attainment* (pp. 43–68). John Benjamins. <https://doi.org/10.1075/llt.35.02spa>
- Stevens, G. (2004). Using census data to test the critical-period hypothesis for second-language acquisition. *Psychological Science*, 15(3), 215–216. <https://doi.org/10.1111/j.0956-7976.2004.01503012.x>
- Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2013). Effects of age of learning on voice onset time: Categorical perception of Swedish stops by near-native L2 speakers. *Language and Speech*, 57(4), 425–450. <https://doi.org/10.1177/0023830913508760>
- Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2015). Effects of age and speaking rate on voice onset time. *Studies in Second Language Acquisition*, 37(1), 71–100. <https://doi.org/10.1017/S0272263114000151>
- Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language. *Language Learning*, 67(4), 747–779. <https://doi.org/10.1111/lang.12241>
- Trenkic, D. (2007). Variability in second language article production: Beyond the representational deficit vs. processing constraints debate. *Second Language Research*, 23(3), 289–327. <https://doi.org/10.1177/0267658307077643>

Performance on second language speaking tasks

Supports and impediments

Peter Skehan

University College London

This chapter will discuss research into the effects of task characteristics and task conditions on second language performance. It explores a range of task characteristics, such as information structure and familiarity, as well as the operations upon such information. It will also consider the effects of task conditions such as pre-task planning and task repetition. The organisational 'frame' for this discussion will be Levelt's distinction between the stages of Conceptualisation and Formulation. Task characteristics and task conditions will be related to these two stages in speaking, indicating how each stage might be positively or negatively affected. This leads into a discussion of how the task research findings which have been covered are relevant in a pedagogic context. In particular, there is a focus on how second language speakers can achieve a parallel rather than a serial mode of speech, through the use of supportive tasks and task conditions, and the avoidance of impediments which could impair their performance.

Introduction

Anyone writing a chapter for this book cannot start without acknowledging their debt to Mike Long. In the present case what I would like to highlight is Mike's impact in demonstrating the importance of theory, research, and pedagogy, separately, but most importantly, in concert. Obviously, all are needed, separately, and each makes a vital contribution. Theory is vital for explanation, to take us beyond individual contexts, and to show how claims need to be, and can be tested. Research is vital to test such claims, to require precision in the questions which are posed, to foster methodological development, and to show how public inquiry is essential. In his publications, Mike made major contributions in each of these areas. But in addition to that, he saw that theory and research have the underlying purpose to contribute to real-world problems, and in the case of our field, to contribute

to pedagogy. Research and theory need to be grounded in second language acquisition to make their contribution to language learning and instruction. Mike Long saw this clearly, and realised that theory, research, and pedagogy need to work together. Building upon this insight, his influence on our field is immense. No one else has argued more influentially that claims should not be data free and should not be based simply on calls to personal experience (Long, 2007a, 2015).

Taking the general points from the previous paragraph, one can make them more specific for the field in which I have worked – task-based learning and instruction – as the need to develop a theory and a research perspective to try to establish which tasks are better, and why, and also to clarify the criteria to establish what ‘better’ means. Mike’s focus in task-based research was on interaction (Long, 2016), as well as the role of feedback (Long, 2007b) within the wider process of acquisition. The present chapter will not have a similar focus regarding interaction and feedback, but will try to cover theory, research, and pedagogy, the principles outlined above and followed by Mike. In my case the emphasis will be on more psycholinguistic accounts, both of tasks and especially performance. Hence a different way to analyse tasks themselves, to explore the conditions under which tasks are completed, and especially how performance itself is measured – the key to making claims about which tasks might be ‘better’ and for which purposes. Acquisition is also important, but less centrally than in Mike’s thinking – more an extrapolation from the research which will be covered rather than something which is directly investigated.

There will be three main sections to the chapter. The first will simply set the scene, outlining the contexts in which the research to be reported on was conducted, the collaborators, and the setting. There will be the briefest of outlines of how my approach to tasks has changed over the years and what it has emphasised. The second section will outline the actual research, as this is organised to reflect two major organising systems: first, the Conceptualisation-Formulation distinction from Levelt’s (1989) model of first language speaking, and second a comparison between general variables that impact on task performance, in contrast to the sorts of influences that are emerging from individual-based contributions to how tasks are done. Then in the third and main section, the research coverage will be drawn on to offer pedagogic suggestions as to how second language spoken performance can be supported, or how it can be negatively influenced. These sections will, hopefully, provide some guidance for pedagogic decision making.

The research background

There are three sources to the research which will be drawn on. First of all, there are the studies I conducted in collaboration with Pauline Foster, at what was called Ealing College when we started out, and which then moved to King's College, London, and where Parvaneh Tavakoli was my Ph.D. student. Then there are the Hong Kong Research Grant Council funds which I was fortunate to receive, as well as lucky to conduct with talented researchers at Chinese University, Hong Kong. Some of these research collaborators were, in fact, Ph.D. students I supervised, and I will also draw upon their research. Finally, there is the research conducted by others, relevant here if it took a psycholinguistic approach, and measured performance in terms of complexity, accuracy, lexis and fluency.

The research itself falls into two main phases. The first is very much focussed on exploring the influence of task characteristics and task conditions, and this in a rather exploratory manner. These earlier studies researched task types, such as narrative and interactive formats (Foster & Skehan, 1996), and also finer grained task characteristics, such as degree of information familiarity and personal relevance (Skehan & Foster, 1997). There was also a lot of attention given to the measurement of performance within the psycholinguistic framework, emphasising complexity, accuracy, and fluency (Skehan & Foster, 2005), and with the beginning of exploration of the complexities within each of these (Skehan & Foster, 2008). To the extent that theory was important at this stage, it was only to import, from contemporary cognitive psychology, the notion that attentional resources are limited, and that this can mean, for the second language speaker, that emphasising one performance area *may* be at the expense of other areas, with a degree of focus here on the tension between complexity and accuracy.

The second (and current) phase saw some important extensions to the earlier research. Perhaps most important was to embrace the Levelt model of first language speaking, and to explore how relevant it might be for second language speaking. De Bot (1992) and Kormos (2006) had offered major accounts of how this extension might work, and the three stages in the Levelt model (Conceptualisation, Formulation, Articulation) seemed to transfer meaningfully to the second language case. In addition, monitoring (Kormos, 1999) was perhaps even more relevant than in first language speaking. Finally, regarding Levelt, the implications of differences in the resources of the second language mental lexicon (SLML) were vital. Beyond the Levelt connection, an interesting development was that the contrast between tasks (types and characteristics) and conditions (such as planning, repetition, time pressure, post-task effects) was sharpening (Skehan, 2016), with conditions seeming to generate more reliable and wider-ranging findings. More theoretically, there was a move from simply discussing trade-off between performance areas to portraying

psycholinguistic resources in terms of the Limited Attentional Capacity approach (Skehan, 2018). This emphasises that trade-off is not always inevitable, but that careful task and condition manipulation can often overcome any tension between accuracy and complexity. Measurement issues also evolved, with greater understanding of the sub-dimensions within CALF and how performance could be more subtly measured (Skehan, 2018, Chapter 10). Finally, there has been a move to raising the importance of the individual, and how different styles and predispositions might impact upon the nature of task performance (Skehan & Shum, 2017; Pang & Skehan, 2014, 2021).

Task research through a Leveltian lens

This section will cover research into tasks and task conditions conducted from a Limited Attentional Capacity perspective. The purpose will be to establish the empirical grounding for the more pedagogic claims made in the next section of the chapter. There are three organisational ‘frames’ which follow from the previous section. First, the Leveltian (1989) distinction between Conceptualiser and Formulator processes will be fundamental, since this distinction, between ideas, their development and organisation (Conceptualisation), and then their expression through language (Formulation) recasts the way we can look at task and task condition effects. This is represented visually in Figure 1.

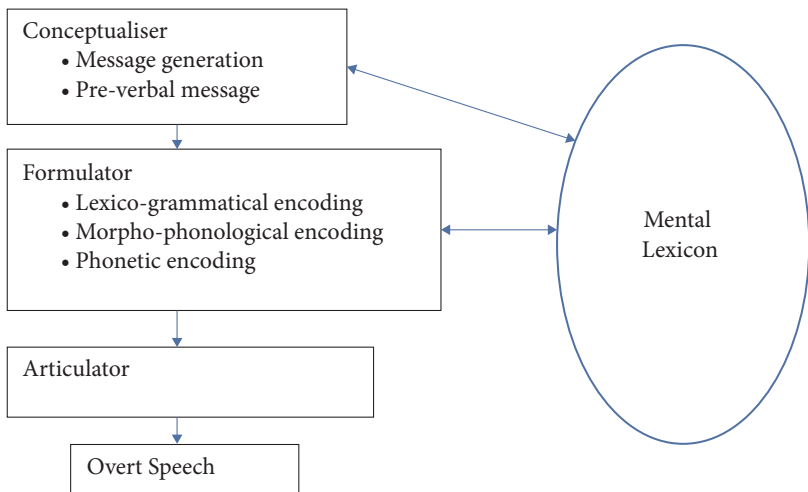


Figure 1. Levelt model of first language speaking

But second, within each of these macro stages, there is a distinction between research which has emphasised task and task condition effects, on the one hand, and research which has explored how individuals, through style or preferences, might approach task performance differently, on the other hand (Pang & Skehan, 2021). Finally, the distinction between tasks and conditions will be sharpened because it helps to organise the presentation of research findings, and also because it connects very well with the Leveltian distinction between Conceptualisation and Formulation.

Conceptualisation and its impact on task and task condition variables

The first set of variables to consider here relate to the information that underlies a task. Broadly the research findings suggest that familiar information and more concrete information ease the task of the Conceptualiser and enable pre-verbal messages to be developed more quickly and with less demand on attentional resources (thus freeing up some attention for subsequent Formulator operations: Foster & Skehan, 1996; Bui, 2014). The impact in performance is likely to be on raised fluency and accuracy, and perhaps lexis also. Complexity, in contrast, may be lowered. In contrast, less familiar information or more abstract information is likely to require attentional resources for the Conceptualiser stage, and while this may lead to more complexity, the other performance areas of accuracy and fluency may suffer (Skehan & Foster, 1997).

Staying with information, a range of research studies (Skehan & Foster, 1999; Tavakoli & Skehan, 2005; Tavakoli & Foster, 2008) have explored the effect of information organisation on task performance (and this will be relevant to Formulator operations also). A puzzle confronting the second language speaker, when required to do a task, is how to handle the need to keep on speaking, while simultaneously coping with ongoing Formulator demands *and* while still generating the ideas that keep speech going (Wang, 2014). Key to this is the relationship between successive ideas within the flow of speech. It is here that Organisation is important, and particularly in the research data base, the role of Task Structure. A series of studies (Tavakoli, 2004; Wang & Skehan, 2014) have shown that clearly structured narrative tasks, especially, where structure is defined in terms of Winter's (1976) problem-solution structure (Kobayashi, 2009), support higher levels of performance. The structure seems to enable larger planning units to be used, and this promotes a more effective flow of discourse, as well as a series of connected pre-verbal messages sustaining greater complexity in speech, and (though this relates more to Formulation), to higher accuracy and fluency.

With ideas, we also have to consider the operations that may be required, and this has generated a significant literature. Perhaps most prominent here has been research stimulated by the Cognition Hypothesis (CH), (Robinson, 2011a, 2015), which proposes that task complexity is the driver for greater structural complexity *and* accuracy in performance. The key to this in the CH is the concept of resource-directing variables, i.e. task characteristics which push the learner to engagement with language. Included here are such factors as time perspective, reasoning demands (causal, spatial, intentional), number of elements, and perspective taking). The evidence on the effects of manipulating tasks through these variables is mixed, as meta-analyses have indicated (Jackson & Suethanapornkul, 2013; Malicka & Sasayama, 2017), with some effects on performance, but rarely both complexity *and* accuracy. But the broader issues of task effects on performance is important, and other variables have also been shown to have an impact. Requiring second language speakers to transform the language in a task (Foster & Skehan, 1996), or to integrate information from different sources (Tavakoli & Skehan, 2005) has been shown to raise structural complexity, for example.

Broadly, then, research findings are accumulating usefully on the impact on performance of various characteristics of the task itself. In addition, though, one can explore how the conditions under which a task is done might impact performance, and here, it can be argued (Skehan, 2016) greater influences are apparent. For example, a significant sub-literature now exists on the effects of planning on performance. This shows that giving learners time to plan, pre-task, is consistently associated with higher complexity and fluency, with medium to large effect sizes (Ellis, 2009; Skehan, 2018), while accuracy is affected, but less consistently and with lower effect sizes. More recently there have been studies which have broadened the nature of pre-task activities to go beyond simply time for planning and these have shown effects for activities such as modelling (Kim & McDonough, 2011). These too show that what happens before a task has an impact on the nature of the performance during the task, and that this is largely beneficial. A major and newer strand of research concerns the impact of task repetition (Bygate, 2018). Studies exploring this area have shown that giving second language speakers the opportunity to repeat a task can have strong and beneficial effects on performance, raising complexity, accuracy and fluency (Wang, 2014), but not so much lexis. In addition, there are suggestions, such as Lambert, Kormos, and Minn (2017) in a study with multiple repetitions, that there is a progression, within the successive repetitions, suggesting that first complexity benefits, then fluency, and finally accuracy. Skehan (2018) also argues that repetition fits neatly within a Leveltian interpretation. Following Bygate (2001), the original performance, he suggests, enables ideas within a message to be related to one another, and even condensed, meaning that the repeated performance can be significantly greater in complexity, and with the connections between

the ideas which are more readily available leading to greater fluency. (There are also repetition implications for the Formulator, below.)

Another influence on Conceptualiser operations is freedom to choose what to say and how to say it, relative to situations where the speaker is rather constrained by the requirements of a task. In the first case the speaker has considerable scope to shape what will be said, what its emphases might be, how ideas can be developed and connected, and even how strengths can be played to. There is even scope to draw upon personal experience more (Wang & Skehan, 2014). This contrasts with a situation where particular input has to be covered, even, as with instructions in a narrative, to cover as much detail as possible (Skehan, 2018). Speakers, in such a situation, are being forced to engage with ideas for which they might not have sufficient linguistic resources. Facility with producing pre-verbal messages in these two situations may be markedly different to those where there is more opportunity to choose, to select and draw upon what is familiar (Wang, 2014).

Conceptualiser operations: The individual

The process of speaking is almost always demanding for second language speakers. It often is so for first language speakers also. One may, therefore, consider whether some people have different approaches to dealing with the problems of keeping up with real-time communication, and whether some of these different approaches lead to more effective performance. Some insights into these possibilities come from qualitative and quantitative research into the effects of pre-task planning. In a first study, Pang and Skehan (2014) used a qualitative approach to ask learners to say what they had done during a ten-minute planning period. They report that second language speakers can vary in how they use the opportunity for ten minutes pre-task planning on a picture-based narrative task. Two aspects of this stand out. First of all, some speakers had much more realism about how they could transfer the results of their planning to their actual performance. They had much greater self-awareness regarding their memory limitations, and so in the planning time they concentrated on ideas rather than detailed language. In addition, they did not try to plan more than they were likely to remember (and then be able to use). Second, some speakers were more likely to use planning time with an awareness of the length of time they were trying to prepare for, and as a result, focussed more on generating a *series* of pre-verbal messages, with more of a discourse orientation. Some speakers even found ways of imposing structure on what they would be going to say. Learners who did these things tended to produce performances with higher structural complexity, and, interestingly, higher accuracy, as though they had used the planning to anticipate the time pressure that would be experienced during performance.

In a subsequent study, Pang and Skehan (2021) confirmed and extended these results. In this study, with the same participants, a second, interactive decision-making task was also involved. Using the technique of cluster analysis with the performances on the narrative and decision-making tasks, they report that, especially for the narrative, there was a contrast amongst the participants between those who had a discourse-oriented style and those who had a more clause-oriented approach. The discourse-oriented cluster of participants were characterised by more structural complexity, if this was measured through subordination, but shorter clauses. They were also more fluent, faster and used less repair. The clause-oriented cluster produced lower subordination complexity, but for a number of words-per-clause measure, they produced longer clauses. But they were less fluent and produced clearly more repair (reformulation, repetition, false starts). In other words, some second language speakers seemed to have used planning opportunities to assemble a number of pre-verbal messages, and focus on the wider discourse picture, rather than the detail. The clause group, in contrast, were more concerned with the details of each utterance, tended to use less subordination, and with the longer clauses they produced, encountered problems which needed attention. Interestingly, there was some consistency of style across the two tasks, with roughly half the participants falling in the same cluster (discourse or clause) with each task, and very few showing any mismatch of cluster (e.g. discourse style in one and clause style in the other).

Formulation: Variables

We turn next to exploring variables which impact on Formulator processes. As a reminder here (and see Figure 1), Levelt (1989) proposes that the Formulator takes a pre-verbal message as input, and then the ideas in that message drive processes of lemma retrieval and syntax building. Vital for this is access to lemmas in the second language mental lexicon (SLML). In addition, the speed of access to such lemmas, linked to the richness of the information that they contain (meaning, syntactic implications, phonology, collocation, connotation, and more), is central to how speaking proceeds in the first language case. So, the issue we have to consider here is how these processes might be different with second language speakers.

Most naturally, this is likely to involve issues like familiarity and frequency of use (Bui, 2014). The more that lemmas triggered by the pre-verbal message have been used before and often, the easier will retrieval be, and the firmer the foundation for syntax building (Foster & Skehan, 1996). A similar influence will come from proficiency level (Bui, Skehan, & Wang, 2018), since as proficiency grows, so, almost certainly, will the SLML, and accordingly the stock of items which is

available, as well as speed and richness of access. These claims are straightforward. Less so are others, such as the role of repetition and time pressure during speaking.

Repetition has already been discussed briefly in relation to the Conceptualiser. It also has great importance in Formulation. In an ideal world, a pre-verbal message will trigger relevant lemma access (Kormos, 2006, 2011), and that access will be rapid and complete. In many second language cases such access may be far from guaranteed. However, it is possible that in a first performance there will be partial lemma access. On this original occasion the access may occur only partially and slowly, and may not lead to satisfactory spoken performance. But it can leave a trace, a prime, which, in the repeated performance can be capitalised upon (Skehan, 2018). The repeated performance can then take as input not the 'cold' elements of the pre-verbal message, but instead what has been semi-retrieved but not used. In this way the repeated performance can more robustly deal with performance time pressure and be more effective. Assuming that lemma retrieval in the repeated performance is not simply faster, but also more complete, the consequence is that syntax-building, articulatory planning, discourse implications, collocational information and so on are all eased, leading to improvements in processing and performance. The burgeoning repetition literature bears this out. Wang (2014) has shown striking effects on Formulator (i.e. accuracy, fluency) dimensions of performance in this respect, with immediate repetition. Lambert et al. (2017), as discussed earlier, have proposed that a series of repetitions make sense in terms of the Levelt model, with complexity (the Conceptualiser) first of all, and then fluency and accuracy (with more Formulator linkage). Fundamentally, then, repetition has a marked effect in easing Formulator operations for the second language speaker.

Another approach takes a different tack to understanding influences on the Formulator, and is concerned with the time pressures that exist for second language speakers. Ellis (2005) makes a distinction between strategic (i.e. pre-task) planning and on-line planning. The former we have covered already. The latter he proposes as the 'on the fly' planning that is possible during speaking when there are more relaxed time pressure conditions. The generalisation from the relevant studies (Ellis & Yuan, 2005; Yuan & Ellis, 2003) is that accuracy is raised when there are on-line planning opportunities. To make the link with the repetition discussion above, it seems that providing second language speakers with just a little more time for the lemma retrieval stage confers considerable benefits for Formulator operations. In particular, the syntax-building which follows lemma access becomes more complete and extensive, more able to avoid error, and even provide a better foundation for monitoring. An interesting study here is Wang (2014) who explores the combined influence of pre-task *and* on-line planning. She reports that in this condition complexity and accuracy are both raised. It appears that with this combination

the speaker has something to say (through the pre-task planning) and then time to be effective in how they speak, (through the greater time available for on-line planning). In other words, Conceptualiser and Formulator can work together in harmony, what Wang, Skehan, & Chen (2019) term the Conceptualiser-Formulator Balance principle. Ellis (2005) also provides a slightly different take on pre-task planning. In addition to the different forms of Conceptualiser-linked planning covered earlier, he also draws attention to what he terms rehearsal planning – planning of specific language elements, including retrieval of lemmas that are thought to be important for the actual performance. Such planning is vulnerable to forgetting, as argued earlier (Pang & Skehan, 2014), but it is likely that some retention of specific items is possible and likely to have Formulator benefits.

The Levelt model next discusses syntax-building, and this driven by the previous lemma retrieval. In fact, it is difficult to separate out the variables which impact upon lemma-retrieval from those which impact upon syntax-building – all the above apply, and the richer and faster the lemma-retrieval, the greater the foundation for effective syntax-building. There are, though, some different emphases. Less time pressure is still important, enabling lemma retrieval to be deeper and syntax building to be more effective (Wang et al., 2019). As a result, language can become more complex and it can also be more accurate. Related to this, structured tasks, where the detail of current performance has a clearer relationship to the wider macrostructure, also confer benefits (Foster & Tavakoli, 2009). The structure, and the clearer part-whole relationship, enables speakers to focus more effectively on current syntax-building, since they can realise more clearly how they can move back to the wider discourse picture (Skehan, 2018).

Another interesting variable here is the role of post-tasks. One potential disadvantage of giving learners tasks to do is that they may sacrifice form in order to get the task done. Ironically, the more interesting the task, the more likely it is that this will happen! One pedagogic proposal to combat this is to give learners post-tasks to do, so that they realise there are connections between the actual task performance and what will happen later (Lynch, 2001, 2007), and the anticipation of the post-task can change how the task itself is done. A series of studies have explored this possibility. Skehan and Foster (1997) showed that telling speakers, in an intact class, doing tasks in pairs, that some of them, randomly chosen, would later be required to re-do the task publicly, in front of the teacher and class, led to higher accuracy on two of the three tasks used, with no change on other performance measures. Foster and Skehan (2013) confirmed this result, but used the different post-task condition of transcribing one's own (recorded) performance. This condition led to higher accuracy on an interactive task (as before) and on a narrative (in contrast to the earlier study where no significant effect was found with this task type). In addition, the interactive task showed higher structural complexity also. Li (2014)

confirmed these findings and extended them by showing that, in addition to the general effects reported in the previous studies (Skehan & Foster, 1997; Foster & Skehan, 2013), *pair*-based transcribing, to a limited extent, increased complexity, while transcription-plus-analysis raised accuracy but had a slightly negative effect on complexity. We can, therefore, draw the conclusion that syntax (certainly accuracy, and to a lesser extent, complexity) can be ‘nudged’ within Formulation. It appears to be the case that anticipation of what is to come can cause attentional resources to be directed, at least to some extent, within ongoing Formulation.

Formulation: Individuals

The central question here is whether there are consistent aspects of performance at the Formulation stage. We saw with the Conceptualiser that there were consistencies in developing ideas with a discourse versus a clause orientation. With the Formulator, the focus is on aspects of performance such as accuracy and fluency. With the first of these, there is little from the literature, perhaps surprisingly, to suggest that accuracy levels are very consistent over different tasks. In within-subjects studies, where several tasks are done, the cross-task correlations, although not close to zero, are not notably high (Skehan, Foster, & Shum, 2016). It appears that task effects are the greater influence with accuracy. The situation is very different with aspects of fluency. In such within-subjects studies there are very high cross-task correlations for indices of repair, such as reformulations, or repetitions, or false starts (Skehan & Shum, 2014). The same is true for unfilled pauses, at clause boundaries, and particularly, mid-clause. In addition, there is some consistency in speed of speech.

To put this another way, tasks and task conditions do not have the major influences on many aspects of fluency. In contrast, personal style seems more important: personal predisposition rather than the result of what they are asked to do. In other words, pausers are pausers, repairers are repairers.

Supports and impediments to second language speaking

We have now covered quite a range of research results, focussing on task performance, but organised in terms of the Levelt model of first language speaking applied to the second language case. Broadly, it can be claimed that the first language stages hold up well in the second language case. Speakers still have to engage in Conceptualisation processes, whose major outcome is one or more pre-verbal messages. The implication of the pre-verbal message for Formulation, for lemma retrieval and subsequent syntax-building is the same, but with the very important

qualification that the basic resource which is drawn on, the second language mental lexicon, may not be anything like as impressive as the first language mental lexicon.

One specific consequence is that the parallel functioning that is typical of first language speech, with different stages working simultaneously, and in concert, becomes an ideal but one that is often not realised (Skehan, 2015). Difficulties at the Formulation stage, the result of the less impressive mental lexicon, mean that the demands that are placed on the Formulator for rapid efficient performance are often not met. As a consequence, parallel functioning is often supplanted by a serial mode of communication, in which one stage, the Formulator, hogs attentional resources, to the detriment of the other stages which can no longer work in parallel.

This analysis is the starting point for the present section. We will explore how speaking, given these constraints, can be supported, and how a parallel mode of functioning, resembling first language speaking, can be achieved more often. Broadly, this involves three things. First, it is helpful to know how the Conceptualiser can be supported, so that it makes more reasonable demands on the Formulator, and also how it can handle the situations when problems occur and serial operations dominate. Second, there is the need to support the Formulator, so that it too is more able to sustain parallel processing, and handle Conceptualiser demands. Finally, we need to integrate the role of the individual more clearly in this speech production process, given that second language speakers vary in style and may have natural predispositions to speaking which may, or may not, be helpful, given these more general constraints. The rest of the chapter will, therefore, be more practically oriented, and make suggestions which could impact on pedagogy. But it is important to say that these more practical suggestions are rooted in the research which was covered in the last section. As Mike Long would argue, the claims are not data-free.

Providing conceptualiser support (and avoiding problems)

One set of suggestions to provide Conceptualiser support is based on ‘loading up’ the Conceptualiser. We have seen that in the Levelt model of speech production, producing *sets* of pre-verbal messages is normal. It is clear from the research covered earlier that, if speaking is quite probably going to be derailed by Formulator problems ‘downstream’, inducing second language speakers to ‘load up’ with sets of pre-verbal messages is a way of pre-empting the difficulties that might occur. If they load up in this way, even though speakers will need to address these Formulator problems, they do have the starting point of a robust set of ideas to restart the discourse. In the research sections we saw that some second language speakers are more likely to do this naturally (Pang & Skehan, 2021). Perhaps it is with clause-oriented speakers that we need to work most on this issue, and there are some practical methods of doing so.

Most prominent here are planning and repetition, two easy to implement conditions with tasks. Taking planning first, and based on the findings of structural complexity and fluency effects (Skehan, 2018), it does seem to be the case that it is the Conceptualiser that is usually advantaged through this condition. In other words, the time spent planning can be effectively linked to ideas and their development. Linked to this is the qualitative data in Pang and Skehan (2014) which suggested that people varied in what they emphasised during the planning period, with some having a greater sense of realism about what they would be able to remember from planning for the actual performance, and the tendency that those who used the time for idea preparation, activation and development were more effective subsequently. The implication is that not only is planning beneficial, but that some approaches to planning are better than others, and there is even the possibility of training people to use planning more effectively so that Conceptualiser 'loading up' is more effective. Repetition also has an effect on Conceptualiser function, and it suggests that second performances benefit from the activation of ideas from the first performance. Bygate (2001) argues that this can lead to elaboration of ideas in the repeated performance, as well as the possibility of more condensed speech. As we will see below, repetition has even stronger effects on the Formulator.

Structured tasks, too, are a means to help Conceptualiser operations load up. Such tasks make the connection between successive stages in a speaking task more obvious and even unavoidable. As a result, they counter the tendency to plan and execute in small, clause-limited units, and push the speaker to be more likely to see the connection between an immediate utterance and the wider discourse structure (Skehan, 2018). Structured tasks also have an important role in combating a serial mode of speech because they provide a form of scaffold that can be picked up again after there has been communication difficulty. In other words, if the Formulator encounters problems, and has to resort to using up attentional resources, the more structured a task is, the more easily can the speaker return to where they were last able to sustain parallel processing, and take up more fluent speech again (Wang & Skehan, 2014).

The second set of suggestions to support Conceptualiser operations is broadly concerned with freedom, with a lack of compulsion in exactly what needs to be said. There are two aspects to this. First, freedom means that the speaker does not have to worry about tripwires, about specific lemmas that have to be used if a task is to be done well, of the sort that would need cumbersome circumlocution to replace the lemma concerned. There needs to be scope, that is, to shape the elements of discourse. Second, freedom means that speakers have room to choose what they say, and play to their strengths and avoid their weaknesses, and therefore pre-empt problems. If they are more comfortable with a particular style of communication, a discourse orientation or a clause orientation, for example, they can rely on such

a style (Pang & Skehan, 2021). Or, if they can divert a task to an area where they have relevant knowledge, they are free to do so (Bui, 2014).

We can illustrate this fairly general point with two specific task-linked examples. First, we can consider there-and-then tasks (relative to here-and-now tasks). Such tasks are done without the stimulus material being visible (Wang & Skehan, 2014). Obviously, a memory issue immediately arises (and this is why such tasks are designated as more complex within the Cognition Hypothesis (Robinson, 2011b, 2015)). But, on the other hand, lack of the immediate presence of the stimulus material confers freedom for the speaker as to how to approach the task. There is an opportunity to shape what will be said, to omit things if so desired, and to re-package what is to be described. In contrast, with a here-and-now condition, there is considerable immediacy of stimulus material, and the speaker has to confront the 'unavoidable lemma' problem. As a result, one condition here, there-and-then, is associated with much greater freedom, and likely easier Conceptualiser operations, where easier does not mean saying less complex things, but rather the opportunity to develop richer pre-verbal messages.

A similar connection to freedom can be seen with the monologic-dialogic contrast (Skehan, 2016). Typically, monologic tasks, especially narratives, are stimulus-driven, and so while details of what is to be said are not fixed, there is an overall set of constraints on what needs to be covered. When instructions talk about providing as much detail as possible in some such tasks, the constraints are even tighter. In contrast, dialogic tasks, even those which require fixed or agreed upon outcomes, give much greater latitude for participants to take their own path to completing the task. In more open interactive tasks there may be even greater freedom as to how the task is done, and indeed, pairs and groups can collaborate to take a task in all sorts of unforeseen directions (Foster & Skehan, 2012; Skehan & Foster, 2005). The result is that speakers can shape their contributions much more freely, playing to strengths, and doing what comes easiest. (In addition, although there is a need to pay attention to the interaction, there is also the added possibility of secretly using interlocutor speaking time as an on-line planning opportunity!)

The final area where Conceptualiser support may figure concerns the issue of individual style, discussed above. We saw evidence from a planning study of a discourse-clause contrast (Pang & Skehan, 2021), with this linked, in the former case, to a loading up of pre-verbal messages and greater subordination complexity and fluency, while the latter was associated with a clausal orientation, less fluency and more repair. Broadly, the discourse orientation here, based on individual style, seems a more effective approach. This suggests that, although clause-oriented speakers may not find this style natural, it might confer benefits if they could, at least some of the time, use it more. One practical approach might be to try to train clause-oriented speakers to use discourse planning more effectively. Second

language learners vary considerably in how they use planning opportunities (Pang & Skehan, 2014). It may well be that such speakers are malleable and that there is scope to provide them with instruction as to how they might change. In addition, one might achieve the same sort of effect by drawing on the suggestions made earlier in this section. Using there-and-then tasks might achieve this, as would using more structured tasks. One might also consider changing the instructions given when narratives are involved to move away from encouraging an excessive reliance on detail. The intention in all this is not to fundamentally change a preferred communicative style, but to widen the repertoire that is available to second language learners so that they can make more effective choices.

The section so far has focussed on Conceptualiser support. But of course, Conceptualiser impediments are an equally important issue, even though we are, fundamentally, only dealing with the opposite of what has been said so far. Just to be clear, there are pedagogic interventions which:

- Make it more difficult to load up with sets of pre-verbal messages
- Deprive the second language speaker of freedom of choice in what they say
- Encourage a clause-emphasising mode of communication

The first of these concerns things like the use of unstructured tasks, or not using planning or repetition opportunities, or using difficult and unfamiliar material. The second suggests ideas and messages dominated by non-negotiable input, too great an emphasis on here-and-now, detailed tasks, hard-to-avoid specific vocabulary, and time pressure so that there is little scope for on-line planning. There may also be little scope for metacognition, as when speakers try to engineer pre-verbal messages to play to their strengths. The third of the above bullet points deals with situations where a clause-at-a-time approach is encouraged, where it is difficult to achieve any sort of macrostructure perspective, and even where personal style to achieve a discourse perspective is impeded. The broad consequence of all this is that a parallel mode of speaking is made more difficult, and it is more likely that a serial mode will have to be used (Kormos, 2006).

Formulator support

The starting assumption for this section is that the Conceptualiser will deliver a pre-verbal message (or set of such messages). This will have implications for lemma retrieval and syntax-building. From a pedagogic perspective, the concern is to ease this transition between the two macro stages of the model, and help the SLML work effectively (Skehan, 2018). The central problem for the second language speaker is a difficult one, but it can be alleviated. There are several possibilities.

Time is of considerable importance here. Providing just a little more time means that more lemmas are retrievable, and, equally importantly, the information they contain is more extensive (Wang, 2014). On-line planning opportunities and less pressured conditions are vital to enable this (Ellis & Yuan, 2005). So is less input, since this too reduces the pressures on what the Formulator has to achieve. Opportunities to avoid remorseless need for speech can also make a major contribution (as can happen naturally during interactive tasks) (Skehan & Foster, 1997). Using there-and-then conditions could also be an important part of this (Wang & Skehan, 2014).

Opportunities for preparation, as well as prior engagement with material, can also make a huge contribution (Bui, 2014). The usual suspects are involved. Planning, if it is rehearsal planning (Ellis, 2005), and if there is realism on the part of the planner about what can be retained and used (Pang & Skehan, 2014), is very important in easing the demands on the Formulator as the specific language to prepare to express ideas is made salient. Repetition is likely to make its greatest contribution here, and it presents an interesting contrast with planning (Skehan, 2018). With pre-task planning, it is possible that speakers will think about particular lexical items, but it is likely they will do this in a rather shallow, unfocussed way – much better than nothing, but with no depth of engagement. Regarding repetition, in a first performance, actual speaking has to take place and so if a lemma is not completely retrieved (as may often be the case), what is retrieved may prime the repeated performance and make it easier and also faster. But the attempt to grapple with how the lemma fits into syntax building may alert the speaker (assuming some details of the attempt are retained) to which aspects of the lemma were only partly retrieved. Material may have been ‘almost there’, but given the fleeting nature of speaking, not properly used. In the repeated performance the half-retrieved elements may become completely available. Broadly, then, the Formulator’s need to access the SLML may be eased considerably.

Freedom and choice are important at this stage also. Difficulty may only become apparent at the point of lemma retrieval, and while the options just described may ease these difficulties on many occasions, there will be times when the speaker’s best option is to revise the detail of what is to be said (something that the Conceptualiser stage may already have started to engage with). The more there is freedom to chart an individual path to what is to be said, the more effectively this can be done (Skehan, 2016). Hence, once more the importance of things like there-and-then tasks, or, more generally, lack of input dominance or tight control to express particular ideas with unusual precision and detail. Interactive tasks again have advantages here (Skehan & Foster, 2008).

So far, the Formulator supports have been directed at easing and mobilising lemma retrieval and syntax-building operations, with all of this founded on the

assumption that attentional limitations have an important impact upon performance. But there is some evidence, as we have seen, that attention can be directed, at least up to a point, and that within the dimensions of performance, particular areas can be advantaged (Skehan & Foster, 1997; Foster & Skehan, 2013; Li, 2014). The studies we covered earlier suggest that using particular post-task conditions can achieve this – anticipation of the post-task to come can nudge second language speakers to giving greater priority to form, specifically accuracy and perhaps complexity. So, the next important area for pedagogic application is the way post-task activities such as the ‘threat’ of a public performance, or the requirement to transcribe some of one’s performance can lead to a greater engagement with the form of language itself, and induce speakers to monitor a little more and avoid making error. We also saw that with an interactive task, this may also have the effect of raising structural complexity. In other words, there is a support for Formulator operations here, but with more of a directing role – insinuating pedagogic norms into an activity where otherwise they would be de-emphasised.

Post-task effects figure because it is assumed that attentional limitations are important, but that careful task and task conditions design can mitigate these limitations. We can develop this further by looking at how combinations of person, task and task condition variables can achieve the same sort of effect – modifying attentional limitations. In some ways, the challenge is posed particularly clearly by the goal of raising accuracy *and* complexity simultaneously – one of the major challenges in a task-based approach. For example, Wang (2014) demonstrated that combining pre-task with on-line planning can lead to higher complexity and accuracy. (She also demonstrated that immediate repetition can produce similar results.) Foster and Skehan (1999) reported that teacher-led planning (rather than individually-based planning) can produce the same effect. These various conditions, that is, seem to enable Conceptualiser and Formulator to work together. Finally, there are suggestions that Formulator operations are improved with the combination of an individual difference and a task condition. This is the case for greater working memory together with on-line planning opportunities (Skehan, 2022). (Pre-task planning is not advantaged by greater working memory.) The same applies to proficiency level – higher proficiency second language speakers seem to benefit more from on-line planning opportunities than do lower proficiency speakers (Bui, Skehan, & Wang, 2018).

It is also worth a brief reminder that, if we return to the focus on parallel vs. serial processing, if second language speakers encounter problems, and need to resort to serial processing, an important issue is how effectively they might regain parallel processing (or not). We have seen several advantages for using more structured tasks already. This is another. If tasks are more structured, they have the considerable advantage that stages within the discourse to relaunch parallel processing are

more common and more obvious. The underlying problem is to have been derailed from fluent speaking, for some reason, and then to need to re-find a thread that has been lost. With structured tasks, when a particular sub-section has come to an end, it will be much easier to start again with a new section, with a clean slate, as it were.

As with the Conceptualiser discussion, there is a reverse situation to consider to all the Formulator supports we have just been covering, and these are worth mentioning, if only briefly. Broadly, little time to retrieve during speaking is going to predispose a serial mode of speaking, as will more input, and unstructured tasks. Lack of preparedness is also important, as with unfamiliar material, no planning or repetition opportunities. Little freedom in how a task will be done or developed is also difficult. Difficult combinations of variables are also likely to impair chance of parallel processing, e.g. low working memory or low proficiency even though there is on-line planning; on-line planning or pre-task planning, but not both.

Conclusions

As mentioned in the Introduction, the focus of this chapter has not been on acquisition, despite Mike Long's major contribution to that area. (Acquisitional issues are addressed in Skehan, 2007, 2013, 2018, where it is argued that pre- and especially post-task stages are the key to interlanguage development). Instead the focus has been on the researchability of task performance, and of how research findings can have pedagogic applications, another issue of major importance to Mike. The central thrust of the chapter has been how task research can help us to identify variables that can lead to more effective second language speaking. The broad assumption has been that while there may be underlying interlanguage competence, it is also important to develop an ability for use (Skehan, 2018), such that second language speakers are able to use underlying competences to maximum effect. As a central aspect of this, it is proposed that we need to find out how such speakers can achieve a parallel mode of communication, and sustain this mode as much as possible.

Drawing extensively on the Levelt (1989) model of first language speaking, the discussion explored the research into the Conceptualiser and Formulator stages in second language speech, and then how this was relevant to pedagogy. Regarding the Conceptualiser, it was proposed that it is important for second language speakers to engage in idea generation, leading to a pre-verbal message (or messages), while having some awareness that there will be problems to come at later stages, and so to be prepared for such difficulties. In particular it was suggested that producing sets of pre-verbal messages might be helpful, and lead to a 'big picture' approach to speaking, which would function in a more robust manner. In addition, task characteristics (e.g. task structure, freedom) and task conditions (planning, repetition)

which support this approach were covered and led to a series of precepts for supporting more effective speaking.

The Formulator stages was approached similarly, and so the variables which emerged (eased time pressure conditions, preparedness, freedom, and potential to recovery from derailed speaking plans) had in common that they recognise the limitations of the second language mental lexicon, and the potential these limitations have to push the speaker into a serial mode of communication, where one stage, lexical retrieval, or slow syntax-building, might be very demanding of attentional resources, and cause disruption. Again, a set of precepts were offered to make the Formulator stage as effective as possible, given underlying lemma limitations.

There are, though, some more complex issues with second language speaking which merit brief mention. First of all, there is the issue of automatism (DeKeyser, 2007). Generally, in second language pedagogy, this is viewed through a declarative-to-procedural lens where focussed material is first presented and then activities are organised to enable the focussed material to be produced with less attention and more automatically (DeKeyser, 2020). But the emphasis is on *focussed* material. Tasks, by definition, do not start with such focussed materials, but there is still a need to develop more automatic, speaking skills. In the present case, these concern the broader issues with Conceptualiser use, such as the generation of pre-verbal messages; access to the second language mental lexicon for the Formulator to work effectively; and perhaps recovery from problems in the Formulator. Using these broader processes also involves automatism, not of specific aspects of language, but of more general speaking stages. So the sorts of variables and influences that have been covered in this chapter are also, in their way, concerned with automatism, but a different level to that which is used in the practice literature. Automatism, in this context, is close to developing a more effective ability-for-use with speaking processes.

A second, although related point, concerns the issues of access and repertoire in second language speaking. Second language speaking is vulnerable to avoidance, as, indeed, speakers play to their strengths and to what they know, and avoid in engaging with areas they need to improve. Part of developing speaking skills therefore highlights the role of pedagogy in:

- Encouraging learners to use language which is only at the emergent stage, and which they may not be sure of – the cutting edge of language development
- Encouraging learners to use language which is in their repertoire, but which they tend to avoid using such as when a simpler form is available and they avoid engaging with a more complex (and perhaps developmentally advanced) form, e.g. with the use of ‘perhaps’ instead of ‘might’

In such cases also, the arguments from the chapter are relevant. Putting pressure on second language speakers will make it less likely that they take chances or avoid relying on simpler forms. More supportive speaking conditions can develop the speaking skill itself, and scaffold the use of emerging and lower repertoire language.

The third, and final, point, is perhaps slightly contradictory to this discussion. The arguments which have gone before might be interpreted as wrapping second language learners in cotton wool when they speak, and as a result, doing them a disservice. Encouraging second language speakers only to talk about familiar structured things, and that after they have had preparation time, maybe with a chance to repeat, only goes so far! And this, of course, is entirely true. So it is important to make clear that the careers of second language speakers should not be confined to the most favourable conditions. But it is equally important to say that speaking, as a skill, needs to be nurtured, and tasks and task conditions chosen to foster that skill. Part of pedagogy, in other words, is to make decisions about the tasks and conditions which are appropriate at any one point. Supporting, that is, when supporting is appropriate. Equally, it means increasing the pressure and extending the challenge when second language speakers are ready. The sort of research, and more pedagogic suggestions covered in this chapter, are intended to provide teachers with more ammunition when they make such pedagogic decisions.

Acknowledgements

The author would like to thank the Editors of this volume and an anonymous reviewer for comments which have strengthened this chapter considerably.

References

- Bui, H. Y. G. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 63–94). John Benjamins. <https://doi.org/10.1075/tblt.5.03gav>
- Bui, G., Skehan, P., & Wang, Z. (2018). Task condition effects on advanced level foreign language performance. In P. A. Malovrh & A. G. Benati (Eds.), *The handbook of advanced proficiency in second language acquisition* (pp. 219–237). John Wiley & Sons. <https://doi.org/10.1002/9781119261650.ch12>
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Longman.
- Bygate, M. (2018). *Learning language through task repetition*. John Benjamins. <https://doi.org/10.1075/tblt.11>
- De Bot, K. (1992). A bilingual production model: Levelt's "Speaking" model adapted. *Applied Linguistics*, 13, 1–24.

- DeKeyser, R. (2007). Introduction: Situating the concept of practice. In R. DeKeyser (Ed.), *Practice in a second language*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511667275.002>
- DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 83–104). Routledge.
<https://doi.org/10.4324/9780429503986-5>
- Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3–36). John Benjamins.
<https://doi.org/10.1075/llt.11.03ell>
- Ellis, R. (2009). The differential effects of three types of task planning on fluency, complexity, and accuracy in L2 oral performance. *Applied Linguistics*, 30, 474–509.
<https://doi.org/10.1093/applin/amp042>
- Ellis, R., & Yuan, F. (2005). The effect of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–192). John Benjamins. <https://doi.org/10.1075/llt.11.11ell>
- Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18(3), 299–324.
<https://doi.org/10.1017/S0272263100015047>
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3, 185–214.
<https://doi.org/10.1177/136216889900300303>
- Foster, P., & Skehan, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: A synthesis of the Ealing research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency investigating complexity, accuracy and fluency in SLA* (pp. 199–220). John Benjamins. <https://doi.org/10.1075/llt.32.09fos>
- Foster, P., & Skehan, P. (2013). The effects of post-task activities on the accuracy of language during task performance. *Canadian Modern Language Review*, 69, 249–273.
<https://doi.org/10.3138/cmlr.69.3.249>
- Foster, P., & Tavakoli, P. (2009). Lexical diversity and lexical selection: A comparison of native and non-native speaker performance. *Language Learning*, 59(4), 866–896.
<https://doi.org/10.1111/j.1467-9922.2009.00528.x>
- Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 63(2), 330–367.
<https://doi.org/10.1111/lang.12008>
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183–199.
<https://doi.org/10.1177/1362168810388711>
- Kobayashi, M. (2009). *Hitting the Mark: How can text organisation and response format affect reading test performance*. Peter Lang.
- Kormos, J. (1999). Monitoring and self-repair. *Language Learning*, 49, 303–342.
<https://doi.org/10.1111/0023-8333.00090>
- Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.
- Kormos, J. (2011). Speech production and the Cognition Hypothesis. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 39–60). John Benjamins. <https://doi.org/10.1075/tblt.2.06ch2>

- Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, 38, 1–30.
- Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge University Press.
- Li, Q. (2014). Get it right in the end: The effects of post-task transcribing on learners' oral performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 129–154). John Benjamins. <https://doi.org/10.1075/tblt.5.05qia>
- Long, M. H. (2007a). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. H. (2007b). Recasts in SLA: The story so far. In M. H. Long, *Problems in SLA* (pp. 43–74). Lawrence Erlbaum Associates.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Long, M. H. (2016). Experimental perspectives on classroom interaction. In N. Markee (Ed.), *Handbook of classroom discourse and interaction*. Wiley-Blackwell.
- Lynch, T. (2001). Seeing what they meant: Transcribing as a route to noticing. *English Language Teaching Journal*, 55, 124–132. <https://doi.org/10.1093/elt/55.2.124>
- Lynch, T. (2007). Learning from the transcripts of an oral communication task. *ELT Journal*, 61, 311–320. <https://doi.org/10.1093/elt/cm050>
- Malicka, A., & Sasayama, S. (April, 2017). The importance of learning from the accumulated knowledge: Findings from a research synthesis on task complexity. Paper presented at the 7th Biennial International Conference on Task-Based Language Teaching, Barcelona, Spain.
- Pang, F., & Skehan, P. (2014). Self-reported planning behaviour and second language performance in narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 95–128). John Benjamins. <https://doi.org/10.1075/tblt.5.04pan>
- Pang, F., & Skehan, P. (2021). Performance profiles on second language speaking tasks. *The Modern Language Journal*, 105(1), 371–390. <https://doi.org/10.1111/modl.12699>
- Robinson, P. (Ed.). (2011a). *Second language task complexity: Researching the cognition hypothesis of language learning and performance*. John Benjamins. <https://doi.org/10.1075/tblt.2>
- Robinson, P. (2011b). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3–38). John Benjamins. <https://doi.org/10.1075/tblt.2.05ch1>
- Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and directions in the development of TBLT* (pp. 87–122). John Benjamins. <https://doi.org/10.1075/tblt.8.04rob>
- Skehan, P. (2007). Task research and language teaching: Reciprocal relationships. In S. Fotos (Ed.), *Form-meaning relationships in language pedagogy: Essays in honour of Rod Ellis*. Oxford University Press.
- Skehan, P. (2013). Nurturing noticing. In J. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 169–180). National Foreign Language Center, Hawaii.
- Skehan, P. (2015). Limited attentional capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–155). John Benjamins. <https://doi.org/10.1075/tblt.8.05ske>
- Skehan, P. (2016). Tasks vs. conditions: Two perspectives on task research and its implications for pedagogy. In A. Mackey (Ed.), *Annual Review of Applied Linguistics*, 24, 34–49. <https://doi.org/10.1017/S0267190515000100>

- Skehan, P. (2018). *Second language task-based performance: Theory, research, and assessment*. Routledge. <https://doi.org/10.4324/9781315629766>
- Skehan, P. (2022). The psycholinguistics of task-based performance. In M. Long & M. Ahmadian (Eds.), *The Cambridge handbook of task-based learning and teaching* (pp. 3–26). Cambridge University Press. <https://doi.org/10.1017/9781108868327.002>
- Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research*, 1, 185–211. <https://doi.org/10.1177/136216889700100302>
- Skehan, P., & Foster, P. (1999). Task structure and processing conditions in narrative retellings. *Language Learning*, 49(1), 93–120. <https://doi.org/10.1111/1467-9922.00071>
- Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language*, pp. 193–216. John Benjamins. <https://doi.org/10.1075/llt.11.12ske>
- Skehan, P., & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & Vedder, I. (Eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching*. University of Brussels Press
- Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *International Review of Applied Linguistics*, 54(2), 97–112. <https://doi.org/10.1515/iral-2016-9992>
- Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 197–210). John Benjamins. <https://doi.org/10.1075/tblt.5.07ske>
- Skehan, P., & Shum, S. (2017). What influences performance? Personal style or the task being done? In L. L. C. Wong & K. Hyland (Eds.), *Faces of English education: Students, teachers, and pedagogy* (pp.28–43). Routledge. <https://doi.org/10.4324/9781315205618-3>
- Tavakoli, P. (2004). Oral narrative tasks and second language performance: An investigation of task characteristics (Unpublished doctoral dissertation). University of London.
- Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 58(2), 439–473. <https://doi.org/10.1111/j.1467-9922.2008.00446.x>
- Tavakoli, P., & Skehan, P. (2005). Planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–276). John Benjamins. <https://doi.org/10.1075/llt.11.15tav>
- Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). John Benjamins. <https://doi.org/10.1075/tblt.5.02wan>
- Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective: Influences on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 155–186). John Benjamins. <https://doi.org/10.1075/tblt.5.06wan>
- Wang, Z., Skehan, P., & Chen, G. (2019). The effects of hybrid on-line planning and L2 proficiency on video-based speaking task performance. *Journal of Instructed Second Language Acquisition*, 3, 53–80. <https://doi.org/10.1558/isla.37398>
- Winter, E. (1976). *Fundamentals of information structure*. Hatfield Polytechnic.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task and online planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1–27. <https://doi.org/10.1093/applin/24.1.1>

A task-based needs analysis framework for TBLT

Theory, purpose, and application

Ellen Grote and Rhonda Oliver

Curtin University

This chapter focuses on the evidenced-based Task Based Needs Analysis (NA) model that Michael Long developed as a tool for Task Based Language Teaching (TBLT). It describes the theoretical underpinnings of the model, its purpose and aims, its function in TBLT, how it contributes to task and syllabus design, and importantly, how the framework foregrounds the ‘real life’ aspects of language learning that is integral to the definition of task as proposed by Michael Long.

Introduction

It is widely recognized that a needs analysis (NA) is the starting point for designing a second language (L2) syllabus (Iizuka, 2019; Long, 2005c; Long & Norris, 2000; Van Avermaet & Gysen, 2006). Although NA research has been in practice for more than fifty years (Trim, 2007), the procedures and aims have changed over time to keep pace with advances in L2 acquisition research and approaches to teaching (e.g., Gass, 1997; Gass, Mackey, Alvarez-Torres, & Fernandez-Garcia, 1999; Gass, Mackey, & Ross-Feldman, 2011; Platt & Brooks, 2002; Robinson, 2011; Sato & Loewen, 2019). Contemporary understandings on L2 acquisition that have emerged from the research, particularly those subscribing to an interactionist perspective (Gass, 2003; Gass et al., 2011; Loewen & Sato, 2018; Oliver, Sato, Ballinger, & Pan, 2019) have guided the development of the task-based language teaching (TBLT) approach (Ellis, 2003; Long & Norris, 2000; Nunan, 2004). This has seen the shift in the emphasis to teaching L2 learners how *to do* things with language rather than how to combine (or synthesize) the components of language (Long, 2015, 2016). As such, syllabus design for TBLT has necessarily moved away from the traditional *synthetic* to an *analytic* syllabus, which foregrounds using language to make meaning in order to execute a task (Long, 1996, 2000, 2015).

Because the centrality of *task* is the central organizing principle of TBLT, identifying the language tasks that learners are expected to engage in constitutes the primary focus of a NA investigation. Long's contribution to NA research in the area of TBLT cannot be understated. His research (Long, 2005a) on TBNA methodology brought to the fore the importance of using evidence obtained through a systematic research process which aims to identify the language tasks that can be used to develop an appropriate and defensible syllabus. Instead of relying on the intuition of teachers, course designers and text book writers to design a syllabus, the TBNA as proposed by Long (2005a) draws on actual data. This NA model advocates using multiple sources and a variety of methods to collect and then triangulate data to ensure that the quality of the results are sound. The findings obtained using such an approach to NA can then be used to map out a TBLT syllabus that is both robustly evidence-based and, therefore, justifiable.

This chapter explores the background and theoretical underpinnings of the TBNA framework advocated by Long (2005a, 2015), its contribution to developing an analytical syllabus (as opposed to a synthetic syllabus), its function as the unit of analysis in a NA and the importance of using multiple methods and sources to triangulate the data. The aims and purpose of a NA are set out with examples from the research literature. The function of a NA is discussed in relation to how the process of conducting a NA contributes to syllabus design from identifying target tasks to organising pedagogic tasks into a coherent, defensible syllabus that will prepare language learners for use of the language in real world settings (Long, 2005a, 2015, 2016).

The 'real life' aspects of language teaching and learning that are integral to the definition of task are illustrated in more detail by describing some of the findings that emerged in the TBNA research that are unlikely to have been predicted by course designers or writers of commercially produced teaching and learning materials who rely on intuition. To provide examples of how a TBNA can bring to light the real-life language and cultural learning needs of students that had previously gone unnoticed, we describe TBNAs undertaken in various settings – including in programs for business, universities and, in particular, we provide a case study of a vocational education training (VET) high school for Australian Aboriginal students.

Background and theoretical underpinnings of the TBNA model

The emergence of TBNA research

Research undertaken in NA for L2 learners was initially formalized in the 1960s and 1970s when there was an influx of international students and workers coming to Britain. The universities needed to accommodate the students from abroad who needed to expand their English repertoire to meet the specialized needs of academia (Long, 2015). The creation of the European Economic Community and subsequent agreements made it easier for European citizens to seek employment across borders. To address the language needs of these workers, the Council of European Modern Language Project was launched and one of its aims was to establish a framework to identify the language required for various occupations (Long, 2005c; Trim, 2007). As L2 teaching programs sought to address the needs of those working in the various trades and professions, the focus shifted from the grammatical components of language to what learners would be required *to do* with the language. Nonetheless, NA endeavours continued to center on linguistic elements and what was thought to be the best ordering of these items in a learning program (Gilbert, 2005; Long, 2005c). As such, environmental and other variables were largely overlooked.

As L2 learning research advanced understandings on how an additional language is acquired, the process began to be viewed as one that progresses gradually over time in a ‘non-linear’ (Long & Norris, 2000, p. 866), non-categorical, developmental fashion. Importantly, the impact of both innate and external influences were also recognized to affect acquisition, highlighting the crucial role that input and the interaction between speakers served for enhancing language acquisition (Gass, 1997, 2003; Long, 1996). Pedagogically this was translated into the communicative approach to syllabus design which became known as Communicative Language Teaching (CLT) (Richards & Rodgers, 2014; e.g., Savignon, 1987) and this was founded upon an interactionist view of language learning.

Although CLT gave greater recognition to the importance of communicative interaction in language teaching, for all intents and purposes, the organisation of CLT syllabuses remained largely synthetic in character. A *synthetic* syllabus is underpinned mainly by the components of language such as the lexis, grammar, syntax and ‘notions and functions’ of the L2 (Long, 2015; Wilkins, 1976). Syllabuses based on these treat language items as discrete building blocks that learners need to synthesize, i.e., combine in a prescribed arrangement, according to certain criteria (Long, 2015), thus, prioritizing a ‘*focus on forms*’ (p. 21) over expressing meaning. The order in which language items are introduced to learners is mostly determined by the writer of the syllabus, or textbook, rather than when the learners are developmentally ready to acquire them. Syllabuses such as these assume that learners

can manage the cognitively demanding process of combining these linguistic components to convey their intended meaning. Furthermore, and as noted previously, designing synthetic syllabuses rely solely on the intuition of textbook writers, course designers or teachers (Long, 2005a).

To address some of these shortcomings, various analytic syllabuses were developed to bring to the fore the core aim of learning while using an L2 to make meaning. A number of methods designed to promote meaning making include the Process (Breen & Candlin, 1980) and Procedural (Prabhu, 1987) approaches and TBLT (Ellis, 2017; Long & Crookes, 1992). A shared feature of these is that tasks are the key pedagogic tool implemented in practice. The notion of *task* varies in the TBLT literature (Ellis, 2017; Long, 2016). In Long's view, *task* is used in 'the non-technical, everyday, real-world use of the term' (2015p, p. 108), that is

a piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes... In other words, by "task" he meant the hundred and one things people do in everyday life, at work, at play, and in between. Tasks are things they will tell you they do if you ask them and they are not applied linguists.

(Long, 1985, cited by Long, 2015, p. 108)

Importantly with respect to TBNA, tasks

are the real-world communicative uses to which learners will put the L2 beyond the classroom – the things they *do* in and through the L2 – and the task syllabus stands alone, not as one strand in a hybrid of some kind'. (Long, 2016, p. 6)

See Long, 2016 and Ellis, 2017 for more detailed descriptions of the different ways that the notion of task has been used in the TBLT literature.

With respect to the TBNA model proposed by Long (2015, 2005c), which is the focus of this paper, communicative tasks that occur in real world settings are perceived to be the cornerstone of TBLT and the unit of analysis for all aspects of TBLT from conducting a TBNA, to designing a TBLT syllabus and implementing TBLT practices in the classroom.

Task as the unit of analysis in NA research

In the TBNA model that Long (2005a, 2015) proposes, he argues that the unit of analysis should be the task because of its capacity to offer a range of practical and theoretical advantages. Perhaps the most utilitarian benefit is that it provides continuity across every stage of TBNA and TBLT, from planning and executing the NA processes, using the data to inform and organize syllabus design, including the development of TBLT activities and assessments.

Using task as a unit of analysis also makes the data collection process more tangible, mutually comprehensible and accurate when talking with ‘domain experts’ (Long, 2005a, p. 22), i.e., those individuals who actively work in a particular profession or trade. (Domain experts as data collection sources are discussed in more detail below.) Because the notion of task is a commonly used term that is neither technical nor linguistic, it is easily understood by most people regardless of age, experience or professional background. Communicating effectively with domain experts about the tasks undertaken in their workplace is critical because they have direct knowledge of the specific routine target tasks (and sub-tasks) performed on their work site. Gilabert (2005) distinguishes between target tasks and target sub-tasks, defining a target task as ‘a differentiated process domain experts have to carry out ... which is divided into steps, each of which must have an outcome, and not be dependent on or part of other tasks’ (p. 184). He goes on to describe a target sub-task as ‘a differentiated process which, while having a number of steps and an outcome, is dependent on or part of another major target task’ (see also p. 184; Long, 2015). In his study of Spanish journalism students learning English for their jobs, the distinction is important because, as Gilbert notes, while some sub-tasks are undertaken in the L2, others were performed in Spanish.

As Gilabert’s research demonstrates the importance of accessing details about the nature of the different interactions required and the settings in which the tasks take place, the interlocutors or types of written texts the learners engage with provides further justification for using task as a unit of analysis. Obtaining such information is essential for designing the syllabus with regard to selecting appropriate learning materials, activities and assessments to ensure that they adequately prepare learners for engaging in real workplace encounters.

Long (2005a) maintains that using task as a unit of analysis accrues additional support when considering language use at the level of discourse. He points out that while analyses of texts at the linguistic or grammatical level centers most of the attention on the results of a communicative interaction between interlocutors, focusing on the task as a dynamic process is by nature more authentic and motivating for the learner.

Importantly, as noted previously, the outcome of the TBNA – the analytic syllabus and its associated learning and assessment activities – aligns closely with contemporary understandings of L2 acquisition processes (Long, 2005a, 2015; Long & Crookes, 1992; Serafini & Torres, 2015). The content and associated interactive learning activities of the TBLT analytic syllabus are necessarily learner-centred (Long, 2005a, 2015; Nunan, 2013) and, therefore, recognized to promote language acquisition inside the classroom just as it is in real world contexts.

The multi-method multi-source approach to TBNA methodology

Advocates of TBNA highlight the advantages of the systematic use of multiple sources, a variety of methods and triangulation in the data collection and analytical processes (Cowling, 2007; Gilabert, 2005; Jasso-Aguilar, 1999, 2005; Long, 2005a; Oliver & Grote, 2016; Oliver, Grote, Rochecouste, & Exell, 2013a, 2013b). However, prior to Long's (2005a) investigation examining the methodologies that could be applied to NA, the information and published research on NA methodology generally and, in relation to TBNA specifically, was somewhat limited. Therefore, in Long's (2005a) research on the tasks performed by flight attendants, the aim was to determine the value of possible sources and various methods for collecting such data. With respect to sources, of particular interest were the quantity and quality of data obtained from 'insiders' versus 'outsiders' (p. 48); in terms of data collection methods, it was the qualitative and quantitative data that was scrutinized. Additionally, the impact of different combinations of methods and sources, as well as the order of the procedures undertaken, were examined and evaluated.

In this study Long assessed the value of the data obtained, including from: documents – such as manuals used in training, particularly those issued by airlines that outlined work procedures; insider informants – comprising flight attendants of different ages and levels of experience; outsider participants – students of Applied Linguistics who were regular airline clientele. A range of methods were implemented to gather the data, including: written introspective responses to prompts, provided by (insider) flight attendants and (outsider) student air travellers; follow-up unstructured interviews with both groups; audio-recordings of flight attendants' speech as they communicated with clients and colleagues; and non-participant observations of flight attendants as they communicated on the job.

The findings indicated that the written documents were useful for providing a list of tasks the flight attendants engaged in. These were complemented by the data gathered from other sources and methods which gave context and more details about the language and the types of discourse they would use. On the other hand, the unstructured interviews served to identify topics that could be explored later in further depth, either through structured interviews or questionnaires. Hence, the sequencing of these two data collection methods, that is, the unstructured followed by structured interviews or questionnaires, highlighted the importance of ordering such procedures strategically. In addition, of particular value were the recordings of the less visible communication occurring between colleagues and out-of-hearing of out-of-view clientele, which provided authentic samples of on-the-job discourse. This type of recorded data had the potential to be useful in providing realistic samples of the type of target language that needed to be learned. Importantly, it was the perspectives of the domain expert flight attendants that proved to be crucial

because many important communicative tasks occurred out of the hearing and beyond the purview of outsiders.

The study underscored the importance of using different configurations of multiple sources and methods to generate a useful corpus of data, which can then be clarified, confirmed, augmented and expanded. Moreover, access to data drawn from multiple sources and methods facilitates triangulation, to ensure that sufficient quality data are collected. In turn, the accuracy of interpreting the data and the validity of the findings can be strengthened (Long, 2005a).

Purpose and aims of NA

The main purpose of conducting a TBNA is to facilitate the development of a TBLT syllabus, one that will effectively prepare learners to do what they will need to do with language as they navigate real world settings and situations beyond the classroom. Because every group of learners will use the L2 for different purposes, Long (2005b, 2015) maintains that a syllabus developed without considering future real world contexts is unlikely to prepare them adequately. Furthermore, the need for accuracy when identifying tasks varies along a spectrum, depending on the aims of the course and the learners. For example, the nature of the language tasks that young children are likely to use may be more generalized than those encountered by adults who will communicate in the L2 for specific occupational purposes.

Identifying the nature of situational contexts in which the learners will engage varies in relation to the types of discourse, genres, registers, potential interlocutors and desired outcomes of the exchange. Nonetheless, Long (2005a, 2015) argues that a NA should be conducted to obtain sufficient data which is necessary to develop a robust syllabus, one that will prepare the learners for the real world situations they are likely to encounter. A course guided by a syllabus of this type can be said to be characterised as a Language for Specific Purposes program (see also Serafini & Torres, 2015).

While it might seem obvious that language teaching programs should address the language needs of the learners, this does not always translate into practice. For example, many international students arrive at university having spent a considerable amount of time in general language courses. Students enrolled in such programs have often learned content for a range of contexts and situations, some of which they are unlikely to encounter. At the same time they undoubtedly could have missed out on the types of language requisite for their chosen fields of study or work (Long, 2005a). Generic (English) language programs that fail to address the specific needs of learners have been disparagingly referred to as TENOR courses, i.e., those Teaching English for No Obvious Reason (Brown, 2016, after Abbott, 1981). They

are said to be common, for example, in Japanese public schools and universities (Lambert, 2010) where the view is that students' future employers will provide the necessary 'fit-for-purpose' language training. While such general courses enable students to qualify for better paying jobs, the lack of a clear idea of what students will be able to do with the language can reduce motivation and fail to adequately prepare them for any particular workplace context (Lambert, 2010).

Although courses in English for Specific Purposes (ESP) have been marketed for some time (often by private language schools to those already in the workplace), university language departments have begun to recalibrate their English language programs to accommodate the increasing demands of international students for English language training that will better prepare them for their future professional lives. Foreign language departments in universities in the United States (US), for instance, are also now responding to growing pressure to adapt their courses to prepare domestic student cohorts to meet the needs of contemporary real world multicultural workplace settings. For example, Spanish for Specific Purposes (SSP) programs which are tailored to enable learners to communicate effectively in the healthcare and business sectors are being developed and offered to cater to the needs of Hispanic communities (Serafini & Torres, 2015).

While TBLT courses should vary according to the learner group and the specific purposes for which the L2 will be used, ultimately it can be said that the overarching aim of any TBNA is to identify the *target tasks* that will inform the design of a TBLT syllabus (Lambert, 2010; Long, 2000, 2015; Nunan, 2004). In order to achieve this, as the TBLT syllabus is designed in more detail, the target tasks are used as models for developing *pedagogic tasks*. These provide the main vehicle to facilitate language learning through interactive meaning making. Specifically, a pedagogic task has been defined as:

a piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form. The task has a sense of completeness, being able to stand alone as a communicative act in its own right with a beginning, middle and an end. (Nunan, 2004, p. 4)

Pedagogic tasks form the nuclei of the language teaching and learning activities, associated materials and assessments that align with a TBLT approach (Klee, 2015; Ogilvie & Dunn, 2010; Oliver, 2020, 2021; Serafini & Torres, 2015). With the aim of facilitating learner proficiency to prepare them to participate effectively and with confidence in real world encounters, pedagogic tasks must be designed to be purposeful, feel authentic and at the same time include a focus on language (Oliver, 2020).

In the sections that follow, some of the aims and purposes that a TBNA can serve are exemplified, including: a study undertaken to develop a TBLT syllabus for a new business English course for English as an Additional Language (EAL) learners enrolled in a bridging course designed to prepare them for entry into a university; several projects designed to identify the specific language needs of L2 learners in particular occupational fields; a long term ethnographic research project conducted to identify and address the language and literacy needs of Aboriginal students in a vocational education training (VET) high school in Australia; and, a study which sought to guide material writers for adult learners of business English.

Developing a new TBLT syllabus

One large scale study (Huh, 2006) aimed to develop a new course on business communication course syllabus to complement a suite of intensive content-based courses within an English language bridging program designed to prepare EAL learners to enter universities in the US. The students in the program needed the requisite English language skills to conduct business either in the US or in their home countries. Although the bridging course was run on the campus at a US university, data were collected from 75 Korean business professionals – domain experts – working in 13 different Korean corporations. The study sought to identify a range of target tasks and then categorise them to inform the business communication syllabus.

Renovating existing syllabuses

Several TBNA research projects that have sought to renovate existing language syllabuses to address the workplace needs of learners have narrowly focused on particular job roles, professions or industries. Within this corpus is a study carried out in the hospitality sector: Jasso-Aguilar's (1999, 2005) study explored the English language tasks required of maids employed by a Waikiki hotel who spoke EAL. Gilabert's (2005) TBNA study sought to renovate an existing EAL program with the aim of identifying the English language and literacy tasks that Spanish speaking students enrolled in a Communications Studies program in Catalonia, Spain would need to undertake in their role as journalists. Cowling's (2007) research examined the English language learning tasks that newly inducted employees in a large Japanese industrial company would be expected to perform in their job roles. The TBNA would guide the development of a syllabus for an intensive three-day English program administered once a year for the first three years of the trainees' employment and therefore needed to cater to a wider range of potential job roles, which for some learners had yet to be determined.

Another TBNA study (Lambert, 2010) took place at a Japanese municipal university for a business English course for students whose post-education ambitions remained in the embryonic stage. For some, their future job roles depended very much on which companies offered them employment. As is common in English departments in Japanese universities, the existing course content was overly generalized, which, as Lambert indicates, not only inadequately prepares learners for using English in the workplace, but can reduce their motivation to learn the language. Hence, the purpose of this research was to design a syllabus that could prepare students more effectively for their post-university careers.

Designing a business Spanish course for advanced students was the aim of research undertaken at two public universities in a northeastern state of the US (Serafini & Torres, 2015). The NA sought to refine the course goals, develop a more appropriate syllabus and aid the selection of learning activities. Although the course was taught by instructors who had extensive knowledge and experience in the area of L2 pedagogy, including TBLT methodology, they had limited familiarity with the business domain.

The purpose of another study (Iizuka, 2019) was to refine the curriculum of a general Japanese language program at a US university to better serve the needs of its learners who had various career goals and who intended to enrol in a summer study-abroad program in Japan. The program involved living with a host family for one month, and then shifting to an apartment while studying in a Japanese language course. As such, it was determined that the NA would focus on the everyday communication tasks that the learners were likely to encounter during their stay (NB: More details about this TBNA are described below).

Integrating TBLT practices into existing vocational skills training programs

A long term institution-wide ethnographic TBNA study took place at a residential VET high school for Australian Aboriginal students (Oliver, 2020; 2021; Oliver & Grote, 2016; Oliver et al., 2013a, 2013b). The school was located outside of small country town in Western Australia on a large tract of land that enabled it to offer a range of VET courses, including Stock and Station (Rural Operations), Outdoor Recreation, Hospitality, Tourism, General Construction, Automobile Mechanics, Metal and Engineering, as well as Business and Administration. The students came from rural and remote communities where Indigenous languages, an English-based creole (Kimberley Kriol) and/or a non-standard dialect of Australian English (Aboriginal English) were spoken. Prior to undertaking the TBNA, a preliminary study (Oliver & Grote, 2010) revealed that the students had difficulty understanding their non-Aboriginal Australian English speaking teachers and the employers and co-workers in their Structured Workplace Learning programs. In addition,

the learners were unfamiliar and uncomfortable with non-Aboriginal cultural discourse practices. It was determined that the overarching aim of the TBNA would be to identify the standard Australian English (SAE) language, literacy and communication practices that trainees needed to learn to increase their chances of succeeding in their current VET programs, the registered training organisation (RTO) institutions in their home communities that they might enrol in or the future places of work. It also sought to gain an understanding of the kinds of training that the teaching staff at the high school might benefit from to enable them to modify their current teaching practices to facilitate the learners' acquisition of SAE as an additional language or dialect.

Guiding material writers for business English adult learners

Evans's (2013) small-scale NA study was part of a larger long-term research project which examined how the transition of Hong Kong's status as a British colony to a special administrative region of China has affected decisions regarding language use in Hong Kong businesses (Evans, 2010). Although not specifically referred to as TBNA research, it explored the English language and literacy tasks undertaken by senior level employees working in the service industries, mainly, 'financial services, trading and logistics, tourism, and professional services' (Evans, 2013, p. 283). The study aimed to provide writers of general Business English materials with the kinds of English language tasks that professionals regularly undertake so that instructors could use this information to develop a 'simulation-based approach' (p. 291) for their teaching practices.

How TBNA contributes to syllabus and pedagogic task design

A strategically planned TBNA can contribute to developing a coherent, defensible TBLT syllabus that is relevant to the language and literacy needs of L2 learners. It can provide evidence to guide the decision-making processes of the syllabus's design, including:

- identifying the course objectives to structure the syllabus;
- ascertaining the target tasks, and then categorising (task types) and sequencing them;
- converting target tasks into pedagogic tasks;
- choosing or expanding on teaching/learning materials to support pedagogic tasks and develop task-based assessments; and,
- integrating the essential cultural matters that came to light in the TBNA so they can be taught explicitly as a part of the TBLT syllabus.

In Long's (2015) work, he provides detailed discussions on issues relating to the selection of tasks, transforming them into pedagogic tasks and sequencing or 'grading' (p. 227) them with respect to various criteria. For the purposes of this chapter, we briefly summarize his conclusions by saying that, in Long's view for a TBLT syllabus to be successful, the selection of tasks should be guided by the findings of the TBNA. He maintains that the ordering of tasks should be undertaken according to various criteria around task complexity, including 'rational sequencing (grading, or gradation)' (p. 246), but also notes that it is an area of ongoing research in TBLT. In this regard, a guiding general principle is that:

pedagogic tasks are simpler, sometimes much simpler, versions of the target task-type or of one or more to its components. They gradually increase in complexity – *the complexity of the pedagogic tasks, not linguistic complexity...* – until they reach the full complexity level of the target task(s) that motivated their inclusion. (Long, 2015, p. 225)

We revisit three of the studies mentioned previously which illustrate some of the ways in which a TBNA contributes to program design. These include: the business Spanish language course for university students (Serafini & Torres, 2015); the Japanese language course aimed to prepare university students for a study-abroad program (Iizuka, 2019); and, the VET skills programs in which existing courses would be modified to include TBLT strategies in order to address the language and literacy learning needs of Aboriginal high school students learning SAE as an additional language or dialect.

The business Spanish language course

Serafini and Torres's (2015) research undertaken to develop a business Spanish course at two American universities illustrates how a TBNA can help determine the course objectives and the selecting, categorising and sequencing of target tasks in the development of a TBLT syllabus for advanced level students. Data were collected from both insider and outsider informants using different methods. The insider group comprised professionals and academics in the business realm while the outsider informants included students majoring in business, having had limited or no experience working in the business world.

The participants recruited as experts for the project were invited to participate in an online survey that asked them to list typical tasks employees would be likely to undertake while at work. Their responses were used to develop a questionnaire made up of 40 communicative tasks. The non-expert group, i.e., the university business majors, were asked to respond to the questionnaire by indicating the level of difficulty they anticipated for undertaking each of the 40 tasks. They were also

asked to estimate the frequency in which they thought they might have to engage in those real life tasks.

Based on the analysis of these data and taking into consideration the difficulty, frequency, modality (reading, writing, listening or speaking) and the appropriateness of the tasks for advanced level learners, a set of criteria was established to narrow the 40 tasks down to a list of 14 'core target tasks' (Serafini & Torres, 2015, p. 455). The course objectives used these criteria to categorise the tasks into each of five target task-types, which constituted the task-based overarching goals of the syllabus (see also Long, 2015). The core target tasks then could be clustered with related task types. Next the target tasks were converted into pedagogic tasks for teaching, learning and assessment activities and were sequenced for the TBLT course syllabus in terms of difficulty. The analytic processes involved in designing the syllabus enabled the researchers to align the core target tasks and target tasks with the American Council on the Teaching of Foreign Languages World Readiness standards in order to meet university criteria (See Serafini & Torres, 2015 for more details of this process).

The Japanese language course for study-abroad students

As mentioned above, Iizuku's (2019) TBNA study sought to revamp a general Japanese language course from one that was mainly grammar oriented to a task-based curriculum in order to serve the needs of American students intending to study at a Japanese university for the summer term. It provides an example of how a TBNA was used to identify and categorise target tasks, to develop pedagogic tasks and to augment the teaching and learning materials to meet the specific needs of learners preparing for living and studying in Japan.

Data were gathered in a two-stage process including exploratory semi-structured interviews with students who had completed the program, as well as families who had experience hosting the students. The research sought to identify the communication tasks that the students had to engage in on a daily basis while living with their host families, potential problems that might arise and the learning objectives with respect to the communicative interactions that needed to be addressed in the course. In addition, topical themes that arose in the interviews were also noted.

The findings from the interviews informed the development of questionnaires for both students and host family members for the second stage of data collection. The analysis of the interviews enabled three categories to be established as organising themes for the questionnaires and the nature of the question: (1) target tasks (importance and problems); (2) homestay matters (occurrence of interactions, issues arising and other topics of a general nature); and (3) the Japanese language course itself (students' goals and views about the classes).

The findings indicated a need for the syllabus to include more opportunities for oral language practice than was provided by the current textbook and other course materials. This included, for example, pedagogic tasks to facilitate ‘small talk’, talking about sensitive topics and asking for (and comprehending) directions to navigate the city. In addition, it was identified that there was a need for more specialised experience in reading labels, particularly on food packages. Iizuka pointed out this would require surveying class members to determine the specific food allergies or preferences of individuals. Based on the target tasks identified, the author suggested potential pedagogic tasks that could offer opportunities for students to learn and practice these target tasks.

The results also highlighted the value of cultural insiders, i.e., host family members, who pointed out a wide range of cultural issues that the students were unaware of, but needed to understand. These included the importance of including activities in the syllabus that would enhance cultural awareness to communicate politeness as well as matters pertaining to being a respectful, courteous guest in the host family’s home. With regard to language issues, the findings indicated a need for assessment criteria to take into consideration the importance of signalling politeness when communicating in Japanese. With respect to the issue of living with a host family, the course designers were alerted to the need to integrate into the syllabus the importance of keeping one’s room orderly, maintaining one’s personal hygiene and not wasting food or electricity.

How the TBNA foregrounds the real life aspects of language learning: Case study

The findings of TBNAs can bring to light the kinds of ‘real world’ conditions and aspects of the communicative interactions that learners encounter in the workplace (eg, Cowling, 2007; Evans, 2013; Gilabert, 2005; Jasso-Aguilar, 2005; Long, 2005a; Oliver, Grote, Rochecouste, & Exell, 2012) and day-to-day situations (eg, Iizuka, 2019) in which they will need to use the L2. Many of these real-world exchanges would not have come to light if the syllabus writers or teachers had simply relied on their intuition instead of applying TBNA methodology involving multiple sources and methods in the data collection process (Long, 2005a, 2015).

In the remainder of this section the real-life language, literacy and cultural issues that emerged in the TBNA undertaken at the VET high school for Australian Aboriginal students are discussed in more detail – describing how they came to light and then how they were subsequently addressed in TBLT practices.

VET high school programs for Australian Aboriginal students

Data for the TBNA research undertaken at the VET high school for Aboriginal students were collected over an extended period of time in order for the researchers to develop a relationships of trust with the Aboriginal students and the school staff (Oliver, 2020; 2021; Oliver & Grote, 2016; Oliver et al., 2012, 2013a, 2013b). A range of ethnographic methods were implemented in the study, including non-participant observation, informal unstructured interviews, photographic records of tasks and document collection. It drew on multiple sources, including teachers, students, community members, teachers at the RTOs in the students' home communities and employers participating in the school's Structured Workplace Learning program. The teachers who delivered the VET subjects were expert insiders in their respective trades, however, some had no formal teaching qualifications. Most teachers also had limited understandings about additional language or dialect acquisition, so were 'outsiders' in relation to these aspects. The Aboriginal students and community members were outsiders with regard to the skilled trades being taught, but were insiders with respect to Aboriginal linguistic and cultural issues, the learners' life experiences, their goals, expectations and desires for the future.

The findings of the study indicated a range of linguistic, literacy and cultural issues that needed to be addressed across the various training courses to help increase the chances of the students succeeding in the program and beyond. These included applying L2 strategies to teach the SAE language and literacy skills required in the workplace, such as terms for tools, procedures, giving and confirming instructions, asking for clarification and reading safety labels. In addition, students needed to develop their confidence and wanted explicit instruction on and practice in how to socialise in the workplace. For example, they were unfamiliar or uncomfortable with non-Aboriginal Australian English conventions and expressions for greeting and leave taking at their Structured Workplace Learning program sites. While it is common and not at all impolite in their communities to use gestures or say nothing at all in these situations, such practices can be misinterpreted in less positive ways by non-Aboriginal bosses and co-workers. Employers also expressed the need for the trainees to learn to engage in small talk at work, exchange personal information and stories and, as one employer noted, being able to "use humor appropriately" and occasionally being "willing to take the mickey", (Oliver et al., 2012, p. 234), that is, willing to be teased, to develop workplace relationships.

When interviewed, several VET teachers pointed out the need for students to develop "soft skills" with regard to displaying non-Aboriginal conventions of "work ethic", "time management" and "showing enthusiasm" (Oliver et al., 2013b, p. 254) on the job. One teacher noted that "we can teach them the skills they need for the job, we can't teach them to be confident to talk to us" (Oliver,

2020, p. 151, italics removed). When students, Aboriginal community members and RTO lecturers were interviewed, the notion of *shame* emerged frequently as something that the students needed to gain mastery over. *Shame* is an Aboriginal English word that extends beyond the way Australian English speakers use it. It refers to an emotional response similar to embarrassment or shyness, usually to a situation that non-Aboriginal people normally would not find confronting, such as when a student is singled out from the group for any reason, negative or positive (Malcolm & Grote, 2007). Shame is experienced in situations when there is a power imbalance, such as when an Aboriginal youth is required to talk to a non-Aboriginal teacher, police officer, medical practitioner or employer (Grote & Rochecouste, 2012). As one student pointed out, being “the only blackfella in that [work] environment” can cause shame (Oliver et al., 2013a, p. 45). Another noted that “communicating with [non-Aboriginal] people is the hardest” (p. 45) and a third reported that speaking on the telephone to an unfamiliar non-Aboriginal person could be “frightening” (p. 45). As another student explained, “Most of them [students who get shame] don’t know any...white kids” [or adults] and need to “be around more [non-Aboriginal] people they don’t know” (p. 45) to get used to it and overcome the shame response.

Later in the study, when the research team were invited by the school to deliver professional development (PD) to the teachers, the concept of shame and how to ameliorate it using tasks was discussed. Informed by the findings of the TBNA, the PDs sought to enhance teacher awareness of the students’ language learning needs and to help develop the skills necessary for them to design linguistically and culturally appropriate pedagogic tasks, ones that were genuinely relevant to their real-life communicative needs, such as those relating to their future occupations (Oliver, 2020, 2021).

An example of an authentic real-life work related pedagogic task was developed by an automobile mechanics teacher following his participation in the PD program (Oliver, 2020). The teacher first implemented a pre-task activity focusing on tools and procedures for changing a spark plug to ensure the students were familiar with and able to use the Australian English terms for the task. Having already labeled the tools, the teacher began a discussion which required students to use the names of the tools to explain how they could be used. The trainer then introduced the pedagogic task in which a more experienced, confident trainee gave step-by-step instructions in SAE to a new student as they worked together to replace the spark plug. This entailed locating and reading the information label affixed to the machine that indicated the type of spark plug required and then accessing the appropriate one from the storage cabinet. The students then had to find the machine’s manual and decipher the instructions on how to replace the spark plug. Thus, the learners collaborated as they undertook the real-life task of changing a spark plug. In order

to complete the task they were motivated to negotiate meaning using their oral and literate SAE skills as they learned and reinforced the names of the tools, parts of the machine and the procedures while executing the task (Oliver, 2020).

Another real-life task was identified in the NA by stakeholders in the community and by employers was that of obtaining a driving license and, importantly, understanding the responsibilities associated with it (Oliver, 2021). A number of staff members worked on different aspects of enabling students to acquire a driving license. Being able to drive legally is an important practical asset for those living in remote communities – not only for work-related purposes – but also for daily life. One of the teachers who has taken on a major role in the endeavour pointed out the central importance of learning SAE oral and written skills for obtaining a learner's permit, as well as for the final written and practical assessment processes. Using only gestures to ask directions, as is common in remote communities, can have dangerous consequences while driving a vehicle. It was noted that while one student was practicing driving skills, he took his hands off the steering wheel to ask directions using only gestures, an action that would have caused him to fail the practical component of the driving test (Oliver, 2021).

To help the students develop their SAE skills in this regard, the teacher developed a number of pedagogic tasks involving toy cars and a road map. Students worked in small groups taking turns to practice the language required for giving directions as a navigator and then following verbal instructions as the driver. When their confidence increased, they progressed to taking turns testing their abilities in pairs with uni-directional information gap activities. The teacher designed other tasks to provide opportunities for students to acquire the SAE skills associated with taking the learner's permit exam, which requires a higher level of literacy than most learners could manage. For example, quizzes were developed requiring them to collaboratively figure out the road rule that applied in different circumstances while driving, again using toy cars and road maps to help them visualise the situations. Simulating scenarios that would appear on the written component of the learner's permit test provided opportunities for the students to have kinetic-like rule learning experiences to process the road rules while collaboratively working on the oral and literate skills they would need for the test.

The teacher also wanted the students to understand the real-life repercussions of road accidents, which occur frequently in remote communities due to disregarding laws about seatbelts, wearing helmets on motorcycles, alcohol and drug use, and the toll such practices take on those involved, their families and friends. As a prompt for discussions, the teacher brought in newspaper articles about recent road accidents for students to discuss in groups, assigning roles to group members such as leader, secretary and a person to summarise their discussion to the class. This was followed by activities in which students designed posters, wrote poems or

recorded their personal reflections in journals, all of which were then shared with the class (Oliver, 2021).

To address the need for students to learn to overcome shame when talking to non-Aboriginal people and the desire they expressed to do so, the school took an institution-wide approach (Oliver, 2020) by organising opportunities throughout the year for students to meet and interact with such people. To prepare students for these encounters, teachers developed scenarios for self-assured students with prior cross-cultural experience to model SAE language they might use in casual communicative interactions with a non-Aboriginal person. Students then had the opportunity to practice these and similar simulated interactions in pre-task and task pair work activities to prepare them for an upcoming annual event. The annual National Aboriginal and Islander Day of Celebrations (or NAIDOC) provides an opportunity for the school to invite members of the public to visit their school campus (Adams & Oliver, 2019). The more experienced students were involved in activities such as traditional storytelling and guiding children's craft activities while others worked in small teams as they showed visitors around the farm and offered beverages at the refreshment table. The process of preparing for and engaging in activities such as this provided an opportunity for students to engage in real-life communicative tasks that are relevant to the kinds of cross-cultural communicative situations they are likely to encounter in real life (Oliver, 2020).

Conclusion

In this chapter we have attempted to provide an overview of the TBNA model proposed by Long (2015, 2005c) with respect to its usefulness as a tool to primarily identify the language learning needs of L2 learners. The overarching aim of the framework is to develop a TBLT syllabus, which is analytic (in contrast to a synthetic syllabus), one that enables learners to use the L2 to make meaning in order to get things done, that is, complete tasks. As such, the TBNA underpins a TBLT approach that is consistent with contemporary interactionist perspectives of L2 acquisition.

To undertake a TBNA it is important to use a variety of methods and multiple sources to obtain a large and triangulated data set. As we demonstrate, by doing this a TBNA framework can be adapted to suit a variety of contexts and purposes in order to develop a sound, evidence-based syllabus (or to renovate an existing one), and to integrate TBLT activities into courses. Using a TBNA can also provide guidance to teachers and material writers in terms of identifying course objectives (task-types), target tasks, and in the development of pedagogic tasks. Finally, and most crucially, a TBNA can illuminate the real-life communicative

tasks, and importantly, provide essential contextualised details, including potential cross-cultural impediments, so that they can be taken into consideration when a TBLT approach is used.

References

- Adams, R., & Oliver, R. (2019). *Peer interactions in classrooms*. Routledge.
- Breen, M., & Candlin, C. N. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics*, 1, 89–112. <https://doi.org/10.1093/applin/1.2.89>
- Brown, J. D. (2016). *Introducing needs analysis and English for Specific Purposes*. Routledge. <https://doi.org/10.4324/9781315671390>
- Cowling, J. D. (2007). Needs analysis: Planning a syllabus for a series of intensive workplace courses at a leading Japanese company. *English for Special Purposes*, 26, 426–442. <https://doi.org/10.1016/j.esp.2006.10.003>
- Ellis, R. (2003). *Task-based language learning and teaching*. OUP.
- Ellis, R. (2017). Task-based language teaching. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 108–125). Routledge. <https://doi.org/10.4324/9781315676968-7>
- Evans, S. (2010). Business as usual: The use of English in the professional world in Hong Kong. *English for Special Purposes*, 29, 153–167. <https://doi.org/10.1016/j.esp.2009.11.005>
- Evans, S. (2013). Designing tasks for the business English classroom. *ELT Journal*, 67(3), 281–293. <https://doi.org/10.1093/elt/cct013>
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Lawrence Erlbaum Associates.
- Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Blackwell. <https://doi.org/10.1002/9780470756492.ch9>
- Gass, S. M., Mackey, A., Alvarez-Torres, M. J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549–581. <https://doi.org/10.1111/0023-8333.00102>
- Gass, S. M., Mackey, A., & Ross-Feldman, L. (2011). Task-based interactions in classroom and laboratory settings. In P. Robinson (Ed.), *Task-based language learning* (Language Learning 61: Supplement, pp. 189–216). Wiley-Blackwell. <https://doi.org/10.1111/j.1467-9922.2011.00646.x>
- Gilbert, R. (2005). Evaluating the use of multiple sources and methods in needs analysis: A case study of journalists in the Autonomous Community of Catalonia (Spain). In M. H. Long (Ed.), *Second language needs analysis* (pp. 182–199). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299.007>
- Grote, E., & Rochecouste, J. (2012). Language and the classroom setting. In Q. Beresford & G. Partington (Eds.), *Reform and resistance in aboriginal education*. UWA Press.
- Huh, S. (2006). A task-based needs analysis for a business English course. *Second Language Studies*, 24(2), 1–64.
- Iizuka, T. (2019). Task-based needs analysis: Identifying communicative needs for study abroad students in Japan. *System* 80, 134–142. <https://doi.org/10.1016/j.system.2018.11.005>

- Jasso-Aguilar, R. (1999). Sources, methods and triangulation in needs analysis: A critical perspective in a case study of Waikiki hotel maids. *English for Special Purposes*, 18(1), 27–46. [https://doi.org/10.1016/S0889-4906\(97\)00048-3](https://doi.org/10.1016/S0889-4906(97)00048-3)
- Jasso-Aguilar, R. (2005). Sources, methods and triangulation in needs analysis: A critical perspective in a case study of Waikiki hotel maids. In M. H. Long (Ed.), *Second language needs analysis* (pp. 127–158). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299.005>
- Klee, C. A. (2015). Content-based programs and Spanish for the professions. In M. Lacorte (Ed.), *The Routledge handbook of hispanic applied linguistics* (pp. 184–201). Routledge.
- Lambert, C. (2010). A task-based needs analysis: Putting principles into practice. *Language Teaching Research*, 14(1), 99–112. <https://doi.org/10.1177/1362168809346520>
- Loewen, S., & Sato, M. (2018). State-of-the-art article: Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329. <https://doi.org/10.1017/S0261444818000125>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Long, M. H. (2000). Focus on form in task-based language teaching. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy* (pp. 191–196). John Benjamins. <https://doi.org/10.1075/z.96.11lon>
- Long, M. H. (2005a). Methodological issues in learner needs analysis. In M. H. Long (Ed.), *Second language needs analysis* (pp. 19–76). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299.002>
- Long, M. H. (2005b). Overview: A rationale for needs analysis and needs analysis research. In M. H. Long (Ed.), *Second language needs analysis* (pp. 1–16). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299.001>
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Long, M. H. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, 36, 5–33. <https://doi.org/10.1017/S0267190515000057>
- Long, M. H. (Ed.). (2005c). *Second language needs analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667299>
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26(1), 27–56. <https://doi.org/10.2307/3587368>
- Long, M. H., & Norris, J. (2000). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597–603). Routledge.
- Malcolm, I. G., & Grote, E. (2007). Aboriginal English: Restructured variety for cultural maintenance. In G. Leitner & I. Malcolm (Eds.), *The habitat of Australia's Aboriginal languages: Past, present and future* (pp. 153–179). Mouton de Gruyter.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667336>
- Nunan, D. (2013). *Learner-centered English language education: The selected works of David Nunan*. Routledge.
- Ogilvie, G., & Dunn, W. (2010). Taking teacher education to task: Exploring the role of teacher education in promoting the utilization of task-based language teaching. *Language Teaching Research*, 14, 161–181. <https://doi.org/10.1177/1362168809353875>
- Oliver, R. (2020). Developing authentic tasks for the workplace using needs analysis: A case study of Australian aboriginal vocational students. In C. Lambert & R. Oliver (Eds.), *Using tasks in diverse contexts* (pp. 146–161). Multilingual Matters.

- Oliver, R. (2021) Developing a task-based approach: A case study of Australian Aboriginal VET students. In M. J. Ahmadian & M. H. Long (Eds.), *The Cambridge handbook of task based language teaching* (pp. 99–108). Cambridge University Press.
- Oliver, R., & Grote, E. (2010). (School name) CAPS student needs: Emerging themes. Edith Cowan University.
- Oliver, R., & Grote, E. (2016). A second language task-based needs analysis for Australian Aboriginal learners: Accessing insider perspectives of the workplace and cultural domains. In A. Mackey & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research* (pp. 224–242). Routledge.
- Oliver, R., Grote, E., Rochecouste, J., & Exell, M. (2012). Addressing the language and literacy needs of Aboriginal high school VET students who speak Standard Australian English as an additional language. *The Australian Journal of Indigenous Education*, 41(2), 229–239. <https://doi.org/10.1017/jie.2012.23>
- Oliver, R., Grote, E., Rochecouste, J., & Exell, M. (2013a). Needs analysis for task-based teaching: A case study of Indigenous vocational education and training students who speak EAL/EAD. *TESOL in Context*, 22(2), 36–50.
- Oliver, R., Grote, E., Rochecouste, J., & Exell, M. (2013b). A task-based needs analysis for Australian Aboriginal students: Going beyond the target situation to address cultural issues. *International Journal of Training Research*, 11(3), 246–259. <https://doi.org/10.5172/ijtr.2013.11.3.246>
- Oliver, R., Sato, M., Ballinger, S., & Pan, L. (2019). Content and language integrated learning classes for child mandarin L2 learners. In M. Sato & S. Loewen (Eds.), *Evidence-based second language pedagogy: A collection of instructed second language acquisition studies* (pp. 81–102). Routledge. <https://doi.org/10.4324/9781351190558-4>
- Platt, E., & Brooks, F. B. (2002). Task engagement: A turning point in foreign language development. *Language Learning*, 52(2), 365–400. <https://doi.org/10.1111/0023-8333.00187>
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
- Richards, J., & Rodgers, T. (2014). *Approaches and methods in language teaching*. Cambridge University Press. <https://doi.org/10.1017/9781009024532>
- Robinson, P. (Ed.). (2011). *Task-based language learning* (Language Learning 61: Supplement). Wiley-Blackwell.
- Sato, M., & Loewen, S. (Eds.). (2019). *Evidence-based second language pedagogy*. Routledge. <https://doi.org/10.4324/9781351190558>
- Savignon, S. J. (1987). Communicative language teaching. *Theory into Practice*, 26(4), 235–242. <https://doi.org/10.1080/00405848709543281>
- Serafini, E. J., & Torres, J. (2015). The utility of needs analysis for nondomain expert instructors in designing task-based Spanish for the professions curricula. *Foreign Language Annals*, 48(3), 447–469. <https://doi.org/10.1111/flan.12150>
- Trim, J. L. M. (2007). *Modern Languages in the Council of Europe 1954–1997*. Strasbourg. Retrieved on 11 February 2022 from www.coe.int/t/dg4/linguistic/.../TRIM_21janv2007_%20EN.doc
- Van Avermaet, P., & Gysen, S. (2006). From needs to tasks: Language learning needs in a task-based approach. In K. Van den Branden (Ed.), *Task-based language education: From theory to practice* (pp. 17–46). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667282.003>
- Wilkins, D. A. (1976). *Notional syllabuses: A taxonomy and its relevance to foreign language curriculum development*. Oxford University Press.

Can focus on form have an effect on language development?

Alessandro G. Benati
University of Hong Kong

This chapter is intended to provide an overview of the term/concept focus on form. We are grateful to Mike Long for providing the original definition of this construct, setting up the debate and actively stimulating more theoretical and empirical research on the role and nature of focus on form. The chapter highlights some of the shortcomings of the current research on the effects of focus on form on language development and provides suggestions for future research in this field of enquiry.

1. Focus on form

Focus on forms vs. focus on form

Originally, the term *focus on forms* was used by Long (1991, 2015) to describe a type of instruction which focus exclusively on the presentation and analysis of isolated discrete linguistic forms outside any authentic communicative context. This type of language instruction would follow a particular sequence determined by the perceived difficulty of the linguistic items of a target language (Doughty & Long, 2003). A typical classroom activity using a *focus on forms* approach consists of (i) explicit grammar rule explanations followed by (ii) mechanical output practice. Mechanical practice would include, but not exclusively: transformation exercises that manipulate a form; explicit error correction; the answering of display questions (i.e., questions to which the inquirer already knows the answers). *Focus on forms* adheres to the obsolete and discredited behaviorist model of second language acquisition.

Research findings in second language acquisition have shown that acquisition sequences do not reflect textbook instructional sequences and that teachability of a linguistic form/structure at any given point is constrained by learnability (Pienemann, 1998; Dyson & Håkansson, 2017). Additionally, due to lack of engagement in true

communicative tasks, *focus on forms* frequently demotivates students who are not interested in linguistic analysis. Overall, the picture which emerges is that *focus on forms* can be detrimental to second language acquisition.

The *focus on form* approach to language instruction emerged in opposition to *focus on forms* and it is sometimes referred to as the form-and-meaning approach. If L2 learners are encouraged to communicate while paying attention to form (Long, 1991), this type of pedagogical intervention would be a viable option. *Focus on form* refers to a variety of pedagogical interventions designed to shift L2 learners' attention briefly to language features in the course of performing a classroom task where the focus is on meaning (Doughty & Williams, 1998). *Focus on form* is drastically different from *focus on forms* (so-called traditional instruction). Long (1991) originally defined *focus on form* as purely *reactive* in nature. Doughty and Williams' (1998) definition of *focus on form* on the contrary includes planned and proactive *focus on form* activities as long as the following three major factors are taken into consideration: (i) the need for L2 learners to be engaged in meaning prior to giving attention to the linguistic code used to express it; (ii) the importance of identifying L2 learners' actual language problems that require intervention; and (iii) the need to keep the grammar component unobtrusive and fairly brief so that it does not detract from the meaning-focused nature of the practice.

The nature of language development and the role of focus on form

What is the best way to provide effective *focus on form*? There is an expectation from adult L2 learners to be taught grammar rules from language instructors. *Focus on forms* is when rules are explained, and this explanation is followed by some form of output practice, usually mechanical drill type of practice. However, as previously said, this approach suffers from a number of shortcomings (Wong & VanPatten, 2003). First of all, language is an abstract and complex system and it should not be taught or learned explicitly (VanPatten, Smith, & Benati, 2019). "Explicit learning is not part of acquisition" (Litchman & VanPatten, 2021: 1). Secondly, there is no clear mechanisms that turns explicit rules into the abstract and complex mental representation we call language. This paper argues for a view of language as a special human faculty and one that is very different than any other types of learning. Thirdly, language development is slow and piecemeal and L2 learners don't acquire one thing and then move on to another, as suggested by typical syllabi and textbooks using a *focus on forms* approach. Instead, L2 learners' minds are constantly working on various aspects of language simultaneously and only over time the internal system builds up and begin to resemble the second language. Fourthly, language development is stage-like and ordered-like. In the acquisition of morphemes/

structures of any language there are orders/stages that all L2 learners go through regardless of their L1. There is no evidence that stages can be skipped or orders can be altered particularly with the use of a *focus on forms* approach. Fifthly, there is no empirical evidence (only anecdotal evidence) that drill practice is useful or effective in second language acquisition. Finally, it is the input that provides the data for language acquisition. Input can be defined as the language that L2 learners hear and/or see in communicative contexts and forms the data on which the internal language mechanisms operates. The main position here is that comprehensible and meaningful input plays a crucial role in language development (Krahsen, 1982; Long, 1996; Litchman & VanPatten, 2021).

The fact that *focus on forms* is not necessary, it does not mean that *focus on form* does not have a role to play (Ellis, 2016). In fact, the position of this paper is to explore the how and why an effective *focus on form* element can be provided in the language classroom. Assuming that *focus on form* might have a facilitative role, the question is to determine what type of pedagogical intervention is more successful in terms of helping L2 learners internalize the grammatical features of a target language. If we are going to instruct L2 learners on formal properties of the language in any way in the classroom, it ought to be input based and meaning oriented. This idea falls out of what we know about the nature of language development. A *focus on form* is the preferred option in language instruction for the following reasons:

1. It might help L2 learners to develop a good level of attainment particularly if opportunities to natural exposure to language input are given;
2. It has a facilitative role when it is used for linguistic features, which are not too distant from the learner's current level of language development;
3. It might have a facilitative role in helping L2 learners to pay selective attention to form and form-meaning connections in the input. Learners make form-meaning connections from the input they receive as they connect particular meanings to particular forms (grammatical or lexical). For example, they tend to connect a form with its meaning in the input they receive (e.g. the morpheme *ed* on the end of a verb in English refers to an event in the past);
4. It should move from input to output practice (from structured input to structured output practice for example, see Benati & Batziou, 2019; Benati, 2021). Initially, input-based and interactional options to instruction might help L2 learners to internalize the grammatical features of a target language. Structured input practice for example is a type of *focus on form* that through input manipulations facilitates the processing of grammatical and syntactic features of a target language. Textual enhancement is another type of *focus on form* through which the input is made more noticeable and eventually processable (Benati, 2021a; Nassaji & Fotos, 2011). In order to help L2 learners notice a particular

feature, textual enhancement could provide L2 learners with typographical cues such as bolding and italics to draw their attention to grammatical forms in the text. A different pedagogical intervention would be to modify a text so that a particular target item would appear over and over again so that the text will contain many more exemplars of the same feature (input flood). One of the questions is: Can these types of ‘focus on form’ in certain cases and conditions enhance and speed up the way languages are learned?

A distinction to be made between explicit and implicit knowledge

Explicit knowledge of language is defined as conscious knowledge. It is often verbalizable knowledge about the language such as to talk about something you are still doing in the present, you add *-ing* to the verb stem. Implicit knowledge is defined as unconscious knowledge and it is not verbalizable. It can be described as the ability to understand or supply *playing* and not *play* in contexts that require the use of present continuous in English, and to do so without a conscious effort to retrieve the form.

Focus on forms approaches (i) instruct L2 learners about some specific grammatical forms (explicit information often using paradigms); (ii) L2 learners practice the target forms through mechanical practice; (iii) and at the end language teachers assess them using paper-pencil tests. There are two problems with this type instruction aiming at developing explicit knowledge: (1) it does not correspond to the way language develop in our mind/brain; (2) it does not correspond to the way L2 learners process information.

The acquisition of grammatical properties is mainly implicit. Language is too abstract and complex to be taught and learnt explicitly. L2 learners create linguistic systems in an organized way that seem little affected by external forces such as instruction and correction (Truscott, 2016). In short, language is not the rules and paradigms that appear on textbook pages (VanPatten & Rothman, 2014).

Explicit rules and paradigm lists can't become the abstract and complex system because the two things are completely different. What winds up in the human mind has no resemblance to anything on textbook pages or what teachers might say. This implication stems from the fact that there are no internal mechanisms that can convert explicit textbook rules into implicit mental representation. *Focus on forms* does little to foster the development of mental representation and tend to develop a learning-like behavior. Instruction does not have an effect on L2 learners implicit knowledge unless it is devised in a way that, on one hand, enhance the grammatical features in the input, and on the other hand, provide L2 learners with opportunities to focus on meaning (facilitating form-meaning connections). L2 learners must be exposed to input and that input must be comprehensible and message-oriented in order to facilitate language development. Language that L2 learners hear and see in

communicative contexts forms the data on which the internal mechanisms operate. The only effective way to facilitate language development (implicit knowledge) is the provision of good quality input (VanPatten, Smith, & Benati, 2019).

2. The effects of *focus on form*

Focus on form has an effect on the development of explicit knowledge. Paper and pencil tests have demonstrated this through a large date of empirical studies (Benati, 2021a). However, acquisition is not about the development of an explicit knowledge system, but the development of an implicit knowledge system called language representation (VanPatten, 2016). This implicit system is also a by-product of input processing in which there is a focus on meaning. Can *focus on form* influence the development of underlying implicit knowledge?

Route, rate and ultimate attainment in language development

Linguistic features of a new language are acquired by L2 learners in a specific order (route of acquisition) and natural route of development. *Focus on form* does not affect ordered development such as morpheme orders and staged-development. There is no clear and substantive evidence to affirm that *focus on form* is solely responsible for the fact that classroom L2 learners seem to acquire language faster than non-classroom L2 learners. What is clear it is that there is no effects on route of acquisition. It seems that it does not change the order in which things are acquired (VanPatten, Smith, & Benati, 2019).

Rate of development refers to the relative speed with which L2 learners progress through ordered development. Long (1983) claimed that classroom L2 learners might progress through ordered development more quickly than non-classroom L2 learners. Instruction might not affect route but it can make the rate of development faster. In the last forty-years, little research has compared the effects of *focus on form* on classroom and non-classroom L2 learners. The majority of the research on the effects of *focus on form* has focused on short-term comparative studies of classroom L2 learners only.

Research measuring input flood and input enhancement, as types of *focus on form*, seems to provide mixed results. Both pedagogical treatments are effective way to integrate *focus on form* component with the provision of opportunities for meaning-focused use of the target language. If L2 learners notice certain salient forms (Schmidt, 1990) because of their frequency in the input, they are more likely to acquire them than they are to acquire forms they have not noticed.

However, even if a learner has noticed a form but does not have the opportunity to process that for its meaning, acquisition might be delayed. Noticing does not guarantee processing. It is therefore vital to help L2 learners making form-meaning connections and to tell them what to pay attention to, what to notice and why they must change the way they process particular linguistic features in the input (sentence and discourse-level). VanPatten (2015b) is more interested in how L2 learners can make form-meaning connections while processing input. The success of this processing is determined by the characteristics of the features involved.

Overall, empirical research measuring the effects of different types of *focus on form* (Wong, 2005; Long, 2007; Nassaji & Fotos, 2011; Benati, 2021a) have used a pre-post-test design model in measuring instructional effects over a short period of time. These studies have been conducted in a classroom or laboratory contexts. Findings from these studies are normally mixed and difficult to interpret. Meta-analyses studies (e.g., Norris & Ortega, 2000; Spada & Tomita, 2010) have often provided some debatable conclusions. Overall, the main findings from studies seems to point to the fact that *focus on form* might have a beneficial role for language development. L2 learners seems to perform better after receiving instruction and post-test measures seem to confirm this main finding.

However, a number of issues would need to be considered (VanPatten, Smith, & Benati, 2019; Goo et al., 2015) before we can be absolutely sure that *focus on form* makes a difference: (i) clarification on what it is measured in this type of research; (ii) clear distinction between measurements of explicit knowledge and implicit knowledge; (iii) clear information about the nature and role of the type of *focus on form* investigated. The truth of the matter is that in the current empirical research database, no systematic analysis of the instructional treatments has been conducted regarding the effects of *focus on form*. In addition, much of the research measuring the effects of *focus on form* is only testing explicit knowledge.

In the case of ultimate attainment (how far L2 learners might become native-like), there is no empirical research measuring the effects of *focus on form* to address this issue.

In conclusion, research on the effects of *focus on form* clearly helps explicit knowledge, at least in the short-term. It is not clear what it does, if anything, for implicit knowledge. As we pointed out, we need more sophisticated research designs that use measurements (e.g., online tests) that attempt to test implicit knowledge and not explicit knowledge.

A number of factors need to be considered more closely in future research: (i) shift the focus of the research from whether or not to the how and why *focus on form* might make a difference; (ii) consider the effects of *focus on form* on implicit knowledge; (iii) find more suitable methodologies to measure implicit knowledge. For the research on the effects of *focus on form* to make advances, it needs to consider ways in which to assess implicit knowledge.

3. How focus on form might interact with development: An input processing perspective

An input processing perspective

The debate around the role of *focus on form* has been mainly concerned with whether or not *focus on form* has an effect on the route and rate of language development neglecting the fundamental questions of how and why *focus on form* interact with language development processes. In addition, much of classroom research has been carried out to measure L2 learners' output.

How do L2 learners process linguistic data? One possible way of researching the causes of the effects *focus on form* is to look at the interaction between instruction to which L2 learners are exposed to and the way they process input. Terrell (1991, p. 56) has posited two crucial questions: "what psycholinguistic processes are utilized in input processing?, how is the resulting information organised and accessed by the L2 learners?" Considering that the primary source of acquisition is the language input, Terrell (1991, p. 56) suggested that in order to comprehend input L2 learners must: (i) use strategies to make sense of unknown forms in the input; (ii) map meaning with new forms. The mental process mapping one form with one meaning is called by Terrell (1991, p. 56) *binding*. What L2 learners must perceive and store during acquisition processes are not the grammatical rules but single and correct meaning-form pairs. Terrell assigns a role for *focus on form* as long as meaningful input is provided and it contains many instances of the same form-meaning connection relationship.

The question is: Does a concentration of many examples of the same form-meaning relationship in one activity result in L2 learners to be able to focus on form and meaning at the same time? Considering that not too much demands should be put on L2 learners' attentional resources, *focus on form* with little lexical load coupled with high frequency of a single form-meaning relationship would result in helping L2 learners to allocate more attention and to accurately process the target form in the input.

VanPatten (1996, 2015a) argues that considering that L2 learners work on an internal schedule when it comes to language development, to investigate the role of *focus on form* in second language acquisition, it is necessary to explain the psycholinguistic processes utilized in input processing. What are these psycholinguistic processes? They are strategies and mechanisms used by L2 learners to process input. Input does not automatically enter L2 learners' brains during exposure for two main reasons: (1) humans have limited capacity of processing information; (2) input is filtered by internal processors (processing strategies).

Research on input processing has focused on issues such as how learners process input, what part of the input becomes intake, and investigated processing

strategies used by L2 learners to decode and store linguistic information. The two main processing strategies utilised by L2 learner are:

- The Primacy of Meaning Principle: L2 learners process input for meaning before they process for form;
- The First Noun Principle: L2 learners tend to process the first noun or pronoun they encounter in a sentence as the subject or agent.

The first processing strategy (The Primacy of Meaning Principle) indicates that L2 learners are driven to look for the message in the input before they look at how the message is grammatically encoded. L2 learners will direct their attention towards the detection of the content words in order to grasp the meaning of an utterance. The use of this strategy might delay form-meaning mappings. The second processing strategy (The First Noun Principle) indicated that L2 learners are driven to assign the role of agent (subject) to the first item they encounter in a sentence. The use of this strategy in certain cases (e.g., passive constructions, causative, case marking) might cause mis-understandings and delays in language processing.

What are the implications of this input processing perspective for *focus on form*?

- *focus on form* should take into account psycholinguistic processes used by L2 learners in input processing;
- *focus on form* should provide L2 learners with comprehension strategies to help them to establish accurate form-meaning connections'
- *focus on form* should provide L2 learners with more meaningful input, and many instances of the same form-meaning relationship.

The main concern from an input processing perspective is how L2 learners initially perceive and process linguistic data in the language they hear or read, and how they make accurate form-meaning connections. A type of *focus on form* called structured input (VanPatten, 1996. 2015b; Lee & VanPatten, 2003; Benati, 2019) can affect the acquisition process as a form-meaning connection focuser.

Structured input

Structured input is a type of *focus on form* aiming at altering the processing strategies L2 learners take to the task of comprehension and to encourage them to make better form-meaning connections. Structured input practice guides L2 learners to focus on small parts/features of the targeted language when they process the input. Given the emphasis on L2 learners' input rather than focusing in on the output, structured input consists of activities which offer the opportunity to interpret form-meaning relationship correctly without any practice in producing the target form or structure.

The main advantage for the use of structured input is that it is a type of *focus on form* that through the manipulation of the input and by providing L2 learners with comprehension strategies help them make accurate form-meaning connections. Structured input practice has the following characteristic: (i) It does take into account the psycholinguistic processes used in input processing; (ii) It alters the way L2 learners process input; (ii); it focuses on the processing of morpho-phonological units; it measures accurate processing (form-meaning connections and parsing of syntactic structures) through online (interpretation tests) and offline (self-paced reading and eye-tracking) testing. Grammatical rules are not tested; (iii) it is not concerned with production skills, but facilitating language processing might have an effect on L2 learners internal developing system. In other words, it might help L2 learners to access the relevant information to express meaning.

In structured input L2 learners are pushed to process the form/structure during activities in which the input is manipulated in particular ways to push the, to become dependent on the form to process the meaning (correct interpretation of the sentence or discourse).

Lee and VanPatten (1995, p. 104) have produced the following guidelines to develop structured input activities:

1. Present one thing at a time.
2. Keep meaning in focus.
3. Move from sentences to connected discourse.
4. Use both oral and written input.
5. Have the learner 'do something' with the input.
6. Keep the learner's processing strategies in mind.

There are two types of structured input activities: referential and affective. Referential activities force L2 learners to focus on a form and its meaning. Referential activities are those for which there is a right or wrong answer and for which the learner must rely on the targeted grammatical form to get meaning. Affective structured input activities are those in which learners express an opinion, belief, or some other affective response and are engaged in processing information about the real world.

A large database (Benati, 2019) exists measuring the offline effects (paper and pencil tests) of structured input on the acquisition of several linguistic features and languages affected by a variety of processing problems. Overall, the findings of this empirical research indicate that structured input has an effect on L2 learners' ability to process input accurately and appropriately (sentence and discourse), and these effects directly affect L2 learners' language development no matter their age, language background or language processing strategy under investigation (Benati, 2021a). Altering L2 learners' processing strategies leads to an alteration of their

developing language system, and these alterations are measurable at comprehension and production-level.

More recently, research evidence has been provided to support the effectiveness of structured input on online measurements such as eye-tracking and self-paced reading tests. Wong and Ito (2018) compared changes in processing patterns between L2 learners receiving structured input and traditional instruction on the acquisition of the French causative. In this pre-and post-test experimental study, a dichotomous scene selection eye-tracking task was used to measure eye movement patterns and accuracy in picture selection while learners were processing auditory sentences. The results from this study indicated that the structured input group gained higher scores for accuracy than the traditional instruction group. A change in eye movement was also observed in learners after the processing instruction training, but not after the traditional instruction training.

Benati (2020a) contrasted structured input and traditional instruction on the acquisition of English causative passive forms. In this study, an eye-tracking task was used to measure the online effects of the two treatments. The purpose of this study was to compare the effects of the two instructional treatments regarding the accuracy of how they process the target forms in the input. An eye-tracking picture selection task was used to measure these eye-movement patterns. The main results of this study indicated that the structured input training did have a positive effect on learners' eye movement patterns. They achieved significantly higher accuracy scores compared to traditional instruction groups. Benati (2021b) also carried out a replication of his previous study. This time, the effects of the two instructional treatments were measured on the acquisition of English passive forms. Results of the eye-tracking task revealed again that it was the structured input group who achieved significantly higher accuracy scores compared to the traditional. Structured input training is successful at causing a positive and relevant change in participants' eye-movement patterns. Chiuchiu and Benati (2020) compared structured input and textual enhancement on the acquisition of the Italian subjunctive of doubt. A self-paced reading test was used to measure sensibility to violations and accuracy in sentence interpretation containing a feature (subjunctive forms) affected by a combination of processing principles. The findings from this study showed that only the structured input group improved from pre-test to post-test on both behaviours observed. The initial findings of this new line of research measuring online effects of structured input does provide some initial evidence that structured input alters the processing principles that L2 learners use to process input in real time.

Benati (forthcoming) investigated online effects of structured input and structured output on the acquisition of English passive construction forms. The self-paced reading test used in this study is a more reliable measurements of

accuracy of response, response and reading time. The main findings from this experimental study confirmed the positive effects of structured input in facilitating the intake of passive constructions in English via correct parsing.

The positive effects of structured input measured by an online test contributes to the current discussion around the role of *focus on form* in second language acquisition. Structured input seems to speed up acquisition rate of particular linguistic features, and despite some psycholinguistic constraints to limit its role, it is successful at improving accuracy in interpretation and moment-by-moment processing.

Overall, the role of *focus on form* might be limited and constrained by a number of factors (e.g., orders and sequences of development, processing constraints). However, *focus on form* might have some beneficial effects in terms of helping L2 learners to pay selective attention to form and to make accurate form-meaning connections in the input. The question addressed by research measuring the effectiveness of structured input is why and how instruction might be beneficial. L2 learners must make form-meaning connections from the input they receive as they connect particular meanings to particular forms (grammatical or lexical). Evidence in second language research shows that the route of acquisition cannot be altered, however, *focus of form*, under certain conditions, might facilitate the speed rate of language development. What are these conditions?

A first condition is that L2 learners must be exposed to sufficient input. A second condition is that L2 learners must be developmentally ready for instruction to be effective. A third condition is that instruction must take into consideration how L2 learners process input. Input plays a key role in the acquisition of a second language and structured input certainly offers a way forward.

4. Focus on form: An agenda for future research

Limitations of current data and further research

Focus on form may help L2 learners build up explicit knowledge quickly, however it may not accelerate the development of implicit knowledge. As previously stated, much of the research within this framework is test-biased toward the measuring of explicit knowledge (VanPatten, Smith, & Benati, 2019). Empirical research must use measurements that probe underlying implicit knowledge that is qualitatively different from explicit knowledge.

What is needed is some methodological tools ready to measure in-depth language processing and implicit knowledge. Using online (real-time) measurements (e.g., self-paced reading, eye-tracking) is desirable for two main reasons: (1) it is a more effective measure of implicit knowledge as L2 learners have no time to access

explicit knowledge while are tested in real time; (2) it is a more effective measure of moment-by-moment language processing as interpretation/comprehension of language input is measured. The use of online methods to measure the effects of *focus on form* might offer us the possibility of more fine-grained information and analysis regarding moment-by-moment sentence comprehension and even a way to better measure implicit knowledge. This approach to empirical research in measuring the effects of *focus on form* has proven to be slightly elusive for researchers and therefore future studies need to address this shortcoming.

In addition to this, empirical research on the effectiveness of different types of *focus on form* (i) should better clarify what the nature of the *focus on form* treatment is, (ii) what it is trying to affect, and (iii) how the *focus on form* treatment reflect what we know about language development more generally. This line of research (online) is very much in its infancy. It should be expanded to include a wider variety of L2s, target forms and structures, and online methods.

Main findings from current empirical research measuring the effects of different types of *focus on form* can be described as mixed. This is because, the results from this research framework are not always clear. One of the main issues with the existing research data is the way scholars measure possible outcomes. This is because there is a general bias toward explicit testing. The question is: How do we know language development has happened after an instructional intervention?

Future research also need to address long-term effects of *focus on form*. Unfortunately, the great majority of existing empirical (except from some few studies such as VanPatten & Fernandez, 2004 and Benati & Batziou, 2019) findings has mainly measured short-term effects (Benati, 2021a) for *focus on form*. Are the effects of *focus on form* durative?

An interesting and emerging line of investigations is the one which investigates whether L2 learners can transfer/apply an effective instructional training received on a particular form/structure to another form/structure affected by the same processing/learnability constraints (Benati & Lee, 2008).

More empirical research is needed to establish the possible facilitative roles of *focus on form*. *Focus on form* of the kind that attempts to manipulate input to facilitate appropriate and accurate language processing seems to provide some promising results. If we were to review and rewrite Long's original paper (1983) on whether or not instruction makes a difference based on the current debate and research findings, we might be considering a new title: Does input manipulation make a difference?

If all the shortcomings highlighted in this paper can be addressed by researchers, the future of *focus on form* will be bright. We must be grateful to Mike Long for first defining this construct, setting up the debate and actively stimulating more theoretical and empirical research in this field to shed some light on the role and nature of *focus on form*.

Acknowledgements

I would like to express my gratitude to colleagues and anonymous readers who provided me with valuable comments and suggestions to improve this paper. The usual caveat of responsibility (mine alone) applies.

References

- Benati, A. G. (2019). Classroom-oriented research: Processing Instruction. *Language Teaching*, 52, 343–359. <https://doi.org/10.1017/S0261444817000386>
- Benati, A. G. (2020a). The effects of structured input and traditional instruction on the acquisition of the English causative passive forms: An eye-tracking study measuring accuracy in responses and processing patterns. *Language Teaching Research*, 24, 1–21. <https://doi.org/10.1177/1362168820928577>
- Benati, A. G. (2020b). The effects of structured input and traditional instruction on the acquisition of the passive constructions: An eye-tracking study. *Instructed Second Language Acquisition*, 4, 158–179. <https://doi.org/10.1558/isl.40642>
- Benati, A. G. (2021a). *Focus on form*. Cambridge University Press. <https://doi.org/10.1017/9781108762656>
- Benati, A. G. (2021b). *Input processing and processing instruction: The acquisition of Italian and modern standard Arabic*. John Benjamins. <https://doi.org/10.1075/bpa.11>
- Benati, A. G. (forthcoming) Structured input and structured output on the acquisition of English passive constructions: A self-paced reading study measuring accuracy, response and reading time.
- Benati, A. G., & Lee, J. F. (2008). *Transfer of training effects for processing instruction: Research and practice*. Multilingual Matters.
- Benati, A. G., & Batziou, M. (2019). Discourse and long-term effects of isolated and combined structured input and structured output on the acquisition of the English causative form. *Language Awareness*, 28, 77–96. <https://doi.org/10.1080/09658416.2019.1604721>
- Chiuocchio, G., & Benati, A. G. (2020). The effects of structured input and textual enhancement on the acquisition of Italian subjunctive: A self-paced reading study. *Instructed Second Language Acquisition*, 4, 235–257. <https://doi.org/10.1558/isl.40659>
- Doughty, C. J. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 206–257). Blackwell. <https://doi.org/10.1002/9780470756492.ch10>
- Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge University Press.
- Doughty, C. J., & Long, M. H. (Eds.). (2003). *The handbook of second language acquisition*. Oxford: Blackwell. <https://doi.org/10.1002/9780470756492>
- Dyson, B. P., & Håkansson, G. (2017). *Understanding second language processing: A focus on processability theory*. John Benjamins. <https://doi.org/10.1075/bpa.4>
- Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, 20, 405–28. <https://doi.org/10.1177/1362168816628627>
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris and Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). John Benjamins. <https://doi.org/10.1075/sibil.48.18goo>

- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Lee, J., & VanPatten, B. (1995). *Making communicative teaching happen*. McGraw-Hill.
- Lee, J., & VanPatten, B. (2003). *Making communicative teaching happen* (2nd ed.). McGraw-Hill.
- Lichtman, K., & VanPatten, B. (2021). Krashen forty years later: Final comments. *Foreign Language Annals*, 54, 1–5. <https://doi.org/10.1111/flan.12543>
- Long, M. H. (1983). Does second language instruction make a difference? *TESOL Quarterly*, 17, 359–382. <https://doi.org/10.2307/3586253>
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). John Benjamins. <https://doi.org/10.1075/sibil.2.07lon>
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp.413–468). Academic Press.
- Long, M. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge. <https://doi.org/10.4324/9780203850961>
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528. <https://doi.org/10.1111/0023-8333.00136>
- Pienemann, M. (1998). *Language processing and L2 development*. John Benjamins. <https://doi.org/10.1075/sibil.15>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263–308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Terrell, T. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal*, 75, 52–63. <https://doi.org/10.1111/j.1540-4781.1991.tb01083.x>
- Truscott, J. (2016). The effectiveness of error correction: Why do meta-analytic reviews produce such different answers? In Y.-n. Leung (Ed.), *Epoch making in English teaching and learning: A special monograph for celebration of ETA-ROC's 25th anniversary* (pp. 129–141). Crane.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex.
- VanPatten, B. (2016). Why explicit knowledge cannot turn into implicit knowledge. *Foreign Language Annals*, 49, 650–657. <https://doi.org/10.1111/flan.12226>
- VanPatten, B. (2015a). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp.113–135). Routledge.
- VanPatten, B. (2015b). Foundations of processing instruction. *International Review of Applied Linguistics*, 53, 91–109. <https://doi.org/10.1515/iral-2015-0005>
- VanPatten, B., & Fernández, C. (2004). The long-term effects of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 273–289). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410610195>
- VanPatten, B., & Rothman, J. (2014). Against “rules.” In A. G. Benati, C. Laval, & M. J. Arche (Eds.), *The grammar dimension in instructed SLA: Theory, research, and practice* (pp. 15–36). Continuum.

- VanPatten, B., Smith, M., & Benati, A. G. (2019). *Key questions in second language acquisition: An introduction*. Cambridge University Press. <https://doi.org/10.1017/9781108761529>
- Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. McGraw-Hill.
- Wong, W., & VanPatten, B. (2003). The evidence is IN: Drills are OUT. *Foreign Language Annals*, 36, 403–23. <https://doi.org/10.1111/j.1944-9720.2003.tb02123.x>
- Wong, W., & Ito, K. (2018). The effects of processing instruction and traditional instruction on L2 online processing of the causative construction in French: An eye-tracking study. *Studies in Second Language Acquisition*, 40, 241–268. <https://doi.org/10.1017/S0272263117000274>

Index

A

accessibility hierarchy 147, 175
accuracy 32, 34, 72, 102, 105,
117, 122, 126, 136, 142, 203, 205,
209, 213–217, 219–221, 227, 231,
233, 241, 266–267, 269
age XII, XIV, XVI, XXIII, 2, 5–7,
11, 15–17, 20, 24–25, 30, 52, 96,
163, 169, 173, 177–179, 183–186,
188, 190, 192–199, 201–210,
239, 265
effects XII, XIV, XVI, XXIII,
6, 177–179, 183, 188, 197–
198, 201–202, 206–207,
209
of acquisition 177–178, 183,
193, 195–196, 207, 209
of arrival 183, 185, 199
agent-based model 5, 29,
43–44, 52, 74
analytic syllabus 235, 239
aptitude XV–XVI, XXII–XXIV,
5, 15–17, 20, 24–25, 95–96,
129, 137, 141–142, 193–194, 198,
201–202, 204–209
aptitude-treatment interaction
XXIV, 137, 204
artificial intelligence (AI) 29
attainment 5, 11, 15–18, 20, 24–
25, 96, 119, 141, 143, 177–178,
192–194, 200–201, 204–205,
207–210, 259, 261–262
attention XVI, 1–2, 23–24, 55,
77, 79–82, 84–85, 87–88, 90,
93–94, 96, 98–99, 105, 116,
118–120, 122, 124, 127–130, 136,
147, 178, 180, 182, 191–192, 195,
201, 213, 215, 218, 220, 224,
227, 229, 239, 242, 258–260,
262–264, 267
attention allocation 79, 84,
87, 90, 93, 182

attentional resources 79,
91–92, 129, 134–135, 213,
215, 221–223, 229, 263
limited attentional capacity
214, 232
autonomy 115
B
backward transition probabilities
105
bad choice hypothesis 40,
42–43, 52
bibliometric analyses 13
bi-directional transfer 158
bilingualism XIX, XXII–XXIV,
27, 72, 98, 174–175, 184, 186,
190–195, 200, 209
blocking 84–85

C

categorisation 180–181, 184–
185, 189, 192
chaos 44–45, 47, 49–50, 72, 74
citation analysis 15, 17
classroom XVII–XVIII, 1–4,
6–7, 21, 23, 26, 79, 99, 114, 119,
123, 130, 142, 187, 197–199,
202, 204–206, 232, 238–239,
241–242, 253, 257–259, 261–
263, 269, 271
clustering 38
cognitive
maturity 204
processes 79, 81, 104, 130,
135–137, 180
restructuring 6, 177–178, 186
communication XVI, XXII,
XXIV, 3, 77, 79–81, 84, 113, 115–
117, 119–120, 122–124, 127–129,
131–132, 134–135, 138–142, 196,
217, 222–223, 225, 228–229,
232, 240, 243–245, 247
communicative
context 33, 133, 257, 270
language teaching 237, 255
task 125, 127, 132
competence 5, 13, 26, 52,
72, 78, 83–84, 88, 101, 103–
106, 108–110, 116–117, 125,
209, 228
complexity XIII, XVIII, 5,
22–23, 41–42, 44, 72–74, 102,
107, 173, 213–221, 223–224,
227, 231–233, 246
comprehensibility 4, 82, 98,
125, 128
comprehensible input 4, 7, 11,
26, 98, 141
computational resources 163
computer-assisted language
learning 139
conceptualisation 192, 211–215,
221
conscious knowledge 107, 260
conscious learning 83
constraints XXI, 4, 7, 11, 16, 26,
36, 60, 72, 74, 96, 123, 143, 173,
175, 178, 194, 208, 210, 222,
224, 267–269
constructed salience 85–86
context XIII, XVI, 4, 14, 30–34,
38–39, 41–42, 45, 52, 57, 67,
70, 77, 79, 82, 85, 87, 91, 109,
113, 116–121, 127–129, 131, 133,
135, 139, 150–151, 155, 158, 171,
179, 185, 196, 198, 204–206,
209, 211, 229, 240, 242, 255,
257, 270
conversational interaction
2, 17, 127–128, 142
corrective feedback XXII, 40,
80, 113–114, 116–118, 120–126,
137, 139–142

- critical period XII, XV–XVI, 6,
16, 24, 26, 174, 191, 193–194,
196–198, 206–210
- cultural learning 116, 236
- D**
- data collection 12, 51, 62, 67,
69, 200, 239–240, 247–248
- declarative learning 202
- decoding ability 163
- detection 5, 77, 81–82, 86–87,
89, 93–94, 264
- developmental dynamics 36,
42, 52
- developmental lock model 41
- developmentally moderated
transfer hypothesis 146
- discontinuity hypothesis 5,
101, 103
- dynamical systems theory 29,
35, 44, 71
- E**
- efficiency XIV, 144, 173
- ellipsis 147, 152, 173–174
- emergence criterion 32–34, 70
- Emergentism XXIII, 143–144,
174–175
- encoded 180, 182, 185–186, 264
- end-state 29, 40, 43
- enhanced incidental learning
XV, 5, 77–78, 81–82, 87, 89,
93–95
- entrenchment 41–42, 55,
58–60, 67, 70–71, 74–76, 146,
188–189, 194–195
punishment 58
- expertise IX, XI–XII, 115, 120
- explicit
- instruction 1, 77, 79, 81–84,
92, 94, 96, 122, 249, 269
- knowledge 77–79, 83–84,
87–89, 93–94, 96, 99,
136–137, 139–140, 200, 202,
205, 260–262, 267–268, 270
- learning XV–XVI, 79, 82–83,
86, 96, 140, 198, 201–202,
204–205, 258, 269
- rules 258, 260
- exposure 186, 189–190, 202
- external influences 237
- F**
- face-to-face 127, 130–132
- facilitative role 259
- familiarity 93, 97, 117, 142, 211,
213, 218, 230, 244
- feedback XVI, XXII, 2, 5–6, 16,
21, 35, 40, 80, 113–114, 116–118,
120–142, 185, 212
- delayed feedback XVI, 6,
127, 129–140
- negative feedback 2, 6, 16,
21, 35, 118, 127–130, 132–135,
139
- timing 6, 127, 129–132, 136,
138, 140
- filler-gap dependency 145
- fluency 72, 78, 116, 202, 213,
215–217, 219, 221, 223–224,
231, 233
- focus on form XVI, 1–2, 4–7,
21, 23, 26, 40, 77, 113–114, 116,
118–126, 128, 134, 140–143,
198, 204, 206, 254, 257–265,
267–270
- focus on forms 1, 4, 118, 128,
134, 237, 257–260
- foreigner talk 16, 26, 55, 71, 81
- forgetting 55–60, 69–71, 220
- form-and-meaning approach
258
- form-meaning connections
259–260, 262, 264–265, 267
- formulation XIII, 6, 72, 180,
211–215, 218–219, 221–222
- fossilization XIV–XV, 4–5, 16,
29–31, 34–35, 40, 43, 70, 73
- freedom 36, 217, 223–226,
228–229
- frequency 15, 80, 82, 84, 89, 93,
98–99, 101, 103, 105, 108–111,
178, 187, 218, 247, 261, 263
of daily usage 178, 187, 218
- G**
- gapping 147–150, 152–153,
173–174
- generative entrenchment
42, 58, 67, 74
- google scholar 9–12, 16, 26
- grammar 4, 36, 40–41, 43,
72, 78, 87, 91–93, 100, 106,
108–111, 125–126, 128, 141, 144,
173–175, 199, 202, 207–208,
237, 247, 257–258, 270
- grammatical learning XVIII,
XXII, 103, 111–112
- group work 11, 16, 23, 26
- H**
- h-index 12–13
- I**
- immediacy 130, 144–145,
148–149, 153–155, 158–162,
166, 170–172, 224
- immediate feedback 130–136,
139–140
- immersion 178, 185–186, 195,
198, 200, 204–206
- immigrants XXII, 198–199, 202,
204, 206
- implicational table 38
- implicit XV–XVI, XXII, 5, 16, 43,
70, 77–89, 92–94, 96, 98–99,
101, 104, 111, 122, 128–129,
135–142, 181, 198, 201–202,
204–205, 210, 260–262,
267–270
- competence 78, 83–84,
88, 101
- knowledge 5, 77–89, 92–94,
136, 142, 205, 210, 260–
262, 267–268, 270
- learning XXII, 79, 83–84,
86–87, 89, 94, 98, 198,
201–202, 205
- negative feedback 128–129
- incidental
- conditions 81, 91
- learning XV, XXI, 5, 77–79,
81–82, 87, 89–95, 99
- individual style 224
- inductive rule 109
- information structure 211, 233
- innatism 109–110
- input XIV, XVI, XIX–XX, 1–5, 7,
11, 15–24, 26, 40, 49, 56, 58,
77–78, 80–90, 93–101, 108–109,
115, 128, 130, 133–135, 141, 143,
146, 200, 204–206, 217–219,
225–226, 228, 237, 253, 259–271
- comprehensibility 128

- enhancement 5, 77–78,
80–82, 84, 86–87, 89,
93–97, 99–100, 261, 271
floored 80, 88, 99, 260–261
limited input 204, 206
processing 261, 263–265,
269–270
- instructed second language
acquisition XXIV, 7, 113–114,
121–122, 124–125, 197, 233,
253–255, 269
- instruction XIV, 1–4, 7, 11, 21,
26, 30, 77–79, 81–84, 87–88,
92, 94–99, 122, 125, 129, 140,
147, 173, 198, 202–209, 212,
225, 249, 257–263, 266–271
amount of 94, 202–203
- intentional learning 80–81,
88–89, 93
- interaction XIV, XVI, XXI–XXIV,
2–6, 11, 13, 15–20, 22–24, 26,
40, 43, 46, 60, 81–82, 113–114,
116–117, 119–128, 130–133,
137–138, 140–144, 157, 171, 192,
201, 204, 206–207, 212, 224,
232, 237, 239, 253–254, 263
enhancement 113–114,
119–122, 125
interaction hypothesis 2, 4,
6, 23, 127–128, 143
interactionist 235, 237, 252
interactive tasks 224, 226
modified interaction 128
- intercultural competence 116,
125
- interface XXI, 52–54, 82–84,
95–96, 111, 210
- interlanguage XIV–XV, 3–5, 11,
23, 26, 31, 33–34, 37–39, 60,
67, 71, 74, 101, 105, 111, 121, 124,
141, 174, 228
- irrationalism 110
- K**
knowability 103, 108
knowledge-driven 109
- L**
L2 competence 5, 103, 108
language aptitude XXIV, 16, 25,
96, 129, 141–142, 193–194, 201,
208–209
language tandem 113
leader 251
learnability 257, 268
learning context 116–117
lemmas 218, 220, 223, 226
lexicon *see* mental lexicon
lexis 24, 177–178, 213, 215–216,
231, 233, 237
linguistic relativity 6, 177–181,
183–184, 191–193, 196
linguistic simplification 5, 29,
52, 55
logistic map 44, 46–47, 49–51,
57
- M**
manner 24, 34, 180–185, 188–
190, 192, 194, 213, 228
mathematics 35, 44–45
maturational constraints
4, 7, 11, 16, 26, 96, 143, 178,
194, 208, 210
maximal input 204
meaningful communication
135, 138
mechanical practice 257, 260
mental lexicon (second language)
213, 218, 222, 229
methodological principles XIII,
21–22, 26, 119, 124
modality 91, 105, 132, 247
modifications 2, 23, 77, 82, 95,
98, 120
monitoring 92, 115, 136, 142,
213, 219, 231
morphology 2, 32–33, 35, 103,
128
morphosyntax 24, 103, 108–
109, 177, 199–200
motion XXIII, 57–58, 180,
182–184, 188–190, 192, 194–195
m-quotient 12
multidimensional model 32, 35
multilingualism XIX, XXIII, 96,
183, 208
- N**
needs analysis XIV, XXIV, 3, 6–7,
15, 22, 27, 143, 235, 253–255
negation 153, 169, 171–174
negative evidence 128
negative feedback *see feedback*
negotiation for meaning XVI,
2, 23, 82
no-interface position 83
non-verbal behaviour 179
noticeable 84–86, 259
noticing 2, 5, 77, 80–82, 85–87,
89, 93–97, 99, 121, 134–135, 137,
232, 262
null arguments 164
- O**
operation 44, 145, 162, 164, 171
output 2–3, 18, 24–25, 95–96,
120, 130, 233, 253, 257–259,
263–264, 266, 269
- P**
parallel mode of speech 6
path 37, 41, 61, 181–185, 188–
192, 224, 226
pattern-sensitive device 109
pedagogic procedures 119
pedagogical intervention 1, 4,
82, 86, 258–260
perceptual salience 5, 77, 80, 85
performance XVIII, XXIV, 3, 6,
34, 62–64, 92, 132, 139, 144,
174, 185, 187, 195–196, 205,
211–217, 219–223, 226–228,
230–233
performance science 103,
106–108, 112
performance (second
language) 195, 211,
232–233
phonology 24, 147, 177–178,
199–200, 218
planning XIII, 211, 213, 215–220,
223–228, 230–233, 238, 253
post-modernism 110
preparation 223, 226, 230
pre-verbal message 218–219,
221, 225, 228
primacy of meaning principle
264

- probability 107
- processing XIV–XV, XVII, XIX, XXI, XXIII, 3, 6, 32, 35–37, 40–41, 43–44, 52, 72–73, 84, 86, 91, 95–97, 99, 103–105, 112, 118, 134–137, 139–141, 143–147, 152–153, 156, 158, 163, 171–175, 178–179, 181, 189, 200, 208, 210, 219, 222–223, 227–228, 230, 232–233, 259, 261–271
- cost XV, 145–147, 153, 156, 158, 171
- possibility theory XIV, XX–XXI, 32, 35, 72–73, 146, 175, 269
- processable 37, 82, 259
- processable leeway 37
- process-oriented approach 3
- proficiency 93, 99, 106, 115–116, 121, 125, 141, 164, 175, 177–178, 184, 186, 199–202, 204–205, 207, 209, 218, 227–228, 230–231, 233, 242
- psychological state 118, 134
- Q**
- qualitative 48, 83, 102–104, 217, 223, 240
- quantifier 153, 157–158, 173–175
- quantitative 98, 102, 104, 115, 217, 240, 270
- R**
- rational inquiry 102, 110
- rationalism XVI, 101–102, 110
- reasoning demands 216
- recast 131, 134
- reciprocity 115
- repetition XIII, 104, 131, 211, 213, 216–219, 223, 225–228, 230, 232–233, 253
- revision 104, 144, 148, 153, 159, 170
- S**
- saliency 5, 77, 80, 82, 84–86, 94, 96, 142, 201, 207
- scopus 9–10, 16–18, 23, 26
- sensitive period(s) 16, 24–25, 96, 141, 177, 186, 194, 196, 208–210
- serial mode of speech 211, 223
- simplifying 2, 38–40, 55, 58, 68
- stabilization 4, 29–30, 34–35, 37, 40, 43, 45, 50–51, 60, 70, 73
- statistical learning 86, 88, 96, 103
- steadiness 34–35
- strong-interface position 83
- structural domain 34, 70
- structured input XIV, XIX, 259, 264–267, 269
- syllabus
design 6, 11, 21, 235–238, 254
synthetic syllabus 118, 236–237, 252
- T**
- tandem learning XVII, 113–126
- task XIII–XIV, XVIII, XXII, XXIV, 2–7, 9, 11, 15–18, 21–22, 25–27, 32, 34, 61, 77, 92, 97, 119–120, 122–125, 127, 130–133, 136–137, 142–143, 151, 157, 167, 170, 184–187, 192, 200, 211–221, 223–224, 226–228, 230–233, 235–236, 238–239, 242, 245–247, 250–255, 258, 264, 266, 270
- characteristics 6, 211, 213, 216, 228, 233
- conditions XIII, 6, 211, 213–214, 221, 227–228, 230
- design 120, 122, 233, 245
- task-based
language teaching XIII, XVIII, XXII, 3, 5, 7, 15, 17–18, 25–27, 77, 97, 119, 124, 143, 232, 235, 253–254, 270
- needs analysis 235
- learning XXIV, 7, 212, 231, 233
- technology XXII, 25, 27, 80, 114–116, 122–126, 139–141
- temporal overlay maps 18
- thinking-
after-speaking 182, 184
before-speaking 182
during-speaking 182
for-speaking 6, 177, 179, 181, 183, 191–192, 194
- time
perspective 216, 233
pressure 213, 217, 219–220, 225, 229, 233
- tipping points 29, 43–44, 52, 58, 74
- training paradigms 187
- transfer XV, XXIII, 6, 35, 143–151, 153–158, 160–164, 167–172, 175, 183, 195, 213, 217, 268–269
- transfer calculus 146–151, 153–156, 158, 160–164, 167–168, 171–172
- typographical enhancement 86–90, 93
- typology 35, 144–145, 173, 182–183, 189, 195
- U**
- ultimate attainment 11, 15–18, 20, 24, 119, 143, 178, 192, 200–201, 204–205, 207–208, 261–262
- uncertainty XVI, 5, 101, 103, 106–107, 110
- unconscious learning 79, 82
- V**
- variability 30, 32, 34–35, 71–72, 183, 210
- verbal behaviour 179
- visual perception 181
- vocational
education training xxiv, 236, 243
skills training programs 244
- W**
- weakness corollary 163–164, 167–168
- web of science 9–10, 16–17, 26
- web-based communication
technology 122
- working memory 93, 98, 129–130, 141–142, 144, 163, 227–228
- wrong track pathway model 52

Professor Michael H. Long (1945-2021) was one of the most influential scholars in the field of second language acquisition. This volume presents a set of chapters that honour some of his key contributions in language teaching and learning. Following a bibliometric analysis of the impact of his research to the field, the volume spans topics such as task-based language teaching, focus on form, age effects, transfer, feedback, interaction, incidental learning, stabilization, among many others.

“A rich and often moving collection of tributes to Mike Long: for lucky ones, as a friend, or as a supervisor, co-researcher or author, not forgetting his work as editor - all sui generis; for all, one of the pre-eminent voices and influences on the development of applied linguistics. This volume, which sometimes reads like a written colloquium, brings together a variety of styles, tones and topics by distinguished scholars, from a welcome range of generations of applied linguists. Together the work they report offers a memorable reflection on Mike Long, the person and the scholar.”

Martin Bygate, *Lancaster University, England*

“The cognitive turn in second language acquisition research in the 1970s triggered a wealth of theoretically informed empirical studies. Few researchers had as profound an impact on these developments over the subsequent 40 years as Michael H. Long. It is gratifying to see that the studies in this volume live up to the depth and breadth of Mike’s own research and honour his memory beautifully.”

Jürgen Meisel, *University of Hamburg, Germany / University of Calgary, Canada*

“This volume represents the collective efforts of outstanding scholars with expertise in key topics in SLA, including age, TBLT, transfer, incidental learning, and context of acquisition. The collection presents new evidence masterfully couched within a long-term perspective on the issues; it is a true testament to Mike Long’s extraordinary academic prowess and to his long-lasting influence in the field.”

Cristina Sanz, *Georgetown University, United States*

ISBN 978 90 272 1256 6



9 789027 212566

John Benjamins Publishing Company