The Bilingual Mind and Brain Book Series 1

Roberto R. Heredia
Jeanette Altarriba
Anna B. Cieślicka *Editors*

Methods in Bilingual Reading Comprehension Research

Springer

The Bilingual Mind and Brain Book Series

Series Editors
Roberto R. Heredia
Anna B. Cieślicka

More information about this series at http://www.springer.com/series/13841

Roberto R. Heredia • Jeanette Altarriba
Anna B. Cieślicka
Editors

Methods in Bilingual Reading Comprehension Research

Springer

Editors
Roberto R. Heredia
Department of Psychology
and Communication
Texas A&M International University
Laredo, TX, USA

Jeanette Altarriba
Department of Psychology
University at Albany
State University of New York
Albany, NY, USA

Anna B. Cieślicka
Department of Psychology and
Communication
Texas A&M International University
Laredo, TX, USA

The Bilingual Mind and Brain Book Series
ISBN 978-1-4939-2992-4 ISBN 978-1-4939-2993-1 (eBook)
DOI 10.1007/978-1-4939-2993-1

Library of Congress Control Number: 2015949638

Springer New York Heidelberg Dordrecht London

Printed on acid-free paper

Springer Science+Business Media LLC New York is part of Springer Science+Business Media (www.springer.com)

Para mis seres queridos y toda esa gente linda que ha dejado huella en mi cuaderno de memorias

Roberto R. Heredia

This work is dedicated to all those who inspire the desire to conduct research, those who carry out that research, and those who contemplate and question the workings of the human mind

Jeanette Altarriba

To my family, friends, colleagues, and students—past, present, and future—who continue to help me become who I am as a person, a teacher, and a researcher

Anna B. Cieślicka

The Bilingual Mind and Brain Series

We take great joy and honor in presenting *The Bilingual Mind and Brain Series* to students, teachers of bilingualism, and the scientific community. This book series is intended to advance and contribute to our understanding of the bilingual/multilingual mind and brain, both as an academic discipline and as a maturing research field. *The Bilingual Mind and Brain Series* is interdisciplinary in its scope and examines the bilingual mind/brain from such perspectives as psycholinguistics, cognitive psychology, cognitive science, and cognitive neuroscience, as well as applied linguistics and pedagogical approaches to second/foreign language learning.

The Bilingual Mind and Brain Series seeks to publish cutting-edge and provocative collective volumes and monographs about how the bilingual mind and brain process, learn, and store information, and it is intended for the growing number of bilingual researchers and practitioners interested in understanding the behavioral aspects and neurobiology of bilingualism, as well as the dynamic character of the bilingual/multilingual/second language learner's mind. Its purpose is to provide updates of the most current work in the behavioral and neuropsychological research fields of bilingualism/multilingualism and second language acquisition. The books in the *Bilingual Mind and Brain Series* are intended to contribute to the development and establishment of the Bilingual Cognitive Psychology and Bilingual Cognitive Neuroscience disciplines as subfields of Cognitive Psychology and Cognitive Neuroscience, as well as to contribute to our understanding of the bilingual mind and brain, and ultimately, the human brain.

We are also honored and humbled to present *Methods in Bilingual Reading Comprehension Research* as the first volume of the series. This volume, for the first time, provides a much-needed set of methodological tools and perspectives (behavioral, connectionist, and brain imaging paradigms) to better understand bilingual reading processes. *Methods in Bilingual Reading Comprehension Research* is timely, and it is the first book of its kind to present a comprehensive overview of the various psycholinguistic and neurophysiological tasks used to measure bilingual language processing.

Roberto R. Heredia
Anna B. Cieślicka

Preface

The presentation of a volume on *Methods in Bilingual Reading Comprehension Research* is timely and provides a much-needed set of methodological tools to further understand reading and sentence processing in bilinguals. Although much has been written on the bilingual's lexical and conceptual representations—mostly at the word level—much less is known about the ongoing bilingual reading processes and the appropriateness of the experimental tasks, as bilinguals comprehend cross-language and mixed language information at the sentence and connected text levels. *Methods in Bilingual Reading Comprehension Research* reviews and presents new bilingual and cross-language reading findings from classic behavioral experimental techniques, such as the *rapid-serial visual presentation* (RSVP) task, the *visual moving window* (VMW), the *cross-modal lexical priming* (CMLP) task, the *eye-tracking paradigm*, and the latest brain-viewing neuropsycholinguistic/neurobiological methodologies including *event-related potentials* (ERPs) and *functional magnetic resonance imaging* (fMRI).

This book is written from an empirical/methodological perspective, and it provides readers and undergraduate/graduate students with the opportunity to acquire *hands-on* experience in the development of basic reading experiments. Each chapter includes a *Suggested Student Research Projects* section. Selected chapters include detailed procedures on how to design and develop reading experiments using sample scripts from experiment builder software (e.g., E-Prime, PsyScope, OpenSesame). *Methods in Bilingual Reading Comprehension Research* has been conceived of as an advanced book for both the undergraduate and graduate levels and represents the first such text of its kind to critically examine the contribution of behavioral, brain-viewing, and computational (i.e., computer simulations) approaches to better our understanding of the bilingual's basic reading processes. To our knowledge, no other published book has addressed these issues directly. It is hoped that this book contributes to the development and establishment of *Bilingual Reading* as a subfield of bilingual sentence processing and will fill a significant gap in the literature on bilingual language processing.

Finally, we would be remiss, however, if we did not acknowledge Besner and Humphreys (1991), Coltheart (1987), and Rayner and Pollatsek (1989) as classic texts that shaped our understanding of reading, and particularly Kieras and Just (1984) that served as the inspiration for *Methods in Bilingual Reading Comprehension Research.*

Laredo, TX — Roberto R. Heredia
Albany, NY — Jeanette Altarriba
Laredo, TX — Anna B. Cieślicka

References

Besner, D., & Humphreys, G. W. (Eds.). (1991). *Basic processes in reading: Visual word recognition.* Hillsdale, NJ: Lawrence Erlbaum.
Coltheart, M. (Ed.). (1987). *Attention and performance: The psychology of reading.* Hillsdale, NJ: Lawrence Erlbaum.
Kieras, D. E., & Just, M. A. (Eds.). (1984). *New methods in reading comprehension research.* Hillsdale, NJ: Lawrence Erlbaum.
Rayner, K., & Pollatsek, A. (Eds.). (1989). *The psychology of reading* (1st ed.). Hillsdale, NJ: Lawrence Erlbaum.

Acknowledgements

We would like to thank the many people who helped us directly and indirectly with the completion of this book. First, we want to thank Welmoed Spahr and Morgan Ryan of Springer for catching the vision of this unique book on bilingual reading. We would like to express our gratitude to our contributors, colleagues, and friends, who were extremely cooperative in meeting our deadlines, and for graciously accepting our suggestions in earlier drafts. We couldn't have done it without their dedication, diligence, and professionalism.

I, Roberto, dedicate this book to the late Elizabeth Bates, my postdoctoral mentor at the Center for Research in Language (CRL) at the University of California-San Diego. She discovered that I was a "Popperian" (I didn't know) and that I somewhat leaned towards the modularity of the mind (I didn't know that either). My only regret is that I did not have the opportunity to thank her in person for helping me rediscover my interests in the "big issues" on the nature of language and language processing. Sometimes I wake up, come to work at 8:00 AM and wonder if Liz is already up (en El Más Allá) typing 300 words per minute, and speaking Italian 100 miles per hour! Indeed Texas A&M International University's Cognitive Neuroscience Laboratory (our lab), to some extent, was created in her own image—a carbon copy of the CRL of the late 1990s! I also dedicate this book to the late David Swinney, my friend, and mentor. Dave, the modularist, the ambiguity, the *cross-modal lexical task* (CMLP which, in my view, is one of the most powerful tools to truly measure online processing), the scholar that designed those very simple and yet powerful experiments! And he was always right! Many experiments could and would benefit from good methodology and the good old CMLP. His influence, of course, is very obvious! Para mis seres queridos, y en especial para *mi querido viejo* que aunque trabajaba de sol a sol en los campos tomateros de California, siempre tuvo tiempo para dedicarle a su familia. Para mi mamá que siempre me consintió todo, pero que sin embargo, forjó las bases de mi futuro académico. Para mi esposa y mi hija que siempre se preocupa cuando llego tarde a casa. Para mis estudiantes y toda esa linda gente que ahora es parte de mi ser.

To follow suit, I, Jeanette, also recollect the important impact of my mentors in the development of my career and in helping me to carve out what I call a *life by design*. To Ronald P. Fisher who inspired me to ask questions of the mind and the brain and to design just the right studies to attempt to answer those questions. Our many talks and discussions led to the confidence in me to pursue graduate studies, though far away from home. It was then, Timothy P. McNamara who taught me how to think clearly and rationally, and how to look at the data, think about the data, and draw the right conclusions, always with a hint towards asking the next set of interesting and provocative questions. *What a life*! I thought. Working towards a clearer understanding of mental processes and the mind by considering clever and insightful ways of pursuing that understanding—and getting paid for it! It has been a journey that has taken me near and far to interesting places, meeting interesting people, and fostering the love of that journey in undergraduate and graduate students. These individuals were fundamental in driving my desire to always move forward and to succeed. Thus, this work is truly dedicated to my many mentors and guides.

I, Ania, would like to acknowledge all those people who inspired me to pursue research and teaching as my career. Professor Waldemar Marton who showed me how fascinating the area of second/foreign language acquisition was and how much of a rewarding experience it would be to design and carry out our own research. Krystyna Droździał-Szelest, my former instructor and friend, who has been a model of a true mentor and a caring professor throughout all my years back at the School of English, Adam Mickiewicz University. I also wish to thank David Singleton, Professor of Applied Linguistics at Trinity College, Dublin, and currently Professor at the University of Panonnia, Veszprém, for inspiring my interest in the bilingual mental lexicon. It was his 1991 *Second Language Research* article on the second language lexicon that caught my attention while I was waiting in a post-communist Poland's long line in Collegium Novum's student cafeteria "Kafka" to get my coffee and flipping through the stack of journals I always carried with me as a young aspiring researcher. The line was long enough to allow me to read the whole article and, from this point on, I knew I would devote all my research energies to solving the mysteries of the bilingual mind/brain. Finally, I would be remiss without acknowledging the ever-inspiring Katarzyna Dziubalska-Kołaczyk, Dean of the Faculty of English, where I first started my academic career, for constantly pushing all of us faculty to strive for excellence, and for her enthusiastic *where there is a will, there is a way* attitude, which is so much needed in pursuing research. Last but not least, dziękuję moim ukochanym Rodzicom i siostrze za ich nieustające wsparcie, to my very dear friends, those far away in Poland and here in Laredo, for their positive energy which knows no boundaries, and to all the students I have had the honor to teach and work with.

The writing of this book was partially supported by grants P031M105048 from the U.S. Department of Education to Roberto R. Heredia, and a National Science Foundation (NSF) Major Research Instrumentation (MRI) Grant (BCS-1229123) to Anna B. Cieślicka. Finally, we wish to acknowledge all our past and present undergraduate and graduate students, our colleagues and friends, too numerous to name, and our wonderful families—related, extended throughout the world (in Poland, Cuba, Miami, Mexico, and California). It is their love and support that always motivates and encourages us. *¡Con mucho cariño para todos ustedes! Z najserdeczniejszymi podziękowaniami dla Was wszystkich! With gratitude and thanks to all of you!*

Con el cariño de siempre,
Roberto, Jeanette, and Anna

Contents

Contributors

Wualú A. Altamira Department of Psychology and Communication, Texas A&M International University, Laredo, TX, USA

Jeanette Altarriba Department of Psychology, University at Albany, State University of New York, Albany, NY, USA

Eva Van Assche Department of Experimental Psychology, Ghent University, Ghent, Belgium

Angélique M. Blackburn Department of Biology, University of Texas at San Antonio, San Antonio, TX, USA

Joanna C. Bovee Department of Psychology, University of Illinois at Chicago, Chicago, IL, USA

Anna B. Cieślicka Department of Psychology and Communication, Texas A&M International University, Laredo, TX, USA

Maria Dimitropoulou Basque Center on Cognition, Brain and Language, San Sebastián–Donostia, Spain

Margaret Gillon Dowens School of English, The University of Nottingham, Ningbo, China

Jon Andoni Duñabeitia Basque Center on Cognition, Brain and Language, San Sebastián–Donostia, Spain

Wouter Duyck Department of Experimental Psychology, Ghent University, Ghent, Belgium

Eva M. Fernández Queens College and Graduate Center, City University of New York, Flushing, NY, USA

Omar García Cognitive Neuroscience Laboratory, Texas A&M International University, Laredo, TX, USA

Patricia G. González Cognitive Neuroscience Laboratory, Texas A&M International University, Laredo, TX, USA

Robert J. Hartsuiker Department of Experimental Psychology, Ghent University, Ghent, Belgium

Roberto R. Heredia Department of Psychology and Communication, Texas A&M International University, Laredo, TX, USA

Dieter Hillert School of Medicine, University of California, San Diego, CA, USA

Daniel Holman Cognitive and Information Sciences, University of California, Merced, CA, USA

Belem G. López Department of Psychology, Texas A&M University, College Station, TX, USA

Clara Martin Basque Center on Cognition, Brain and Language, San Sebastián–Donostia, Spain

Jennifer M. Martin Department of Psychology, University at Albany, State University of New York, Albany, NY, USA

Nicola Molinaro Basque Center on Cognition, Brain and Language, San Sebastián–Donostia, Spain

Yoko Nakano School of Human Welfare Studies, Kwansei Gakuin University, Nishinomiya City, Hyogo, Japan

Irina Pivneva Department of Psychology, Centre for Research on Brain, Language & Music, McGill University, Montreal, QC, Canada

Gary E. Raney Department of Psychology, University of Illinois at Chicago, Chicago, IL, USA

Ricardo Augusto Souza Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Michael J. Spivey Cognitive and Information Sciences, University of California, Merced, CA, USA

Debra Titone Department of Psychology, Centre for Research on Brain, Language & Music, McGill University, Montreal, QC, Canada

Veronica Whitford Department of Psychology, Centre for Research on Brain, Language & Music, McGill University, Montreal, QC, Canada

Chapter 1
Introduction to Bilingual Research Methods

Roberto R. Heredia, Anna B. Cieślicka, and Jeanette Altarriba

Abstract The current chapter introduces the reader to the vision and the contents of the present volume—*Methods in Bilingual Reading Comprehension Research*. The focus is on traditional as well as newly developed methodological approaches to the study of bilingual reading. Findings are critically reviewed stemming from the well-known behavioral approaches to those that are neuropsycholinguistic in nature. The ways in which reading comprehension is measured are critically important to the eventual outcomes of empirical work and to their theoretical significance. An emphasis is placed on the advantages and the challenges of using particular methods to examine language representation and cognition with caveats where necessary to inform the researcher as to the limitations or benefits of employing a particular technique. Overall, methods are used to understand how cognitive processes operate and how the mind interprets stimuli, regardless of the particular language that is known or spoken. This compendium is meant to provide a comprehensive overview of the methodology and empirical techniques that can be used to accurately assess reading behavior in bilingual readers and to stimulate new research directions regarding bilingual reading comprehension.

R.R. Heredia, Ph.D. (✉) • A.B. Cieślicka, Ph.D.
Department of Psychology and Communication, Texas A&M International University,
5201 University Boulevard, Laredo, TX 78041-1900, USA
e-mail: rheredia@tamiu.edu; anna.cieslicka@tamiu.edu

J. Altarriba, Ph.D.
Department of Psychology, University at Albany, State University of New York,
1400 Washington Avenue, Albany, NY 12222, USA
e-mail: jaltarriba@albany.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_1

Introduction

Theory is interesting because it attempts to provide and explain answers. Data are interesting because they provide researchers with the so-called evidence to support or not support a particular hypothesis generated by a theory. But the real issue is that data are as good as the methods used to obtain them. We continuously remind our students and ourselves that *it is not what you get*, *but how you did it*, both in relation to results (i.e., data) and research methods used to obtain them. As researchers, we sometimes are highly influenced by powerful statistical innovations, but in the end, everything boils down to methodology. *How did you measure it*? *What task did you use*? *Why did you choose that particular task*? *What demand characteristics are afforded by a particular task*? A few years ago, one of us attended a research presentation (a dissertation defense to be exact) in which many in the audience were absolutely enamored with the straightforward and interpretable crossover interaction (see, for example Garcia-Marques, Garcia-Marques, & Brauer, 2014; Loftus, 1978). No one in the audience dared to ask the *how* or *what* question. To paraphrase the late David Swinney, one of our mentors, *It*'*s all about the task*; *it boils down to the particular task and whether the task is sensitive enough to measure language processing in real time*. To this we add Blackburn's (Chap. 12) observation that *each fMRI experiment is as good as its design* [*and task*]… and *it is important to define a simple and robust design without too many conditions*… (Duñabeitia et al., Chap. 11).

Although other excellent bilingual research books exist (e.g., Wei & Moyer, 2008) that provide a "know-how" for researchers to investigate bilingualism using such methodological approaches as experimental, correlational, psychophysiological, and archival methods, the purpose of *Methods in Bilingual Reading Comprehension Research* is to present an overall review of some of the psycholinguistic techniques and approaches that have been typically used in bilingual reading research and to propose other possible tasks that may prove viable in investigating such theoretical issues as bilingual lexical ambiguity resolution, or how bilingual speakers might resolve multiple sources of potentially conflicting information as they comprehend sentences and discourse during the communicative process. The focus of this book is on reading, broadly defined. For other experimental techniques that have been used or could be used in bilingual language research, the prospective bilingual researcher is strongly encouraged to consult Grosjean and Frauenfelder's (1996) classic guide to spoken word recognition paradigms, Kieras and Just's (1984) excellent volume on monolingual reading research paradigms, and Heredia and Stewart (2002) for bilingual on-line spoken language research paradigms.

How to best measure bilingual language processing, and bilingual reading in particular? A somewhat related issue is whether bilingual reading research is sufficiently different from monolingual research, so as to necessitate its own methods specifically designed to measure *bilingual processing*. The answer to the first question is that it would be difficult to pinpoint good vs. bad experimental techniques, but it would be quite acceptable to list advantages and disadvantages of each task. For example, one would argue that eye-tracking (Chap. 8) is one of the most precise and ecologically valid techniques to measure bilingual reading, since it provides insights into early and late processing stages of language comprehension (e.g., first pass reading time vs. total reading time). Measurements such as the number of fixations or regressions are also important in understanding the overall pattern of basic reading behavior. However, eye movement research requires special equipment that in practice is rather costly (but see the Internet Related Section for free/open source hardware and software alternatives); the same could be said of other highly technologically advanced approaches, such as *event-related potentials* (ERPs, Chaps. 10–12), capable of identifying the fast time course of linguistic processes on account of the exquisite temporal resolution of the technique, or brain viewing technology (Chap. 12). Lack of such advanced equipment would make the *visual moving window* (VMW) and its variants, as well as the *maze task* (Chap. 5), fairly attractive possibilities based on their ease of use and availability. Other tasks such as *rapid serial visual presentation* (RSVP, Chap. 4; see also Altarriba & Soltano, 1996; Schwartz & Kroll, 2006) may also prove useful, especially when used in conjunction with eye-tracking paradigms (Altarriba, Kroll, Sholl, & Rayner, 1996) and other tasks such as the VMW or self-paced reading (cf. Witzel, Witzel, & Forster, 2012).

In order to answer the second question concerning specificity of bilingual vs. monolingual research methodology, we need to acknowledge that some areas of bilingualism do indeed require bilingual-specific methodology, such as for example, translation-related research. In a typical *translation task* (e.g., De Groot, 1992; Kroll & Stewart, 1994), participants are presented with a word in one language (e.g., *house*) and are asked to generate a word translation in the other language (e.g., *casa*). Reaction time is taken as an index of translating or retrieving a word/concept from the second language mental lexicon. In the *cued translation* task, the participant might be provided with a cue of a possible translation (e.g., *c_* for *casa*). The translation task has been widely used in the bilingual literature and it has proven successful, as its effects are robust and easy to interpret in the context of existing bilingual models. However, because it relies on bilingual performance, it is not clear if its effects might be due to ease or aptness in translating across languages. The other exclusive bilingual technique is the so-called *translation-recognition task* (De Groot, 1992). In this task, participants are typically presented with translation (e.g., *house-casa*) and nontranslation word pairs (*house-dedo*: translation: *finger*). During the course of an experiment, the participant's task is to determine if a target word is a translation (e.g., *casa*) or nontranslation (e.g., *dedo*) of the preceding prime (*house*). Assuming that the translation and nontranslation targets are properly controlled (e.g., word frequency, length), a facilitation effect might be computed where the response times obtained for the nontranslation target are subtracted from the times obtained for the translation target. Translation pairs are typically faster than nontranslation pairs. Although this task provides a much needed improvement on the normal/cued translation task, one criticism is that this task, too, relies heavily on "conscious awareness" (e.g., Roediger, 1990), and because the task is so predictable (i.e., intentional task), participants are very likely to develop strategies such as putting themselves into a "bilingual mode" as they perform it.

A much more improved task could be the one that could be called a *translation lexical decision task* or a *translation naming task*. This translation lexical decision task would be similar to the translation-recognition task, except that instead of determining whether word pairs are translation equivalents, participants would determine whether a target is a real word (e.g., *casa*, *dedo*) or a nonword (*dedi*), in the particular language(s) of interest. For the translation naming task, participants would simply name a related or unrelated target. The facilitation effect, in this case "priming," would be computed the same way as in the translation-recognition task. The priming effect, in this case, would be called the *translation priming effect*. Notice that it would also be possible to come up with another "purely" bilingual task, similar to the bilingual translation lexical decision task, in which participants decide whether a target in a second language is a word or a nonword. In this case, however, one of the target words would be a word associate (e.g., *window* for *house-window*). We would call this task the *cross-language lexical decision task*, and the effects resulting from this task would be called *cross-language priming*. However, these proposed tasks have already been used in bilingual and monolingual research and are collectively known as *lexical decision tasks* (see for example, Heredia & Cieślicka, 2014 for a review on translation and cross-language priming). Although developed to investigate monolingual language-related issues, these tasks can be used with bilingual, multilingual, or monolingual populations. However, even when some measures can be used cross-linguistically, researchers must know and understand the nature of the task (i.e., task demands; e.g., Witzel et al., 2012) or what exactly the task measures and whether it requires a special kind of expertise (as in variants of the translation task). Case in point is the *sentence verification task*, which has been used to measure semantic memory and has been critical in the development of models of memory storage. Typically, participants in this task are presented with a sentence such as *Canaries are birds* and their objective is to respond if the sentence is true or false. Their reaction times are recorded. This task has been classified as a semantic memory task that measures concept activation in relation to mental space and time. But the question is, what does this task really measure and what are some of its task demands? Does it require "conscious recollection" or explicit knowledge, in that participants have to actually "think about" the relationship between *canaries and birds*? Or is the task automatic enough that the connection is made implicitly without having to think about the semantic relationship? It would not be surprising, for example, to find out that amnesics experience the same memory difficulties with these tasks, as they do with performance on conceptually driven tasks that require conscious recollection during the retrieval process (see for example, Baddeley, 1990). Perhaps, a much better task would be an implementation of Swinney and Cutler's (1979) *phrase classification task* in which participants determine if a string of words (e.g., *The canaries are birds*) is natural or unnatural (e.g., *The are bird canaries*), and the task would be more automatic and tacit. Such a variant would not require explicit knowledge and recollection. However, we are not aware of any studies conducted across tasks to determine the effect of task demands on the sentence verification task. In our view, there are too many unknowns about this task to make it readily adaptable to studying bilingual conceptual categories with bilingual sentences such as *Los perros son ANIMALS* (note: in a real experiment, *SON* would be excluded because of its homographic nature to English as in *son-daughter*) "Dogs are animals," or *Los pájaros tienen SKIN* "Birds have skin."

Overall, we take the general view that cognitive tasks, with some exceptions, as pointed out above, are language-free and that they could be adopted to study both bilingual or monolingual processes, as it is the case in models of reading integration (Chap. 7), connectionism (Chap. 9), and the neurobiology of bilingualism (Chaps. 10–12). It would be difficult to argue about the "monolingualism" or "bilingualism" of a task, especially when the task is intended to measure a linguistic or cognitive process, which we view as universal. As argued by Raney and Bovee (Chap. 7), *…there are no uniquely bilingual research methodologies. Instead, we take the perspective that there are research methodologies for exploring cognitive processes that can be more or less easily applied to study bilingualism. If one assumes unique methodologies must be used for studying bilinguals, one is assuming bilingual and non-bilingual cognitive processes are also unique.* However, as argued in some of the chapters, it is critical that some of these tasks are studied to determine the particular processes they measure (see Chap. 5), and how some tasks are different than others, as well as how reliable they are. As we pointed out in the introduction, the task matters, and what happens or how the experiment is designed (i.e., demand characteristics) are crucial in obtaining valid and precise data. As Altarriba and Basnight-Brown (2007) have pointed out in their review article, for example, in relation to the priming literature, some of the variables that need to be accounted for include stimulus onset asynchrony (SOA), relatedness proportion (RP), nonword ratio (NR), word length and frequency, language proficiency, and cognate status (see also Chaps. 7, 8, and 10 for excellent discussions about other possible linguistic variables influencing experimental outcomes). To summarize, psycholinguistic techniques can be viewed as "language-free" and could be adapted to any language or combination of languages to inform theory and processing, in general.

Scope of This Volume

Methods in Bilingual Reading Comprehension Research consists of 12 chapters. Chapters 2–8 look at the behavioral and the more traditional, reaction time-dependent tasks, as well as the eye-tracking methodology (Chaps. 2 and 8). Chapter 9 provides a much needed view of bilingual reading in the framework of connectionist modeling. Chapters 10–12 address the neuropsychology of bilingual reading.

Van Assche, Duyck, and Hartsuiker (Chap. 2) discuss how the choice of a research technique, the experimental task itself, and the stimuli used in the study (presented in context or out of context) may differentially affect the results and subsequently lead to diverging views concerning models of bilingual visual word recognition. Lexical decision, naming, and translation tasks are critically discussed, as well as the eye-tracking paradigm. In a similar vein, Fernández and de Souza (Chap. 3) focus on the distinction between off-line and on-line (i.e., measuring language processing in real time) methodologies, and discuss the importance of asking the right question, and choosing the appropriate empirical paradigm to answer the particular question. Accordingly, different research questions (i.e., lexical level processing vs. how bilingual parsing strategies work, for example) require different methodological techniques. Fernández and de Souza argue for incorporating the notions of linguistic *competence* (i.e., capacity) and *performance* (ability to produce) into the interpretation of the results stemming from on-line and off-line tasks, and show how the strict divide between the two is impossible to maintain. The chapter also discusses speeded vs. unspeeded tasks and stresses the necessity of implementing an incremental approach to studying bilingualism, as no single empirical paradigm can successfully address all the questions in all the possible bilingual contexts. Martin and Altarriba (Chap. 4) focus specifically on the RSVP task. While the technique has been so far mostly used to explore monolingual processing and such issues as lexical access, repetition blindness, attentional blink, or executive control in monolinguals, the authors argue for extending its use to bilingual populations on account of its usefulness for exploring reading and attention in bilinguals. As in the previous chapter, also here, the necessity of taking into account factors related to experimental design (such as stimulus type, target word type, modality of the required response, or language blocking during stimulus presentation) is strongly emphasized in determining the suitability of one experimental technique over the other.

Heredia et al. (Chap. 5, see also, Chap. 3) look at the self-paced VMW and its variants, such as the *stationary moving window* (SMW) and *auditory moving window* (AMW), as well as reviewing bilingual studies that have employed this technique to investigate bilingual processing of code-switched or mixed sentences. Bilingual studies focusing on cross-language effects exploring grammatical gender, and the processing of homographs and cognates that employed the VMW technique are discussed, as well. From the discussion of the technique's strengths and weaknesses provided in the chapter, the VMW emerges as a powerful tool to investigate bilingual reading processes. Cieślicka and Heredia (Chap. 6; see also Chap. 10) provide a critical analysis of the strengths and weaknesses of the *cross-modal lexical priming technique* (CMLP), which has been used extensively in bilingual processing studies. In particular, they focus on research which has looked into multiple language activation in the course of bilingual lexical processing and describe a set of their own studies investigating effects of context in connected speech with the use of the variant of the CMLP technique known as the *cross-modal naming task*. As in the previous chapters, also here the role of task demands and task-specific instructions is emphasized as crucial in influencing bilingual lexical processing. Raney and Bovee (Chap. 7) provide a comprehensive overview of research methodologies that have been employed to examine both the surface (i.e., word or lexical) and higher levels (i.e., meaning-based and conceptual-based) of the representation of languages in the bilingual mind. More specifically, the focus of the chapter is on those research methods that have been used to address the question concerning the degree to which lexical entries are linked across languages and whether meaning and memories for text are integrated in the bilingual mind. Throughout their discussion, the authors stress the importance of controlling for many variables that can affect experimental results, such as participants' language background and their proficiency level, frequency and length of experimental stimuli, the order in which reading passages are presented, as well as context and task demands.

Chapter 8 (Whitford, Pivneva, & Titone) elaborates on the eye-tracking methodology and provides a description of its historical background, major processing assumptions, and measures most typically reported in eye-tracking studies. The authors discuss strengths and weaknesses of the eye-tracking paradigm and stress the importance of carefully selecting language materials to be used in the experimental design. Factors such as word length, frequency, predictability, or orthographic neighborhood are all likely to significantly affect eye reading measures, and so should be controlled for. A number of studies employing the eye-tracking methodology to investigate cognitive processes involved in bilingual reading are carefully examined. The overall conclusion of the chapter is that, with properly controlled linguistic materials and experimental variables, the eye-tracking technique is a very powerful and sensitive measure able to accurately account for bilingual processing at both sentence and discourse levels. Chapter 9, by Holman and Spivey examines the contribution of the connectionist modeling tradition and the neural network approach to addressing questions regarding the architecture of the bilingual mind and processes underlying bilingual reading comprehension. A fascinating characteristic of connectionist models is that they respect the physical constraints governing neuronal networks, and thus offer the closest possible approximation to reflecting the neural processes that underlie bilingual reading comprehension. The chapter first looks at the earliest connectionist monolingual models, such as the *Interactive Activation Model*, and then goes on to discuss its bilingual extensions, such as the *Bilingual Interactive Activation* (BIA and BIA+) and other connectionist formulations of bilingual word recognition. What emerges from the chapter is that connectionist models, rather than providing an opposing alternative to more traditional box-and-arrow bilingual word recognition models, can actually be seen as complementing them and coexisting with them, with each class of models driving different processing predictions and addressing different research questions.

The last three chapters focus primarily on the neuropsychological approaches to studying bilingual reading. Hillert and Nakano (Chap. 10) look at psycholinguistic (such as *probe recognition*, CMLP, self-paced reading, and *plausibility judgment tasks*) behavioral research techniques, as well as the latest electrophysiological ERPS, magnetophysiological (e.g., *magnetoencephalography*), and hemodynamic (e.g., *magnetic resonance imagining* [MRI] *and functional magnetic resonance imaging* [fMRI]) measures used to investigate second language sentence processing. Strengths and weaknesses of each technique are elaborated on, and selected studies are reviewed. While behavioral measures can address many questions regarding sentence processing, the latest advancements in the neurocognitive domain research techniques, such as electrophysiological and hemodynamic studies, have offered a fascinating insight into neural correlates of bilingual processing. Duñabeitia et al. (Chap. 11) first elaborate on the EEG technique itself and its many advantages, such as excellent temporal resolution and the ability to capture language processing in real time, and then offer an overview of the bilingual studies that have employed ERPS. More specifically, they discuss studies into the processing of individual words, both in a single-language context, where words belong to only one of the bilingual's two languages, and in a dual-language context, where words from both languages are presented. In addition, ERP studies investigating semantic, syntactic, and morphological processes during bilingual sentence reading are critically examined as well. Factors such as age of acquisition, language proficiency, and L1–L2 similarity are discussed as relevant in explaining potential inconsistencies in the ERP studies examining bilingual sentence reading. Finally, Chap. 12 by Blackburn reviews the most recent developments in bilingual reading research using such brain viewing techniques as MRI, fMRI, *functional connectivity analysis*, and *diffusion tensor imaging* (DTI). A case for combining various methodologies is made, such as electrophysiological with hemodynamic and behavioral, as findings from neuroimaging studies seem to converge with findings from behavioral measures. In addition, neural networks involved in bilingual processing and identified through neuroimaging techniques can be explained within the framework of the existing psycholinguistic models. A number of neuroimaging studies are reviewed that have investigated such aspects of bilingual sentence processing as the neural overlap of the bilingual's first and second languages, as well as the processing of orthographic, semantic, and syntactic levels of information in the course of bilingual reading.

List of Keywords

Bilingual mode, Bilingual processing, Conscious recollection, Cross-language lexical decision task, Cross-language lexical priming, Cued translation task, Event-related potentials (ERPs), Eye-tracking, Facilitation effect, Language-free, Lexical decision task, Nonword ratio, Stimulus onset asynchrony (SOA), Phrase classification task, Priming, Relatedness proportion, Sentence verification task, Translation lexical decision task, Translation naming task, translation priming effect, Translation task, Translation-recognition task

Related Internet Sites

Eyewriter: http://www.eyewriter.org/
Gaze Analyze: http://gazealyze.sourceforge.net/
OGAMA: Open Gaze/Mouse Analyzer: http://www.ogama.net/
openEyes open source hardware and software: http://thirtysixthspan.com/openEyes/
Pupil mobile eye-tracking platform: http://pupil-labs.com/pupil/
Starburst eye-tracking software: http://thirtysixthspan.com/openEyes/software.html

Suggested Further Reading

Chin, N. B., & Wigglesworth, G. (2007). Bilingualism: *An advanced resource book. London: Routledge.*

Mackey, A., & Gass, S. M. (2012) *Research methods in second language acquisition: A practical guide*. New York: Wiley-Blackwell.

Marinis, T. (2003). Psycholinguistic techniques in second language research. *Second Language Research*, 19, 144–161.

References

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic and translation priming experiments across languages. *Behavior Research Methods, 39*, 1–18.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Altarriba, J., & Soltano, E. G. (1996). Repetition blindness and bilingual memory: Token individuation for translation equivalents. *Memory & Cognition, 24*, 700–701.

Baddeley, A. (1990). *Human memory: Theory and practice* (Rev. ed.). Needham Heights, MA: Allyn & Bacon.

De Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 1001–1018.

Garcia-Marques, L., Garcia-Marques, T., & Brauer, M. (2014). Buy three but get only two: The smallest effect in a 2×2 ANOVA is always uninterpretable. *Psychonomic Bulletin & Review, 12*, 1415–1430.

Grosjean, F. & Frauenfelder, U. H. (Eds.). (1996). *A guide to spoken word recognition paradigms (Special issue). Language and Cognitive Processes*, *11*(6), 553–558.

Heredia, R. R., & Cieślicka, A. B. (2014). Bilingual storage: Compound-coordinate and derivates. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of bilingual memory* (pp. 11–40). New York: Springer.

Heredia, R. R., & Stewart, M. T. (2002). On-line methods in spoken language research. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 7–28). Amsterdam: Elsevier Science.

Kieras, D. E., & Just, M. A. (Eds.). (1984). *New methods in reading comprehension research*. Hillsdale, NJ: Erlbaum.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6*, 312–319.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*(9), 1043–1056.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18*, 523–534.

Wei, L., & Moyer, M. G. (Eds.). (2008). *The Blackwell guide to research methods in bilingualism and multilingualism*. Malden, MA: Blackwell.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparison of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research, 41*, 105–128.

Chapter 2
Context Effects in Bilingual Sentence Processing: Task Specificity

Eva Van Assche, Wouter Duyck, and Robert J. Hartsuiker

Abstract This chapter provides an overview of bilingualism research on visual word recognition in sentence context and relates this work to task-specific context factors. Many studies examining bilingual word recognition out-of-context have shown that words from both languages become activated when reading in one language (i.e., language-nonselective lexical access). A recent research line investigated whether presentation of words in a sentence context, providing a language cue and/or semantic constraint to restrict lexical access to words in the target language, modulates this language-nonselective activation. Recent *lexical decision*, *translation*, *naming*, and *eye-tracking* studies suggest that the language of the sentence context cannot restrict lexical access to words of the target language. Eye-tracking studies revealed that semantic constraint of a sentence does not necessarily restrict language-nonselective access, although there is evidence that it has a relatively late effect, and that it affects language-nonselective activation in lexical decision, translation, and naming studies.

Introduction

A fundamental issue in the domain of bilingualism concerns the organization of the bilingual language system. One viewpoint is that bilinguals have two separate lexicons that can be accessed selectively so that they can effectively function like monolinguals. Another viewpoint is that they have an integrated lexicon containing all words in both languages that can be accessed in a language-nonselective way. In the last decade, more and more researchers have provided evidence for this latter view.

E. Van Assche, Ph.D. (✉) • W. Duyck, Ph.D. • R.J. Hartsuiker, Ph.D.
Department of Experimental Psychology, Ghent University,
Henry Dunantlaan 2, Ghent 9000, Belgium
e-mail: eva.vanassche@ugent.be; Wouter.Duyck@ugent.be; Robert.Hartsuiker@ugent.be

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_2

It has become clear that lexical representations of the first language (L1) are accessed when bilinguals are reading single words in their second language (L2; Dijkstra, Grainger, & Van Heuven, 1999; Duyck, 2005; Jared & Kroll, 2001; Lemhöfer & Dijkstra, 2004) and vice versa (e.g., Duyck, 2005; Van Hell & Dijkstra, 2002). Only recently has this question been addressed in relation to how context and the semantic constraint provided by a sentence might modulate this language-nonselective activation for single word reading (e.g., Schwartz & Kroll, 2006; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009; Van Hell & De Groot, 2008). In the present chapter, we provide an overview of single-word and sentence processing studies on bilingual visual word recognition and discuss how task characteristics might modulate the results.

Bilingual Visual Word Recognition Out-of-Context

To investigate whether bilinguals activate words in both languages or only in the contextually relevant language when reading, the processing of words that are similar across languages is often compared to the processing of language unambiguous words. For instance, *cognates* are translation equivalents with a similar or equal spelling across languages (e.g., Spanish-English *papel-paper*). These words are typically read faster than noncognates that have no orthographic overlap across languages (e.g., *silla-chair*). This *cognate facilitation effect* is typically explained by assuming language-nonselective activation in which words from both languages are activated in parallel (e.g., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Dijkstra & Van Heuven, 2002). The presentation of a word in one language co-activates orthographically, phonologically, and semantically similar words in the other language. Since cognates share orthographic, phonological, and semantic information across languages, whereas noncognates only share semantic information, the convergent cross-lingual activation of these representations speeds up the recognition of cognates as compared to noncognates. The cognate facilitation effect has been consistently found for word reading in L2 (e.g., Caramazza & Brones, 1979; Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004; Lemhöfer, Dijkstra, & Michel, 2004) and even for word reading in L1 (e.g., Van Hell & Dijkstra, 2002).

Van Hell and Dijkstra (2002) used an L1 (Dutch) *lexical decision task* to investigate whether knowledge of a second language influences native-language reading. In this task, which is the most frequently used experimental task to study cross-lingual interactions in bilingual processing, bilinguals see letter strings on a computer screen and they have to decide as quickly and as accurately as possible whether a presented letter string is a real word (e.g., *blouse*) or not (e.g., *flouse*), in English, for example. Van Hell and Dijkstra tested reading in L1 in two groups of Dutch-English-French trilinguals: one group was highly proficient in English and relatively low in their proficiency in French, and the other group was highly proficient in both English and French. The stimuli were L1–L2 cognates (e.g., Dutch *hamer*: *hammer* in English: *marteau* in French), L1–L3 cognates (e.g., Dutch *citroen*: *lemon* in

English; *citron* in French) or matched control words (e.g., Dutch *kelder*: *basement* in English: *cave* in French). These three groups of Dutch (L1) words were matched for word length, word frequency, and number of *orthographic neighbors* (i.e., words differing by a single letter from the target such as *snow*, an intralingual neighbor of *slow*) in Dutch (Coltheart, Davelaar, Jonasson, & Besner, 1977) because these factors have been shown to significantly influence word processing (e.g., New, Ferrand, Pallier, & Brysbaert, 2006; Segui & Grainger, 1990). This matching ensured that any observed differences in processing between cognates and noncognates could be attributed to the difference in cross-lingual overlap between cognates and noncognates and not to any uncontrolled stimulus characteristics. For both groups of trilinguals, the results showed a cognate facilitation effect for L1–L2 cognates. There was also an effect for L1–L3 cognates, but only for the participants who were highly proficient in both English (L2) and French (L3). Apparently, the occurrence of cross-lingual activation in L1 reading required a certain level of proficiency in L2 and L3. Nevertheless, these results show that even a second or third language gets activated strongly enough to influence native-language word processing.

Other evidence for dual-language activation comes from studies investigating the recognition of *interlingual homographs* (i.e., words that have the same orthographic form in both languages but have a different meaning; e.g., English *red*: *net* in Spanish). However, in contrast to the consistent replication of cognate facilitation effects, studies using homographs have yielded mixed effects depending on task requirements, stimulus list composition, and relative frequency of the homographs in the two languages (e.g., Dijkstra et al., 1999; Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998). Dijkstra et al. (1998) tested Dutch-English bilinguals in three lexical decision experiments. In Experiment 1, they performed an L2 (English) lexical decision task including interlingual homographs (e.g., English *room*: *cream* in Dutch) and monolingual control words that have no cross-lingual overlap (e.g., Dutch-English *stoel-chair*). Reaction times to homographs did not differ from monolingual controls suggesting that the Dutch reading of the homograph did not influence English word recognition. However, in Experiment 2, Dutch nonhomographic filler words were included in the English lexical decision task. Participants were instructed to respond with "yes" to English words (homographs and controls) and to respond with "no" to non-English words (Dutch words and nonwords). In this experiment, homographs were responded to more slowly than monolingual controls. The presence of Dutch words as nonwords might have boosted activation in the L1 lexicon, leading to stronger interference effects for homographs. In Experiment 3, the same stimuli were presented in a generalized lexical decision task in which a "yes" response had to be given to a word in either language. Under these task requirements, homographs were processed faster than monolingual control words, indicating that the fact that a homograph is a word in both languages speeds up its reaction time relative to controls. It seems that bilinguals then react to the fastest available representation in either language, leading to faster processing of homographs as compared to monolingual control words. Dijkstra et al. (1998) nicely illustrated that cross-lingual interactions and word reading can differ depending on task characteristics.

In another study, Dijkstra et al. (1999) showed that the null results for interlingual homographs in Dijkstra et al.'s (1998) Experiment 1 could be clarified by distinguishing the orthographic and phonological overlap components of the homographs. Interlingual homographs have the same orthography in both languages but they can differ in the degree of phonological overlap. To investigate the effect of phonological overlap in interlingual homograph processing, they included English homographs that were either pronounced very similarly to Dutch words (e.g., English *pet*; *cap* in Dutch), or that were pronounced very differently to Dutch words (e.g., English *glad*; *slippery* in Dutch). The results showed that the processing of homographs with no phonological overlap was facilitated. This suggested that the facilitative influence of orthographic overlap and the inhibitory influence of phonological overlap led to the null effects in Dijkstra et al. (1998) and highlighted the importance of controlling for phonological similarity in the homograph stimuli.

To summarize, these studies on cognate and homograph word processing indicate that bilinguals cannot effectively function like monolinguals and that both languages interact and influence word recognition. Ever since, many studies have provided evidence for the viewpoint of language-nonselective activation of words in the two languages (e.g., Dijkstra et al., 2010; Dijkstra & Van Heuven, 2002; Van Assche, Duyck, & Hartsuiker, 2012). Does this language-nonselective activation similarly apply for bilinguals reading in context? Recently, studies began to test the ecological validity of the experiments presenting words out-of-context. After all, people rarely read lists of isolated words, but instead, words are embedded in meaningful sentences. It is possible that the presentation of words in a sentence context restricts lexical activation to words of the target (sentence) language only or allows for earlier language selection during lexical access. Indeed, studies in the monolingual domain have shown that semantic and syntactic restrictions imposed by a sentence context are used to speed up recognition of upcoming words (e.g., Schwanenflugel & LaCount, 1988; Stanovich & West, 1983). The question now is whether these monolingual sentence context effects generalize to bilingual sentence processing. Altarriba, Kroll, Sholl, and Rayner (1996) were the first to investigate word recognition in mixed-language sentences; a number of other studies investigating more natural unilingual sentence reading were carried out more recently (e.g., Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Schwartz & Kroll, 2006; Van Assche et al., 2009; Van Hell & De Groot, 2008).

Bilingual Visual Word Recognition in Sentences

Studies of bilingual word recognition in sentences have used several different tasks to investigate how sentence context might modulate the cross-lingual activation effects observed in single-word studies. Different tasks may tap into different processes and may consequently lead to different result patterns related to the time course of word processing. In the following sections, we first discuss research using tasks requiring an overt response such as lexical decision or naming tasks, before presenting research using more natural reading tasks such as *eye-tracking*.

Reading Tasks Requiring Overt Responses

One of the first investigations on bilingual word recognition in sentences is presented in Van Hell and De Groot (2008). Proficient Dutch-English bilinguals performed an L2 (English) lexical decision task or a *translation task* in forward (from L1 to L2) or in backward direction (from L2 to L1). The critical stimuli were cognates (e.g., Dutch-English: *kapitein-captain*) and matched control words (e.g., *rok-skirt*) presented out-of-context or preceded by a low- or high-constraint L2 sentence context. The sentence constraint manipulation allowed testing whether merely presenting words in a low-constraint sentence is sufficient to restrict lexical access to words of the target language or whether only a semantically constraining sentence can direct lexical access to the language of the sentence. Sentence completion ratings and plausibility ratings ensured that mean production probabilities did not differ for cognates and controls.

Experiment 1 investigated whether a meaningful sentence context and a semantically constraining sentence can guide lexical access to words of the target language and modulate the cognate facilitation effect. Dutch-English bilinguals were presented with an L2 (English) sentence context in which the target word was omitted (e.g., *A green ____and a yellow banana lay on the fruit dish*). After the sentence context disappeared from the computer screen, the target (e.g., *apple*) for L2 lexical decision was presented. A control condition also presented target words out-of-context. The results showed that cognates were processed faster than controls when presented out-of-context. Cognate facilitation remained after reading a low-constraint sentence context, but not after a high-constraint sentence context. This finding suggests that the semantic constraint of a sentence, but not the linguistic context and language cue provided by a sentence, can restrict cross-lingual activation effects.

Experiments 2 and 3 investigated how contextual information influences the translation of words. Sentences were presented as a whole with the target omitted (Experiment 2) or were presented word-by-word (Experiment 3). The target's translation had to be spoken out loud. In both experiments, results were comparable across the two translation directions (from L1 to L2 and from L2 to L1): cognate facilitation effects observed out-of-context remained in the presence of a low-constraint sentence, but were strongly reduced when the words were presented after a semantically constraining sentence. The results of the lexical decision and translation tasks suggest that the feature restrictions imposed by a high-constraint, but not a low-constraint sentence, delineate the lexical and conceptual information of the upcoming words.

Schwartz and Kroll (2006) observed similar cognate results for word production in two sentence context experiments with highly proficient Spanish-English bilinguals living in a bilingual community and intermediate proficient Spanish-English bilinguals living in a monolingual community. They presented English-Spanish cognates (e.g., *piano*), interlingual homographs (e.g., *fin*), and monolingual control words in L2 (English) low- and high-constraint sentences. The sentences were presented word by word using a rapid *serial visual presentation task* (RSVP) and participants had to name the target word (printed in red) as quickly as possible (e.g., high-constraint cognate sentence: *Before playing*, *the composer first wiped the keys of the piano at the beginning of the concert*; high-constraint control sentence: *Before the test*, *the student looked for some paper and a sharp pencil to write with*). Cognate facilitation was observed in low-constraint sentences, but not in high-constraint ones. No reaction time differences were found for homographs and controls in either low- or high-constraint sentences, but bilinguals of intermediate proficiency made more errors than highly proficient ones, especially in low-constraint sentences. Although the results for homographs were somewhat inconclusive, the results for cognate processing show that the semantic constraint of a sentence can restrict cross-lingual activation effects for both intermediate and highly proficient bilinguals.

A semantic priming study by Elston-Güttler, Gunter, and Kotz (2005) also investigated homograph processing in an L2 sentence context and tested how a more general language context may influence cross-lingual activation effects. To this end, German-English bilinguals saw either a German or English movie prior to the experiment, boosting L1 or L2 activation. Additionally, Elston-Güttler et al. tested how these language context effects change over time by analyzing the first and second halves of the experiment. German-English homographs (e.g., *gift*: *poison* in German) or control words (e.g., *shell*) were presented at the end of a relatively low constraining sentence (e.g., *The woman gave her friend a pretty gift* vs. *The woman gave her friend a pretty shell*). The sentence was then replaced by a target word for L2 (English) lexical decision (*poison*). If the L1 influences the L2 during word recognition, reading the homograph *gift* should influence subsequent processing of the related word *poison*. Targets were recognized faster after the related homograph sentence than after the unrelated control sentence, but only in the first half of the experiment and only for participants who saw a German movie prior to the experiment, boosting L1 activation. This *semantic priming effect* of the targets and their related homographs was also present in the recordings of *event-related potentials* in the modulation of the N200 and N400 components. The N200 component has been linked to word access and/or orthographic processing (e.g., Bentin, Mouchetant-Rostaing, Giard, Echalier, & Pernier, 1999). Elston-Güttler et al. suggested a translational word form link between *gift*-*poison* so that reading the prime *gift* leads to faster lexical access of the target *poison*. The N400 component has been linked to semantic integration processes (Brown & Hagoort, 1993) suggesting that the target *poison* was easier to integrate and resulted in less negative N400 amplitude after the related prime *gift* than after the unrelated prime *shell*. This study showed that sentence context can eliminate the activation of the nontarget L1 homograph representation and that this effect is very sensitive to language context. Semantic priming effects were only observed when L1 activation was boosted prior to the experiment and only in the first half of the experiment. This suggests that context effects such as activating an L1 or L2 prior to the experiment can influence word recognition and that the bilingual language system can quickly zoom into the L2 processing context.

These homograph studies did not distinguish word class overlap, and Baten, Hofman, and Loeys (2011) reasoned that this variable might interact with how a sentence context influences cross-lingual activation for homographs. They explicitly distinguished between interlingual homographs sharing the same word class and those that do not. For example, *angel* has the same word class as the Dutch reading's meaning *sting*, whereas *breed* is a verb or a noun and has a different word class from the Dutch reading's meaning *wide*. Baten et al. reasoned that the influence of word class might be particularly important when presenting words in a sentence context in which word meaning and sentence constraint interact. Dutch-English bilinguals performed an L2 (English) lexical decision task to target words appearing as final words in a sentence. Both the homograph and its control word could appear in the same low-constraint sentence (e.g., *She looked up and there seemed to be an angel/alien*: where *angel* is the homograph and *alien* is the control). Reaction times for homographs were faster than for controls when the two readings of the interlingual homograph had the same word class, but this homograph facilitation was eliminated when there was no such word class overlap. So the overlap in orthographic representation for homographs only led to faster processing times in sentence context for homographs that had the same word class. These results suggest that the presence of a sentence context indicating the word class of upcoming words can have a direct impact on cross-lingual activation. That is, only orthographic overlap for homographs did not lead to cross-lingual activation effects in a sentence context, but orthographic and word class overlap did.

Overall, these studies using tasks that require overt responses converge on the conclusion that the degree of language-nonselective activation is influenced by the semantic constraint of a sentence because a high-constraint, but not a low-constraint sentence context affected lexical access in bilinguals. However, even though the presentation of words in a sentence provides a more natural reading situation than word recognition out-of-context, the procedure still requires a response from the participant (e.g., a word/nonword response in lexical decision), which is not necessary in natural reading. Moreover, studies often presented target words at the end of a sentence context (e.g., Baten et al., 2011; Elston-Güttler et al., 2005). Sentence-final words are typically read more slowly than sentence-internal words (e.g., Just & Carpenter, 1980), and this sentence wrap-up effect has traditionally been explained by integrative processing that occurs at the end of sentences (e.g., Rayner, Kambe, & Duffy, 2000). These processes might interfere with cross-lingual activation processes. Recent studies therefore have used the eye-tracking paradigm, in which participants can read normally as in everyday life, and no overt task other than comprehension is required. The time-sensitive eye movement measures allow researchers to investigate the time course of lexical activation by dissociating several early (reflecting initial lexical access) and late reading time measures (reflecting higher-order processes; Rayner, 1998). Indeed, eye-tracking studies in the monolingual domain (e.g., Duffy, Kambe, & Rayner, 2001; Rayner, Binder, & Duffy, 1999) suggest that the degree of competition between multiple meanings of an ambiguous word (e.g., *bank* as a financial institution or as a *river side*) depends on the relative time course of their activation. The time course of meaning activation, in turn, is determined by the relative frequencies of the ambiguous word's meaning, and this activation can be modulated by a biasing context (e.g., Duffy et al., 2001). In the next section, we discuss how the use of eye movements has deepened our knowledge of sentence context effects on bilingual visual word recognition in L2 and in L1.

Natural Reading and Eye-Tracking

L2 Processing

Duyck et al. (2007) investigated how the linguistic context provided by a sentence can constrain language-nonselective access in normal second language reading while measuring eye movements. The cognate facilitation effect was used as a marker of cross-lingual interactions because cognate effects have been shown to be strong and reliable in out-of-context studies (e.g., Dijkstra et al., 1999; Van Hell & Dijkstra, 2002). Duyck et al. selected both identical (e.g., Dutch-English: *ring*) and nonidentical cognates (e.g., *schip*-*ship*) to examine how cross-linguistic overlap between translation equivalents may interact with the cognate effect in sentence context. In Experiment 1, the L2 cognate facilitation effect, as found in earlier single-word studies (e.g., Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004), was replicated. Proficient Dutch-English bilinguals were presented with cognates, matched control words, filler words, and nonwords in an L2 (English) lexical decision task out-of-context. Reaction times were faster for cognates than for controls, and this effect interacted with the degree of cross-linguistic overlap: cognate facilitation was stronger for identical than for nonidentical cognates. This experiment validated the stimulus materials for use in the sentence studies.

In the second experiment, the same cognates and controls were presented as the final words of low-constraint sentences (presented word by word using RSVP), to which an L2 lexical decision had to be made. Both the cognate and its matched noncognate fit the same sentence (e.g., *Hilda was showing off her new ring/coat*; *ring* is the cognate; *coat* is the control). Similar to the presentation out-of-context, a cognate facilitation effect was obtained, and this effect was stronger for identical than for nonidentical cognates. This finding again shows that the unilingual linguistic context provided by a sentence does not eliminate cross-lingual interactions (cf. Schwartz & Kroll, 2006; Van Hell & De Groot, 2008).

The third experiment presented the cognates and controls in the middle of L2 low-constraint sentences while measuring eye movements. The eye-tracking technique can investigate reading in its most natural way in a laboratory situation and can distinguish between several early and late reading time measures. Early measures typically include *first fixation duration* (i.e., the duration of the first fixation on the target word) and *gaze duration* (i.e., the sum of fixations from the moment the eyes land on the target until they move off again). Late measures typically include *go-past time* (i.e., the time elapsing from encountering a target for the first time until a region to the right of the target is fixated), which also takes into account regressions originating from the target. The results showed cognate facilitation effects on the reading times for identical cognates on first fixations from 249 ms onwards after first encountering the word and also on later go-past time. Such cognate facilitation was not present for nonidentical cognates. This result indicates that the amount of cross-lingual activation is a function of the similarity between the translation equivalents. A sentence context providing a language cue might eliminate L2 cognate effects when cross-lingual activation is weaker (i.e., nonidentical cognates), but not when overlap is complete (i.e., identical cognates). Furthermore, the fact that nonidentical cognate effects in low-constraint sentences were observed in lexical decision but not in normal reading, as measured via eye movements, indicates that context and lexical variables (i.e., the degree of cross-lingual overlap between translation equivalents) may also interact with task-specific factors.

Libben and Titone (2009) later showed that even in a high-constraint sentence context, lexical activation is initially language-nonselective, although previous studies using tasks requiring overt responses (e.g., lexical decision; see for example, Schwartz & Kroll, 2006; Van Hell & De Groot, 2008) suggested that a semantically constraining sentence can constrain lexical selection to the target language. They presented French-English identical cognates (e.g., *divorce*) and interlingual homographs (e.g., *chat*: *cat* in French) in L2 (English) sentences that were either low or high in terms of semantic constraint for the target (e.g., high-constraint homograph sentence: *Since they like to gossip*, *they had an extended chat that lasted all night*; control sentence: *Since they liked to compose songs*, *he made an extended tune that was very catchy*). Highly proficient French-English bilinguals read the sentences while eye movements were measured. The results of early reading time measures (e.g., first fixation, gaze duration) revealed that homographs were read more slowly than matched controls in both low- and high-constraint sentences. Cognate facilitation was present on early reading time measures in low- and high-constraint sentences.

Thus, lexical access was nonselective and not modulated by semantic constraint in approximately the first 350 ms upon fixating the word. However, in the time range of approximately 350–600 ms of later reading time measures (e.g., go-past time), cognate facilitation and homograph inhibition was still present in low-constraint sentences, but not any longer in high-constraint sentences. Libben and Titone (2009) suggested that lexical access is initially language-nonselective, but that this cross-language activation is nullified by top-down factors such as semantic constraint of a sentence at later word processing stages. The absence of cognate facilitation in high-constraint sentences for later stage results (e.g., go-past time) is consistent with the lexical decision results of Schwartz and Kroll (2006) and naming results of Van Hell and De Groot (2008) and suggests that these tasks may reflect comprehension processes occurring after lexical access had taken place. Furthermore, Libben and Titone also suggested that the absence of homograph interference effects in Schwartz and Kroll (2006) may be related to task characteristics. RSVP and word naming may be less sensitive than eye-tracking to detect cross-lingual interference effects.

The absence of cross-lingual activation effects at later stages of comprehension in semantically constraining sentences reported in Libben and Titone (2009) contrasts with the results of Van Assche, Drieghe, Duyck, Welvaert, and Hartsuiker (2011), who did observe dual-language activation on late eye movement measures. This difference may originate from the fact that the Dutch-English bilinguals in Van Assche et al. were less balanced than the bilinguals tested in Libben and Titone. This may lead to stronger L1 activation in the bilinguals tested in Van Assche et al. (2011). Indeed, Titone, Libben, Mercier, Whitford, and Pivneva (2011) suggested that the bilinguals in Van Assche et al. may have experienced greater L1-to-L2 cross-language activation, so that semantic context may be insufficient to diminish cross-language activation.

The Dutch-English bilinguals in Van Assche et al. read cognates and matched control words in low and high semantically constrained sentences in their L2 while eye movements were recorded (e.g., low-constraint cognate sentence: *He went to the shop to buy a book that he needed for school*; low-constraint control sentence: *She did not want to look at her face while she was crying*). Cognate facilitation was shown on early and late eye movement measures, both for low- and high-constraint sentences. Moreover, facilitation increased gradually as a function of cross-lingual overlap between translation equivalents: higher orthographic overlap between translation equivalents on Van Orden's (1987) word similarity measure for cognates and controls led to faster reading times. These results indicate that semantic constraint does not affect cross-lingual activation in the bilingual language system at any stage of word recognition.

The cognate eye-tracking results in semantically constraining sentences in Van Assche et al. (2011) and Libben and Titone (2009) contrast with the previous studies of Schwartz and Kroll (2006) and Van Hell and De Groot (2008) who observed no cognate facilitation on lexical decision and naming times in high-constraint sentences. Van Assche et al. tested whether this difference between studies may be related to the different methodology used in an additional experiment, in which the stimulus materials of Van Assche et al. were presented using the paradigm of Van Hell and De Groot. They observed a weak cognate facilitation effect in high-constraint sentences, and this effect only emerged after running many more participants than did Van Hell and De Groot. These findings illustrate that the eye-tracking paradigm may be more sensitive to detecting cross-lingual activation effects than tasks requiring overt responses.

Balling (2012) recently tested an even more natural reading situation than word recognition in sentences. She had Danish-English bilinguals read cognates in texts or paragraphs. Cognate facilitation was observed that was modulated by morphological complexity. There was cognate facilitation for simple cognates (e.g., Danish-English *rolle-role*). This observation extends the evidence for language-nonselective access for word recognition in sentences to reading in texts. There was also an inhibitory effect for complex cognates (e.g., words that contain at least one cognate morpheme as in *onsdag-Wednesday*, where *dag-day* is the cognate morpheme). Balling suggested that problems in the integration of cognate and noncognate morphemes might lead to this inhibition.

Note that the above studies all used noun stimuli to investigate language-nonselective activation in bilinguals and that theoretical accounts of bilingual language processing (e.g., Dijkstra & Van Heuven, 2002) are almost exclusively based on noun processing. Van Assche, Duyck, and Brysbaert (2013) therefore examined lexical access for verbs during sentence reading in L2 with Dutch-English bilinguals. Although verbs have generally smaller degrees of formal and semantic overlap between languages than nouns (Gentner, 1981), there was cognate facilitation for cognate and control verbs presented out-of-context and cognate facilitation remained on a late reading time measure (go-past time) when targets were presented in low semantically constraining sentences. Early reading time measures did not show cross-lingual activation effects though. Thus, although cross-lingual activation effects for verbs were weaker than for nouns, these results show that cross-lingual activation is strong enough for verb cognate effects to arise.

L1 Processing

As in the literature on bilingual word recognition out-of-context, the majority of published sentence context studies have focused on L2 processing. Influences of L1 on L2 processing are indeed generally stronger than influences of L2 on L1 processing (e.g., Duyck, 2005; Haigh & Jared, 2007; Jared & Kroll, 2001) and so cross-lingual activation effects are more likely to be observed for L2 processing. However, in order to demonstrate the existence of a profoundly language-nonselective bilingual language system, influences of the weaker L2 on reading in the dominant language should be investigated. Van Hell and Dijkstra (2002) were the first to show cognate facilitation effects in L1, indicating that the bilingual's L2 knowledge influenced native-language reading. Van Assche et al. (2009) replicated this cognate facilitation effect for words out-of-context and then tested how language information of a sentence context may influence this cross-lingual activation effect. Dutch-English bilinguals read L1 low-constraint sentences which contained a cognate or a control word (e.g., *Bert heeft een oude oven/lade gevonden tussen de rommel op zolder*: "Bert has found an old *oven/drawer* among the rubbish in the attic"). Early reading time measures (i.e., first fixation duration) were shorter for cognates than for controls. Moreover, cognate facilitation was shown to be a continuous effect because cognate facilitation gradually increased as a function of cross-lingual similarity. The results show that the mere presentation of words in a sentence context does not restrict cross-lingual interaction effects in bilinguals during native-language reading. This indicates a limited role for top-down lexical restrictions generated by sentences on the cross-lingual activation in the bilingual lexicon.

Titone et al. (2011) investigated whether semantic constraint would modulate cross-language activation during native-language reading. In Experiment 1, they measured the eye movements of English-French bilinguals reading identical cognates (e.g., English-French: *divorce*) and interlingual homographs (e.g., c*hat*: *cat* in French) in low- and high-constraint L1 sentences (e.g., high-constraint cognate sentence: *Because of the bitter custody battle over the kids*, *the expensive divorce was* *a disaster* vs. high-constraint control sentence: *Because the maid of honor and best man were late*, *the expensive wedding was a disaster*). Cognate facilitation was present on early reading time measures in both low- and high-constraint sentences, but this effect was modulated by age of L2 acquisition: only bilinguals who acquired their L2 early in life showed cognate facilitation. Age of L2 acquisition did not modulate cognate effects on late reading time measures, but here, semantic constraint did: cognate facilitation was smaller in high- than in low-constraint sentences. There were no early homograph interference effects. Homograph interference was only present on total reading times and, contrary to the L2 results of Libben and Titone (2009), was unaffected by the semantic constraint of the sentence.

In the second experiment, L2 (French) filler sentences were intermixed with the experimental English sentences to examine whether making the L2 more salient would increase cognate and homograph effects during L1 reading. Indeed, under these experimental conditions, cognate facilitation was not reduced. The inclusion of L2 filler sentences seems to have increased cross-lingual activation during L1 sentence reading, and this process may have counteracted the semantic constraint effect. Homograph interference was present on total reading times, and this effect was stronger in Experiment 2 than in Experiment 1.

Theoretical Accounts

Although cognate and homograph effects have often been taken as evidence for an integrated lexicon in which words from both languages are represented and/or for lexical access of words from both languages in parallel, the precise representation of cognates and homographs and the modeling of task and context effects is a strongly debated topic (cf. Costa, Santesteban, & Cano, 2005; Dijkstra et al., 2010; see Van Assche et al., 2012 and Degani & Tokowicz, 2010, for reviews). A theoretical explanation of the cross-language activation effects discussed above can be given within bilingual language processing models such as the *Bilingual Interactive Activation Plus Model* (BIA+; Dijkstra & Van Heuven, 2002) and a bilingual extension of the *Re-ordered Access Model* of Duffy, Morris, and Rayner (1988; Arêas Da Luz Fontes & Schwartz, 2010; Degani & Tokowicz, 2010).

The BIA+ model (Dijkstra & Van Heuven, 2002) assumes that L1 and L2 words are represented in an integrated lexicon and that representations from both languages become activated in parallel. Lexical representations are activated depending on the overlap with the input stimulus and the resting level activation of the representations (based on frequency, proficiency, etc.). Cognates have similar orthographic and phonological representations and the same semantic representation. This high degree of similarity across languages speeds up their activation and recognition, as compared to noncognates. However, other theoretical accounts of the cognate facilitation effect assume qualitative differences in the representation of cognates and noncognates at a conceptual (e.g., De Groot & Nas, 1991) or a morphological level (e.g., Sánchez-Casas & García-Albea, 2005; see e.g., De Groot, 2011; Dijkstra et al., 2010, for an overview). Interlingual homographs, on the other hand, have different semantic representations but they have the same orthographic representations in both languages. Control words only activate representations in one language. This difference in activation levels gives rise to the homograph effect.

Another theoretical account of cross-lingual homograph effects is an extension of the monolingual Re-ordered Access Model of Duffy, Morris, and Rayner (1988; Arêas Da Luz Fontes & Schwartz, 2010; see also Degani & Tokowicz, 2010; Schwartz & Van Hell, 2012). According to this monolingual model, the extent to which each meaning of a homonym (e.g., *bank* as a riverside or a financial institution) is activated depends on the relative frequency of the meanings and on the syntactic/semantic context biasing a certain meaning. For instance, for homonyms presented without a biasing context, the relative frequency of the meanings determines the time course of their activation. A strong biasing context can reorder this activation. For the bilingual case, Arêas Da Luz Fontes and Schwartz (2010) propose that, in addition to frequency and context, cross-language activation may influence the time course of meaning activation. All three factors can interact with each other to activate the meaning of interlingual homographs in each language and therefore affect cross-lingual homograph effects.

In the BIA+ model, language membership is represented via language nodes such that all words from the same language are connected to a corresponding language node. The language nodes also reflect the global activity of each language. In the earlier BIA model (Dijkstra & Van Heuven, 1998), the language nodes could suppress the activation of words in the other language through inhibition mechanisms. Later, in the BIA+ model, the language nodes served only a representational function; they can be pre-activated by the sentence, but they cannot influence the activation of words in the other language. As such, Dijkstra and Van Heuven (2002) predicted that the mere presentation of words in a sentence does not constrain language-nonselective activation. Indeed, the fact that cross-lingual activation effects were preserved in low-constraint sentences in L2 (e.g., Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche et al., 2011; Van Hell & De Groot, 2008) and in L1 (e.g., Titone et al., 2011; Van Assche et al., 2009) provides support for the assumption of limited influence of the language of the sentence on language-nonselective activation.

Turning to the effect of semantic constraint on lexical activation, Dijkstra and Van Heuven (2002) suggested that syntactic and semantic context might directly affect the word identification system. This may change the degree of language-nonselectivity in bilingual word recognition in a similar way, as sentence context influences monolingual word recognition (e.g., Schwanenflugel & LaCount, 1988). Indeed, lexical decision and naming studies have revealed that a semantic context could constrain lexical access (e.g., Schwartz & Kroll, 2006; Van Hell & De Groot, 2008), but eye-tracking studies did not find the same results (e.g., Van Assche et al., 2011) or found an effect of semantic constraint only in later processing stages (e.g., Libben & Titone, 2009). This suggests that the semantic context effect on lexical activation may occur during later stages of word recognition, although not all studies support this suggestion (Van Assche et al., 2011).

In order to account for differences between experiments and nonlinguistic context effects (e.g., task characteristics, participant's expectations) in the BIA+ model, a distinction is made between the word identification system containing orthographic, phonological, and semantic representations and the task/decision system, analogous to Green (1998). This additional task/decision system allows distinguishing processes that influence the activation of lexical representations in the word identification system from processes that influence participants' decision criteria. Cross-experimental differences are thus handled by the task/decision system affecting the output of the word identification system. Dijkstra and Van Heuven (2002) propose that nonlinguistic information affects only the decision criteria related to task demands rather than the activation level of lexical representations in the two languages (for more information, see Dijkstra & Van Heuven, 2002).

Creating a Sentence Context Experiment

The studies presented above illustrate that the design of an experiment and its specific task choice or stimulus materials can influence the results and subsequent conclusions substantially. In this section, we discuss the procedures and points of interest for designing a sentence context experiment using the eye-tracking paradigm.

Participants

In selecting the bilinguals to take part in the experiment, it is important to carefully consider several factors that have been shown to influence cross-language activation effects in visual word and sentence processing such as L2 (and L3) proficiency (e.g., Van Hell & Dijkstra, 2002) and age of L2 acquisition (e.g., Titone et al., 2011). Bilinguals are often asked to rate their speaking, reading, writing, and comprehension abilities in each language as a measure of proficiency in language history questionnaires, even though the validity of these questionnaires has only rarely been tested. Recent studies explicitly addressed this issue, and this has led to the development of validated instruments such as the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007). There are also more direct tests of proficiency. The LexTALE is a short lexical test for advanced learners of English as an L2 and has been shown to be a good predictor of English vocabulary knowledge and general English proficiency (Lemhöfer & Broersma, 2012). A similar test to measure language proficiency in French has been developed by Brysbaert (2013). Another recent test to measure language dominance and language proficiency in spoken production is the Multilingual Naming Test (MINT; Gollan, Weisberger, Runnqvist, Montoya, & Cera, 2012). A paper-and-pencil dominance scale to quantify the language dominance of bilinguals was presented by Dunn and Fox Tree (2009).

Stimulus Materials

Selection of stimulus materials can often be a time-consuming process in language research because word and sentence stimuli have to be carefully selected and controlled on a number of factors. For instance, testing cognates or homographs in a factorial design (i.e., comparing processing of the set of cognates with the set of control words) requires the selection of control words that are matched to the cognates on word characteristics such as word length, word frequency, and number of orthographic neighbors (Coltheart et al., 1977). These factors have been shown to significantly influence word processing (e.g., New et al., 2006). The WordGen program (Duyck et al., 2004) can be used to calculate the values of these variables for selected words in Dutch, English, German, and French. It can also be used to select control words and to generate nonwords (for use in a lexical decision task) adhering to any combination of linguistic constraints such as number of letters, neighborhood size, frequency, and summated bigram frequency. Another example of a resource for psycholinguistic research is LexicALL that contains useful datasets such as Chinese, Dutch, and English word frequencies based on film and television subtitles (e.g., Cai & Brysbaert, 2010; Cuetos, Glez-Nosti, Barbon, & Brysbaert, 2011). It also includes Wuggy, which is a multilingual generator of nonwords (Keuleers & Brysbaert, 2010). More information on LexicALL and Wuggy can be found on www.crr.ugent.be (see Related Internet Sites).

Selection of stimulus materials does not necessarily have to include the selection of cognates and homographs and matched controls. Cross-lingual overlap can also be investigated as a continuous measure (see Van Assche et al., 2009, 2011). For instance, orthographic overlap of a set of stimulus materials including cognates and noncognates can be calculated using the Van Orden (1987) word similarity measure or Levenshtein distance (Levenshtein, 1966). By calculating an orthographic overlap score for each translation word pair (e.g., Dutch-English: *schouder-shoulder* has a Van Orden overlap value of .81; *leraar-teacher* of .30; for more information on the calculation of Van Orden overlap scores, see Related Internet Sites), it can be investigated whether more orthographic overlap across languages facilitates word processing. These measures do not take into account phonological overlap, and therefore, additional ratings will have to be collected.

In a sentence context experiment including a constraint manipulation, low- and high-constraint sentences have to be created for each target word and matched control. Sentences for targets and controls are preferably matched on number of words, syntactic structure, and the length of the word preceding the target. Critical words cannot be presented as the final word of the sentence because of sentence wrap-up processes on sentence-final positions (Rayner et al., 2000). To qualify sentences as low or high in terms of semantic constraint, sentence completion ratings have to be collected in a separate cloze probability study, in which participants are presented with the sentence frames up to the target word. They are instructed to write down the first word that comes to mind when reading the sentence. In order for the constraint manipulation to be successful, high-constraint sentences should be completed with one specific word, whereas low-constraint sentences should be completed with a variety of words. The resulting cloze probabilities allow one to verify the constraint manipulation and to further optimize the sentences. In order to avoid having the participants see the same target word twice, sentences can be divided across two presentation lists, so that each participant sees the target word and its control in either the low- or high-constraint sentence context.

Even though stimulus materials are carefully controlled, it is always useful to conduct a control experiment with monolinguals if possible, to ensure that effects are not due to any uncontrolled stimulus characteristics. A control experiment consists of testing a group of participants who have no knowledge of the nontarget language on the same stimulus materials. These participants should not be influenced by cross-linguistic overlap.

Procedure

An eye-tracking experiment typically starts with camera setup and calibration. After calibration is completed, the instructions are given to the participants. They are instructed to read the sentences as naturally as possible for comprehension (as if one were reading a book or a newspaper). Sentences are presented as a whole on the screen and participants can press a button indicating that they have finished reading the sentences. For single-sentence experiments, it is advisable to display the sentence on no more than two lines and in monospaced Courier font. If a sentence has to be presented on two lines, make sure that target words are never the final word of a line, nor the first word of the second line. Comprehension of the sentences and attention to the reading task are typically examined by presenting comprehension questions following some trials. Participants can respond "Yes" or "No" to these questions using the appropriate buttons or keys. Sentences have to be presented in a random order to each participant. It is advisable to start with some practice sentences so that participants get used to the reading task.

Summary and Conclusion

The L2 and L1 studies on bilingual visual word recognition in sentence context show that the language of the preceding words is an insufficient cue to restrict lexical access to words of the target language (e.g., Duyck et al., 2007; Schwartz & Kroll, 2006; Van Assche et al., 2009; Van Hell & De Groot, 2008). This literature offers strong evidence for a bilingual language system that is profoundly language-nonselective. Furthermore, studies measuring eye movements revealed that the degree of semantic constraint for a sentence does not necessarily eliminate lexical activation of the nontarget language (Van Assche et al., 2011), although there is evidence that it has an effect that occurs relatively late (Libben & Titone, 2009; Titone et al., 2011) and that it influences cross-lingual activation effects in lexical decision, translation, and naming studies (e.g., Schwartz & Kroll, 2006; Van Hell & De Groot, 2008). The difference in result patterns across studies indicates that the influence of a sentence and semantic context on language-nonselective activation is dependent on experimental factors such as task demands (e.g., lexical decision vs. eye-tracking; Van Assche et al., 2011; Van Hell & De Groot, 2008), type of bilingual tested (e.g., proficiency and age of acquisition; Libben & Titone, 2009; Van Assche et al., 2011), cross-overlap of translation equivalents (e.g., identical vs. nonidentical cognates; Duyck et al., 2007), and stimulus list composition (e.g., Titone et al., 2011).

List of Keywords

Bilingual Interactive Activation Plus Model (BIA+), Bilingual word recognition, Cognate facilitation effect, Cognates, Dual-language activation, Eye-tracking, Factorial design, First fixation duration, Gaze duration, Go-past time, High-constraint sentence context, Higher-order processes, Homograph facilitation effect, Homographs, Initial lexical access, Interlingual homographs, Language-nonselective lexical access, Lexical access, Lexical decision task, Low-constraint sentence context, Naming task, Proficiency, Rapid serial visual presentation (RSVP), Re-ordered Access Model, Semantic priming effect

Review Questions

1. The studies presented in this chapter all involve bilinguals who speak languages with the same scripts. There are, however, also languages with completely different scripts (e.g., Chinese, Hebrew). How do you think cognates will be processed in these languages?
2. Think about the bilinguals living in your country. How proficient are they in both languages? Do they use both languages regularly? How do you think proficiency can influence the degree of language-nonselective activation in the bilingual language system?
3. Can you think of other tasks that can be used to investigate whether bilinguals activate words in one or both of their languages? What advantages or disadvantages can you think of for each task?

Suggested Student Research Projects

1. *Textbook assignment*. Choose a page in a magazine or newspaper. See whether you can find cognates. How many cognates were you able to find? Some languages share many words across languages while other language pairs do not,

so depending on the relevant language pairs in your language context, you will find many or only a few. Check the word categories of the cognates. Most of the studies discussed in this chapter have focused on the processing of nouns. Do you think that the same results can be found for verbs or adjectives? In answering this question, especially consider the degree of semantic, orthographic, and phonological overlap between the words.

2. *Creating a lexical decision experiment*. In this project, you will try to find evidence for the hypothesis you generated above for cognate effects for other word categories, such as verbs or adjectives. Try to create your own lexical decision experiment in which you will examine whether cognates are processed more quickly than controls.
3. *Semantic processing in experimental tasks*. The processing of homographs in lexical decision tasks has shown that the orthographic overlap for homographs can lead to facilitation. However, how do you think homographs will be processed in tasks that obligatorily involve semantic processing (e.g., semantic categorization)?

Related Internet Sites

Experimental materials: http://www.tamiu.edu/~rheredia/materials.html
LexicALL: Data-sets: http://lexicall.widged.com/repository/listing.php
Software and data-sets: http://crr.ugent.be/
SUBTLEXus: Word frequency American English: http://expsy.ugent.be/subtlexus/
Van Orden overlap score: http://users.ugent.be/~rhartsui/Applet1.html
Word generator: WordGEn: http://www.wouterduyck.be/?page_id=29
Word frequencies: http://crr.ugent.be/programs-data/subtitle-frequencies
Wuggy: Multilingual pseudoword generator: http://crr.ugent.be/programs-data/wuggy

Suggested Further Reading

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. *The Quarterly Journal of Experimental Psychology*, *63*, 1266–1303.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals*: *An introduction*. New York: Psychology Press.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism*: *Language and Cognition*, *5*, 175–197.

Kroll, J.F., & De Groot, A. M. B. (2005). *Handbook of bilingualism*: *Psycholinguistic approaches*. New York: Oxford University Press.

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology*, *3*, 174.

References

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Arêas Da Luz Fontes, A. B., & Schwartz, A. I. (2010). On a different plane: Cross-language effects on the conceptual representations of within-language homonyms. *Language and Cognitive Processes, 25*, 508–532.

Balling, K. W. (2012). Reading authentic texts: What counts as cognate. *Bilingualism: Language and Cognition, 16*(3), 637–653.

Baten, K., Hofman, F., & Loeys, T. (2011). Cross-linguistic activation in bilingual sentence processing: The role of word class meaning. *Bilingualism: Language and Cognition, 14*, 351–359.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echalier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*, 235–260.

Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*, 34–44.

Brysbaert, M. (2013). LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*, 23–37.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One, 5*, e10729.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society, 13*, 212–214.

Coltheart, M., Davelaar, A., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 535–555). New York: Academic.

Costa, A., Santesteban, M., & Cano, A. (2005). On the facilitatory effects of cognates in bilingual speech production. *Brain and Language, 94*, 94–103.

Cuetos, F., Glez-Nosti, M., Barbon, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 32*, 133–143.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York: Psychology Press.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representations of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90–123.

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. *The Quarterly Journal of Experimental Psychology, 63*, 1266–1303.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*, 284–301.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language, 42*, 445–464.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Lawrence Erlbaum.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1*, 51–66.

Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity. Decade of behavior* (pp. 27–44). Washington, DC: American Psychological Association.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition, 12*, 273–289.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1340–1359.

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments & Computers, 36*, 488–499.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 663–679.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25*, 57–70.

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory, 4*, 161–178.

Gollan, T. H., Weisberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition, 15*(3), 594–615.

Green, D. W. (1998). Mental control and the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81.

Haigh, C., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning Memory and Cognition, 33*(4),623–44.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44*, 2–31.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*, 627–633.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44*, 325–343.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory and Cognition, 32*, 533–550.

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one echo: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*, 585–611.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady, 10*, 707.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 381–390.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q). *Journal of Speech, Language, and Hearing Research, 50*, 940–967.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English lexicon project. *Psychonomic Bulletin & Review, 13*, 45–52.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K., Binder, K. S., & Duffy, S. A. (1999). Contextual strength and the subordinate bias effect: Comment on Martin, Vu, Kellas, and Metcalf. *Quarterly Journal of Experimental Psychology, 52*, 841–852.
Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology, 53*, 1061–1080.
Sánchez-Casas, R. M., & García-Albea, J. E. (2005). The representation of cognate and non-cognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). Oxford, England: Oxford University Press.
Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 344–354.
Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.
Schwartz, A. I., & Van Hell, J. (2012). Bilingual visual word recognition. In J. Adelman (Ed.), *Visual word recognition* (pp. 131–150). New York: Psychology Press.
Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 65–76.
Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General, 112*, 1–36.
Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1412–1431.
Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*, 88–107.
Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science, 20*, 923–927.
Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb processing by bilinguals in sentence contexts: The effect of cognate status and verb tense. *Studies in Second Language Acquisition, 35*, 1–23.
Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology, 3*, 174.
Van Hell, J. G., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*, 431–451.
Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*, 780–789.
Van Orden, G. C. (1987). A rows is a rose: Spelling, sound, and reading. *Memory & Cognition, 15*, 181–198.

Chapter 3
Walking Bilinguals Across Language Boundaries: On-line and Off-line Techniques

Eva M. Fernández and Ricardo Augusto Souza

Abstract This chapter examines *off-line* and *on-line* methodologies used to study bilinguals. We demonstrate how methodological choices in experimental design are linked to the theoretical frameworks within which the research is cast. We illustrate how to identify appropriate methodological paradigms drawing from research on the integration of languages in bilinguals, specifically work on how bilinguals process argument structures with different restrictions in the standard grammars of their languages. We report data from Portuguese-English bilinguals and their monolingual counterparts performing three different tasks: *off-line acceptability judgments* using magnitude estimations, *on-line self-paced reading*, and *sentence recall/sentence matching* (i.e., providing whole sentence reading times, speech initiation times, and oral recall errors). With both on-line and off-line measures, bilinguals have different restrictions in argument structures than their monolingual counterparts, in their first language. The overall pattern suggests that these differences are rooted in grammatical representations rather than being driven by performance variables.

E.M. Fernández, Ph.D. (✉)
Queens College and Graduate Center, City University of New York,
65-30 Kissena Blvd., Flushing, NY 11367, USA
e-mail: eva.fernandez@qc.cuny.edu

R.A. Souza, Ph.D.
Universidade Federal de Minas Gerais,
Av. Presidente Antônio Carlos, 6627 Campus Pampulha,
Belo Horizonte, Minas Gerais, Brazil
e-mail: ricsouza@ufmg.br

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_3

Introduction

A perennial question in psycholinguistic research in the area of bilingualism asks whether language-specific knowledge is separated or integrated in the bilingual mind. A related question is to what extent such knowledge, be it integrated or separated, underpins performance in one or the other language of the bilingual. Evidence suggesting that there is some degree of integration between languages comes from the well-documented phenomenon of first-language (L1) transfer in second language (L2) production (Jarvis & Pavlenko, 2007; Odlin, 1989). A classic example of transfer is the sentence in (1), whose structure in English, the speaker's L2, has a word order that is recognizably German, the speaker's L1 (Weinreich, 1953):

1. *He comes tomorrow home*.

Transfer phenomena have led investigators to ask whether such instances of language transfer reflect the application of momentary learning strategies, used by learners to test hypotheses about the grammar of the language they are learning (Schachter, 1993). Some scholars refer to processing-based transfer effects as interference (Grosjean, 2012), but we will not make such a distinction here. Alternatively, L1 transfer could be a reflection of the way languages are represented and managed in the mind of the second language learner (Cook, 2002). What methodological tools are available to psycholinguists interested in studying these kinds of questions? Are some techniques, perhaps *on-line techniques*, better for capturing performance difficulties, and others, like *off-line techniques*, more sensitive to linguistic knowledge? There is no one-size-fits-all answer to these questions; this chapter offers some guidance for beginning investigators.

In the following sections we will explore how various standard on-line and off-line psycholinguistic techniques have been used in research on whether language-specific knowledge is integrated or separated in the bilingual's mind. As a point of reference, we will focus on a recent investigation from our own work, a series of experiments that exploit a cross-linguistic difference specified lexically and realized syntactically. We will describe how this phenomenon can be studied with techniques that reflect on-line processing, but we will argue that the critical questions emerging from this investigation can also be asked off-line. Our overarching objective is to illustrate the range of empirical tools available to investigators, tools designed to probe on-line processes or off-line intuitions. Thus, in this chapter, we hope to show:

- How to examine a broad question in the psycholinguistics of bilingualism using a specific linguistic phenomenon.
- How to link up an empirical technique to theoretical distinctions (such as on-/off-line and competence/performance).
- How to use a set of findings to design future studies.

The On-line/Off-line Distinction as a Theoretical Construct

Asking about the way bilinguals represent language-specific knowledge is only possible if we draw a distinction between the mechanisms that support *linguistic performance* and the knowledge of language that constitutes *linguistic competence* (Chomsky, 1965). Ambiguous sentences help us motivate empirically the theoretical distinction between competence and performance. Consider the following newspaper headline (LAPD Shot 2013):

2. *LAPD shot two women delivering newspapers by mistake*

Your knowledge of English allows you to construe two interpretations for the sentence in (2): one in which the mistake referred to in the sentence involves delivering newspapers, and the other (obviously the meaning intended by the author of the headline) in which the mistake was the shooting of the two women. Your knowledge of the grammar of English supports both interpretations, but one of the two (the one that might have made you wonder how delivering a newspaper could ever be done in error) likely came to you first, making it harder for you to retrieve the alternative intended meaning. Several decades of psycholinguistic research have demonstrated that speakers of English have systematic preferences in interpreting certain kinds of ambiguities, and have attributed such preferences to the performance mechanisms that facilitate language comprehension (for an overview, see Pickering & Van Gompel, 2006). In the example in (2) the prepositional phrase *by mistake* is interpreted as a constituent of the phrase headed by the verb *deliver*, because it is closer (and therefore computationally easier to attach to) than the phrase headed by the verb *shot*. Even though both interpretations are licensed by the grammar (i.e., linguistic competence), one interpretation is preferred when the sentence is processed in real time (i.e., linguistic performance).

The theoretical distinction between what you know about a language (competence) and how you put that knowledge of language to use (performance) is further complicated by the fact that language is used in real time: participants in a conversation produce and perceive sentences with a speed and agility that does not include time for conscious reflection on what is being produced or perceived. The kind of access you have to your knowledge of language is very different when you are consulting your linguistic competence on the fly (in real time) during production or perception, as compared to when you are thinking about whether a sentence you just wrote or read is ambiguous, ungrammatical, or a good example of L1 transfer. Let us introduce here the contrast that is the topic of this chapter, the distinction between processing that takes place *on-line*, on the fly, in real time, and reflections on language that take place *off-line*, after normal production or perception mechanisms have done their work. Notice that defining the on-line/off-line distinction this way, by tying it directly to processing that happens in real time (on-line) vs. processing that happens after the performance mechanisms have been applied (off-line) makes the distinction more theoretical than strictly methodological. It is not possible to talk about on-line methods if the framework one is considering for studying language production or perception does not propose performance mechanisms specific to production or perception. These performance mechanisms (see Fernández & Cairns, 2011) are responsible for extracting meaning from signals and producing signals from ideas, and rely on both grammatical knowledge (to which we turn next) and their own mechanism-specific routines—like the preference presented above, with example (2), for new phrases to attach to closer constituents.

People who are fluent users of a language have repositories of knowledge about that language. The discipline of linguistics is devoted to describing the nature of this kind of knowledge (for an introduction, see Cruz-Ferreira & Abraham, 2011), and compartmentalizes knowledge of language (grammar) into components that include what a language user knows about combining individual sounds into sequences of sounds (phonology), words into bigger words (morphology), and words into phrases (syntax). An additional repository of linguistic knowledge is the lexicon, where information about a word's form and meaning is stored. This knowledge is understood to be implicit, that is, not at all accessible consciously, though reflected in the judgments that fluent speakers can easily make about grammatical vs. ungrammatical sentences—judgments that are best understood as happening off-line and may be outside of the constraints imposed by performance mechanisms. Knowledge of language also guides the mechanisms that control language production and perception, constraining how sentences are encoded and decoded, processes that are considered to happen during real-time, on-line processing of language. In sentence comprehension, the parser is the mechanism that decodes the syntactic properties of the sentence being processed.

On-line linguistic performance is limited by the language processing machinery itself: there are limitations on working memory and on computational resources. It is difficult to process very long sentences because you cannot keep all the information active in your *working memory*—the temporary storage of information for immediate processing (Baddeley, 2003). It is also difficult to process constructions with heavy computational demands, like the following sentence, in which the computationally easier interpretation of *the sock* as the direct object of the verb *mend* leads to an incorrect structural representation of the sentence:

3. *While Mary was mending the sock fell off her lap.*

Linguistic competence is not limited in this way: the grammar of any language tolerates infinitely long sentences (natural language grammars are recursive, after all), as well as sentences with heavy computational demands (all natural languages have subordination, even though subordination is hard to compute). One way to think about this, relevant for methodological design, is that linguistic competence is hard to tap directly through an experiment, because to get to it one must go through the linguistic performance mechanisms, using some sort of task that measures a behavior, a brain response, or an ocular movement, and all of these measurable responses are limited in ways that linguistic competence is not. The task must also be designed to avoid complications that arise from conscious *metalinguistic knowledge*. Metalinguistic knowledge is the knowledge you access when you reflect consciously about language, when you think about it as a formal object with identifiable structural properties. The depth of this knowledge varies from person to person, and is influenced by language instruction and by training in formal linguistic analysis—in contrast to the unconscious grammatical competence of fluent speakers of a language, which is not much influenced by explicit instruction or training. Your metalinguistic knowledge is what triggers you to recognize the unintended ambiguity in (2), leading you to recover the meaning that the author of the headline intended. Your metalinguistic knowledge might also prompt you to think that a comma after *mending* in (3) would make the sentence much better. Notice, though, that your understanding of rules about comma use in standard writing might be much more robust than somebody else's, so a psycholinguistic task that invites metalinguistic ruminations might elicit very different responses from people with different types of training. Such differences are undesirable if what one is after is a better understanding of the competence and performance of ordinary speakers, regardless of their expertise in the formal analysis of language.

Before we move on to exploring different types of techniques that tap linguistic competence and performance in bilinguals, let us consider one more aspect of task design which intersects in complex ways with the on-line/off-line distinction. Some psycholinguistic paradigms are designed to make participants respond quickly—for instance, techniques where participants read or hear a linguistic stimulus and are asked to press a button as soon as the stimulus appears, as soon as it ends, or as soon as it is recognized as a grammatical string. Such paradigms are categorized as *speeded tasks*, and the speeded response is encouraged by some aspect of the design: instructions to "respond as quickly as possible" in combination with a response time-out (i.e., the task goes on to the next item if the participant does not respond within a small window of time). In *unspeeded tasks*, in contrast, participants may take as long as they need to respond and there is no time pressure for a response. Notice that an unspeeded task (e.g., a questionnaire in which participants read a sentence and judge the acceptability of that sentence, or a neurophysiological procedure with auditory stimuli in which the participant is required to do nothing more than listen) could certainly elicit responses that reflect the preferences of the on-line processing mechanisms. Regardless of whether the participant is asked to respond swiftly, the measured response itself may consist of data with extremely fine-grained temporal resolution (millisecond by millisecond), as is the case with techniques that capture eye movements during reading (Rayner, 1998) or in visual world paradigms (Ferreira & Tanenhaus, 2007), or with techniques that capture neural responses with measures such as *electroencephalogram* (EEG), *event-related brain potentials* (ERPs), or *magnetoencephalogram* (MEG; Kutas, Federmeier, & Sereno, 1999).

Thematic and Methodological Approaches

The question of whether linguistic knowledge is integrated or separate in the single mind of the bilingual is one that has been asked by many researchers of the psycholinguistics of bilingualism. In bilinguals, two languages reside within a single person, so how much integration is there between the two languages, and how much do they influence each other (or stay apart from each other)? Let us consider three different types of evidence, and the methods that the relevant studies have used: studies of bilingual lexical space, studies of transfer, and studies of cross-linguistic priming.

Studies of the bilingual *lexical space* have demonstrated that at least some integration exists between the bilingual's two languages. According to much of this research, lexical representations are shared. For example, translation equivalent words facilitate retrieval in certain types of lexical access tasks (Dijkstra, 2005), and this facilitation is greater when the translation equivalent words are also cognates (Sánchez-Casas & García-Albea, 2005). Lexical access is frequently studied with *lexical decision tasks*, where study participants are asked to judge whether a word presented visually or auditorily is a word (e.g., *blouse*) or not a word (e.g., *flouse*) in their language. In such tasks, the measure is the time it takes to make the lexical decision, and what is of interest is whether this time varies as a function of some experimental manipulation in the materials (responding to words in one category or another), or as a function of words that were presented just before (i.e., *lexical priming*). In lexical decision tasks, participants are not asked to reflect (i.e., off-line) about the nature of the stimuli. They are being asked to make rapid binary-choice responses about whether the target is a word or not. This kind of measure is constrained by on-line lexical processing routines, but in a study interested in determining whether integration exists in the bilingual lexicon, the task itself is used to probe the nature of lexical representations, so the fact that a lexical decision task is of an on-line nature is not the main feature of interest; what is of interest is that the patterns of behavior measured on-line reflect the constraints on lexical access imposed by the lexicon and its internal organization. On-line tasks can be used to study the nature and organization of lexical knowledge.

One way to examine how much two languages might be integrated in the single bilingual mind is to examine *transfer between the two languages*. As mentioned above, transfer is what happens when a structure in one language emerges in the other. Transfer from L1 to L2 has been documented in many investigations. Transfer can occur at any grammatical level (MacWhinney, 2005): phonology, morphology, syntax, and even pragmatics or conceptual structure. Documented cases of transfer from L1 to L2 suggest that knowledge of a dominant L1 can mediate performance in a second nondominant language. Consider for instance the case of the bilingual who speaks L2 with an "accent" that evokes the phonology of the L1. This can be studied from multiple perspectives: What characterizes the accent in L2 speech? How exactly does a very particular characteristic of the accent differ from native-like production? Does the nonnativeness also affect perception of contrasts in the L2? (Note that similar questions can be asked about transfer at other grammatical levels.) Investigations of phonological transfer from the perception angle use techniques that probe whether the speaker can perceive different types of contrasts. Such techniques not only measure some (off-line) behavioral response to a stimulus, like in *discrimination tasks*, where participants have to determine whether a stimulus differs from another, or whether it is more similar to one or another baseline.

Tasks like this could also reflect on-line processing, if they measure with high temporal resolution how long it takes to make this judgment. They can be coupled with ERP or eye-tracking, which allows the investigator to collect additional indirect evidence that reflects on-the-fly processing. *Elicited production* techniques, in which participants produce target utterances based on some stimulus provided by the experimenter, yield recordings of speech that can be analyzed acoustically (at very fine-grained temporal resolutions) or judged by experts or nonexperts on some parameter, like degree of accent (an off-line measure). During an elicited production technique, participants' eyes could be tracked (yielding an on-line measure of how visual stimuli guide speech planning). Evidently, a complex phenomenon like transfer can be studied from multiple methodological angles, and the best approaches are the ones that combine more than one measure reflecting the phenomenon of interest.

A third source of evidence about integration vs. separation in the bilingual linguistic architecture comes from studies of *cross-linguistic structural priming*, which have demonstrated that a syntactic structure just experienced in one language can prime (i.e., facilitate) production in the other language. For example, Spanish-English bilinguals in a study by Hartsuiker and colleagues (Hartsuiker, Pickering, & Veltkamp, 2004) were more likely to produce passive constructions in English when they had just heard a passive construction in Spanish. In structural priming experiments, participants are asked to produce sentences given some sort of a prompt (a set of words, or a picture to be described). The target sentence could be produced using two (or perhaps more) structures, and the objective is to see what structure the participant will choose to use when a preceding trial has involved a sentence instantiating one of the two alternatives. Some experiments take this a step further by using primes in one language and targets in the other. For example, in a study of German-English bilinguals, Loebell and Bock (2003) presented participants with an auditory prime in English, which the participant had to repeat out loud, followed by a picture to be described in German.

4. (a) *The lawyer sent his client the contract.*
   (b) *Eine Frau zeigt einem Mann ein Kleid.* [*A woman shows a man a dress.*]

This kind of procedure cannot be characterized as on-line, since its measures are not directly about moment-by-moment processing. What is measured is the proportion of responses that match the prime, and in this way it is an indirect measure of linguistic choices made during language production. Evidence of cross-linguistic structural priming, as reported by Hartsuiker and colleagues and by Loebell and Bock, suggests that the bilinguals' two languages are deeply integrated: residual activation of the structure just experienced in one language prompts the speaker to be more likely to produce that same structure in a sentence about to be uttered in the other language.

The broad set of studies just surveyed all have in common the fact that they examine how knowledge of one language affects performance in the other. Is what you know about language X a performance mediator for how you process (produce or perceive) language in language Y? In such research, investigators need to carefully control a set of interrelated variables: order of acquisition (which gives us the distinction between L1 and L2), age of acquisition, linguistic proficiency (which gives us the distinction between the dominant and nondominant language), and frequency of use (Grosjean, 2008). We can identify which is the first or the more dominant or the more frequent language of the bilingual and even make predictions about these if we know the bilingual's age of acquisition of the L2 and the length of residence in an L2 environment (Fernández, 2003; Marian, Blumenfeld, & Kaushanskaya, 2007).

Another fact in common in those studies is that they all examine effects of L1 into L2. The alternative, exploring whether knowledge in L2 affects performance in L1, avoids some of the problematic issues revolving around sequence and age of acquisition, language dominance, and perhaps even frequency of use. The investigation of argument structure in bilinguals we introduce below addresses the issue of integration of languages among bilinguals by exploring whether knowledge of L2 affects performance in L1. The investigation will allow us to further ask whether the L2 influence on L1 actually affects long-term representations in the L1 grammar.

Cross-Linguistic Variation in Grammatical Representations and Argument Structure

As the previous section suggests, questions about integration or separation of the bilingual's two languages are always cast in the context of specific linguistic phenomena and always require invoking some phenomenon that varies cross-linguistically. Descriptions of the grammars of the world's languages point to many grammatical properties that are subject to cross-linguistic variation. Those properties include restrictions on whether a language allows sentences with null (empty) subjects; Spanish does (5a), while English doesn't (5b):

5. (a) *Hablábamos sobre el libro*. [*Spoke*-1pl *about the book.*]
   (b) **Spoke about the book*.

(Throughout this chapter we will use asterisks to indicate ungrammaticality. For examples in a language other than English, we will provide relevant word-by-word glosses or translations in brackets and use conventions like "1pl" to indicate grammatical features, here: first person plural).

Other restrictions have to do with the placement of certain kinds of words, like adverbs inside verb phrases; French permits post-verbal adverbs (6a), while English does not (6b):

6. (a) *Jean prend souvent le bus*. [*Jean takes often the bus.*]
   (b) **John takes often the bus*.

One domain of linguistic knowledge that is subject to substantial cross-linguistic variation is the realization of argument structure. *Argument structure* is a property of the meaning of certain words, like verbs that require other accompanying concepts for their meaning to be fulfilled. For example, consider the meaning of a verb such as *kick*. The meaning of *kick* is only fulfilled if there exists reference to a kicker (the verb's *agent*) and also to an entity being kicked (the *patient*). The verb's agent and patient are referred to as the verb's *arguments*. Verbs are a word class whose meanings are typically associated to argument structures. The realization of argument structure of a verb is the mapping of that verb's arguments to specific morphosyntactic elements, such as noun phrases or prepositional phrases, and to specific grammatical roles such as subject or object. With the verb *kick*, the realization of argument structure requires for the agent to be linked to a noun phrase that is also the subject of the sentence and for the patient to be linked to another noun phrase that is the object of the ensuing sentence. You know exactly who performed the violent act when you hear a sentence like (7). This knowledge is a consequence of your knowledge of the realization of argument structure for the verb *kick*.

7. *Bob kicked Tom.*

The realization of argument structure varies significantly between different languages. Let's illustrate such variation with a simple example: consider the meaning of the verb *listen*, whose two arguments can also be labeled as agent and patient. You probably have already noticed that, unlike with *kick*, where the patient maps onto a noun phrase that is the object of the verb, with the verb *listen*, the patient maps onto the object of a preposition, which is by force of usage the preposition *to*:

8. (a) *Bob listened to Tom.*
   (b) **Bob listened Tom.*

In Portuguese, the equivalent of the verb *listen* is *ouvir*. But the difference between English and Portuguese goes beyond this lexical difference, as in Portuguese the equivalent arguments are realized as noun phrases that are the subject and object of the verb. So, the grammatical Portuguese equivalent of (8a) is (9b), where *Tom* is the direct object (the word-by-word translation of (8a) in (9a) is ungrammatical in Portuguese):

9. (a) **O Bob ouviu ao Tom.*
   (b) *O Bob ouviu o Tom.*

Cross-linguistic variation in argument structure may involve more than just the type of syntactic constituent arguments are mapped to. There are situations in which verbs that are cross-linguistic synonyms when it comes to their most basic meanings will show differences concerning the very number of arguments they require and permit. This means that many examples of translation-equivalent verbs do not share argument structures. One example is the English verb *shine* and its translation equivalent verb in Spanish, *brillar*. Both of these verbs subcategorize for a subject and may be used intransitively, that is, in sentences without a direct object:

10. (a) *The sun shines.*
    (b) *El sol brilla.*

But *shine* is optionally transitive, while *brillar* is not:

11. (a) *The machine shines shoes.*
    (b) **La máquina brilla zapatos.*

Thus, one of the tasks facing the L2 learner as he or she learns verbs in the new language is acquiring not only the possible subtle differences in the basic meaning and usage of such verbs, but also the differences that may exist between L1 and L2 argument structure and its realization. Ultimately, this means that bilinguals' knowledge of verbs in more than one language also encompasses knowledge of the language-specific syntactic structures where those verbs obtain well-formed sentences. Studies with L2 learners and bilinguals have demonstrated that L2 argument structure poses learning difficulties in an L2 (Juffs, 2000; Montrul, 2001; White, 2003).

The realization of argument structure is not only a consequence of the meaning encoded by individual words, such as verbs. Some syntactic patterns can be interpreted as linked to the realization of specific arguments, and such patterns are referred to as argument structure constructions (Goldberg, 1995, 2006). This means that verbs that would not be typically assumed to require a given set of arguments when taken as isolated lexical items may eventually participate in argument structure constructions that instantiate that set of arguments, provided there are no aspects of their meaning that contradict the meaning of the construction. Let's examine an example from Goldberg (2006, p. 154):

12. *Frank sneezed the napkin off the table.*

Surely, *sneeze* is not a verb that typically requires both agent and patient arguments. In fact, intransitive verbs like *sneeze*, requiring only an agent, have been used in sentence processing research, including some classic work on whether a verb's subcategorization restrictions can help readers avoid the difficulty in sentences like (3; Adams, Clifton, & Mitchell, 1998; Mitchell, 1987). Still, a core aspect of the meaning of the verb *sneeze*—namely, forced movement (i.e., of air, in this particular case)—corresponds well with a construction consisting of the sequence Subject-Verb-Object-Prepositional Phrase Adjunct, which supports an interpretation of caused motion. Familiarity with constructions may be what explains the fact that argument structure is subject to diachronic change. An interesting and relevant question in the study of bilingualism is whether innovations like the example in (12) can *walk across* linguistic boundaries as a result of knowledge and use of more than one language.

One specific case of cross-linguistic variation in argument structure is the expression of induced movement with verbs of manner of motion in English and in Brazilian Portuguese, which vary systematically between the two languages. In English, agentive and normally intransitive verbs of manner-of-motion such as *run*, *march*, and *jump* participate in a type of causative construction referred to as the *induced movement alternation* (Levin, 1993):

13. (a) *The researcher ran the mouse through the maze.*
    (b) *The general marched his soldiers along the street.*
    (c) *The coach jumped the students around the gym.*

The induced movement alternation construction entails a slight change in the conceptualization of the event typically expressed by verbs of manner-of-motion when they appear in intransitive constructions. This conceptualization change involves a shift from direct to indirect agentivity (Brousseau & Ritter, 1992), in other words, from direct to indirect responsibility for the action the verb expresses. Consider as an example the verb in sentence (13b) above. In a transitive construction such as *The soldiers marched along the street*, there is direct responsibility of the agent (*the soldiers*) for the marching event, so we can classify this usage as implying direct agentivity. In contrast, in the induced movement alternation construction, there is a shift of responsibility for the event from the agent of the verb to another argument. Thus, in (13b), the meaning of the sentence could be expressed as *The general caused the soldiers to march along the street*, a less economical sentence that captures that direct responsibility for the event has shifted from the argument expressing the agent of marching (*the soldiers*) to another argument, expressed by the noun phrase *the general*. All sentences in (13) are expressions of indirect agentivity.

As we can see, a key semantic property of the induced movement alternation is that it has a causative reading. In other words, it conveys the meaning that a given entity expressed by an argument caused an event to take place. Verbs of manner-of-motion are not the only kinds of verbs participating in causative constructions whose syntactic configuration resembles the induced movement alternation. For example, verbs denoting change of state—such as *melt*, *dry*, and *warm*—will naturally occur in sentences with a causative reading.

14. (a) *The cook melted the butter before adding the onions.*
    (b) *The dancers dried their costumes after the evening show.*
    (c) *The researcher froze the blood samples for her experiment.*

There are also verbs that will nearly always produce ill-formed sentences if they are forced into the same syntactic pattern of the induced movement construction, and if they sometimes produce acceptable sentences, it is because such sentences entail unusual, specific connotations. Intransitive verbs such as *arrive* and *appear* are examples of verbs that resist this type of causative construction:

15. (a) **The bus driver arrived the students late for school.*
    (b) **The detective appeared the evidence during the trial.*

In Brazilian Portuguese, constructions of the type Subject-Verb-Object-(Prepositional Phrase Adjunct) with causative reading also occur with some verbs. It resembles English concerning the fact that change-of-state verbs will naturally occur in this syntactic configuration, and also concerning the fact that some intransitive verbs tend to never participate in this construction. But interestingly, Brazilian Portuguese differs from English with respect to verbs of manner of motion, which will not produce causatives like the English induced movement alternation (Cambrussi, 2009). In Brazilian Portuguese, induced movement with verbs of manner-of-motion is expressed through periphrastic constructions with light verbs like *fazer* ("make"):

16. (a) *A pesquisadora fez o rato correr emu ma Caixa*. [*The researcher made the rat run in a box*.]
    (b) *O general fez seus soldados marcharem ao longo da rua*. [*The general made his soldiers march along the street*.]
    (c) *O treinador fez os estudantes pularem ao redor do ginásio*. [*The coach made the students jump around the gym*.]

In fact, expressing manner of motion with the induced movement alternation construction is not licensed in Brazilian Portuguese:

17. (a) **A pesquisadora correu o rato em uma caixa*. [*The researcher ran the rat in a box*.]
    (b) **O general marchou seus soldados ao longo da rua*. [*The general marched his soldiers along the street*.]
    (c) **O treinador pulou os alunos ao redor do ginásio*. [*The coach jumped the students around the gym*.]

Empirical evidence about this comes from a study that collected acceptability judgments from native monolingual speakers (Souza, 2011). The data showed that sentences like those in (17) are at least highly marginal, if not totally unacceptable, for monolingual speakers of Brazilian Portuguese. A similar restriction exists in Spanish (Montrul, 2001). Let us now turn to how bilinguals judge such sentences in their second language, English.

Exploring Bilinguals' Knowledge of the Induced-Movement Alternation

A method frequently used to explore people's knowledge of grammar is the *acceptability or grammaticality judgment task*. In acceptability judgment tasks, participants are asked to rate the well-formedness of linguistic stimuli, generally presented as sentences. Acceptability judgment tasks vary in the kinds of responses elicited from participants: some force participants to make a binary choice ("good" or "bad"), while others ask participants to use a ratings scale, seeking gradience in the perceptions of how well formed the materials might be. Acceptability judgments are typically considered to be off-line tasks, even in variants of the procedure that involve speeded responses. Participants are being asked to make a judgment about well-formedness, which will not necessarily be made on-the-fly using normal processing routines and which might be informed by metalinguistic knowledge. Despite their off-line categorization and the possibility that metalinguistic knowledge is being consulted, acceptability judgments are assumed to be informed by participants' grammar, so acceptability judgment tasks are generally accepted as reasonable tools to explore aspects of grammatical knowledge.

We recently collected data using an acceptability judgment procedure aimed at comparing how native speakers of English, native speakers of Brazilian Portuguese, and bilinguals of Brazilian Portuguese and English perceived the well-formedness of sentences whose structures instantiated the induced-movement alternation. The acceptability judgment task was designed according to what is known as the *magnitude estimation paradigm* (Bard, Robertson, & Sorace, 1996; Sorace, 2010). In magnitude estimation acceptability judgment tasks, participants are asked to assign a number of their choice rating an initial sentence, and then to assign numbers to each new sentence they are asked to judge according to their perception of how well- or ill-formed it is in comparison to the initial sentence. For example, if a participant thinks the initial sentence should receive a number 10, and she thinks another sentence in the experiment is two times better than that, she is to assign it a number like 20; if she thinks a sentence is about half as acceptable as the initial sentence, she should give it a number like 5. Magnitude estimation tasks generate fine-grained scales of relative acceptability. Although this technique is not universally considered to be superior to standard acceptability judgment tasks eliciting binary decisions (for critiques, see Sprouse, 2011; Weskott & Fanselow, 2011), such estimates of judgment ratios allow researchers to explore the degree of divergence in acceptability across constructions. The ability to capture relative acceptability is an advantage for investigations of phenomena that may cause *optionality and instability* in linguistic intuitions (Sorace, 2010, p. 67), so magnitude estimation tasks are popular in studies that probe possible changes in the linguistic representations of bilinguals resulting from their development in the L2. Our objective in using magnitude estimation was to gauge whether bilinguals and monolinguals differed in their judgments of induced movement alternation sentences and to estimate the size of such a difference, if any did in fact emerge.

Participants were native speakers of Canadian English (recruited in Canada) and native speakers of Brazilian Portuguese (recruited in Brazil)—monolinguals and Portuguese-English bilinguals. The bilingual participants were screened with a test of vocabulary knowledge in English, the *Vocabulary Levels Test* (VLT, Nation, 1990). The VLT places test takers into one of five levels of vocabulary mastery, based on scores that range between 0 and 90. Participants were either placed at level 3 ("low proficiency," scores of 28 or higher, presumably having mastered the 2000 highest frequency lexical items in English) or level 5 ("high proficiency," scores of 60 or higher, having sufficient vocabulary to cope with academic texts, newspaper articles, and similar writing). None of the participants was placed below level 3 on the VLT, and therefore, none had a really low proficiency in English as L2; our "low" vs. "high" labels merely served to distinguish the two groups with respect to each other. Bilinguals ($N=23$ per group) completed the task in English, and monolinguals ($N=23$ per group) in their respective native languages.

In this acceptability task, participants judged 72 sentences of various syntactic structures, half of which were ungrammatical. In all, 8 sentences in the set of 72 sentences contained the induced movement alternation—which, according to traditional accounts, is grammatical in English but ungrammatical in Portuguese.

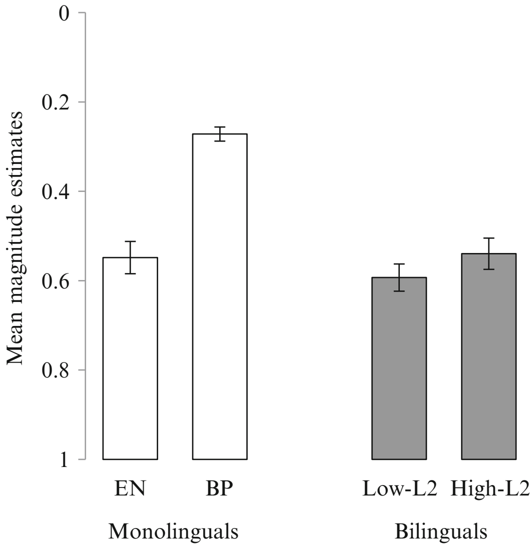


Fig. 3.1 Magnitude estimation scores for materials with induced movement alternation for English (EN) and Brazilian Portuguese (BP) monolinguals (*left*) rating sentences in their L1, and low proficiency (low-L2) and high proficiency (high-L2) Portuguese-English bilinguals (*right*) rating sentences in English, their L2. The scale represents mean magnitude estimates from 0 (best) to 1 (worst)

To analyze the magnitude estimation data for the induced movement alternation sentences, we converted the numbers assigned by participants into a 0 (best, most acceptable) to 1 (worst, least acceptable) scale. Mean magnitude estimates for the four groups of participants are plotted in Fig. 3.1, in which bar height corresponds with degree of unacceptability (the higher the bar, the less acceptable the materials). Analyses of variance (ANOVA) with participants (F_1) and items (F_2) as random variables revealed statistically reliable differences between monolinguals and bilinguals, $F_1\,(3,88) = 23.10, p < .001$; $F_2\,(3,21) = 9.27, p < .001$; post-hoc tests confirmed that Portuguese monolinguals differed reliably from the other three groups ($p < .001$), which did not differ from each other ($p > .20$).

The magnitude estimation data indicated that Portuguese monolinguals find induced alternation constructions objectionable, compared to bilinguals and English monolinguals. The results of this acceptability judgment task confirm the cross-linguistic contrast previously identified in the literature: Portuguese monolinguals judge sentences containing the induced movement alternation construction as less grammatical than English monolinguals. Bilinguals, even those with lower proficiency in L2, rated the materials as grammatical as the English monolinguals. Bilinguals appear to have no special difficulty learning that manner-of-motion verbs such as *run* and *jump* participate in the induced movement alternation structure in English. Comparing the data displayed in Fig. 3.1 to data reported by Souza (2011) supports this speculation. In that study, participants with a VLT level of 2 or below (i.e., very low proficiency) judged materials in English like their Portuguese monolingual counterparts. All of this is as expected, but what happens when we examine preferences for bilinguals in their L1?

Processing Induced-Movement in English and Brazilian Portuguese

A related investigation (Souza, 2012) examined how induced-movement alternations are processed on-line by monolingual and bilingual speakers of the two languages. This study included five groups of bilinguals ($N=9$ each): English and Portuguese monolinguals reading in English or Portuguese, bilinguals with high- or low-L2 proficiency reading in English, and bilinguals with high-L2 proficiency reading in Portuguese. (As with the previous study, low-L2 bilinguals had VLT scores of 3, and high-L2 bilinguals had VLT scores of 5.) This study did not focus on whether participants had a mental representation of the induced-movement alternation construction. Instead, this study explored whether the different grammatical status of this construction in Portuguese and English would be reflected in the relative ease or difficulty that Brazilian Portuguese-English bilinguals experience when processing sentences instantiating the induced-movement alternation structure. Would the different states of linguistic knowledge possessed by speakers of English and speakers of Brazilian Portuguese impinge on their linguistic performance in relation to the induced-movement alternation with verbs of manner of motion? And would the participants' bilingualism result in differences in on-line processing of induced-movement alternation sentences?

The task was a *self-paced reading procedure* in which sentences appeared first as a sequence of dashes on a computer screen. When the participant pressed a button on the keyboard, the dashes were replaced by groups of words, left-to-right. On subsequent button presses, the previous group of words was replaced by dashes and the next group appeared, so the participant only had visual access to a chunk of the sentence at a time. This type of task in known as *noncumulative moving window self-paced reading*. The example in (18) shows, line by line, how one of the items of the English sentences used in Souza (2012) study would have been seen by the participants (each line corresponds to one screen).

18.	-------------	---------	----------	---------------------
	The trainer	---------	----------	---------------------
	-------------	*jumped*	----------	---------------------
	-------------	---------	*the lion*	---------------------
	-------------	---------	----------	*through the hoop*

The stimuli were 32 sentences in each language, 8 of which contained the induced-movement alternation; the remainder 24 sentences were distractor sentences. All stimuli were divided into four presentation frames, as in the example. Two versions of the task were constructed: one with materials in English and the other with equivalent sentences in Portuguese. Consequently, in the Portuguese version, there were sentences that forced the induced-movement alternation with verbs of manner of motion into Portuguese. The task was presented to participants as a sentence comprehension task, so each item was followed by a yes/no question.

In self-paced reading tasks, the measure of interest is participants' reaction times (also called *reading times*) for each frame of the sentence. The data of interest, typically reported in milliseconds (ms), are the time lags between the moment a participant first sees a new frame of text on the screen and the moment they press a button to advance to the next frame. Measurements of reaction times as indicators of the cost of cognitive processes (i.e., the longer it takes, the more costly the processing event is) have a long history as a method for the study of mental processing in psychology (Goodwin, 2003). The critical region in Souza (2012) was the object noun phrase (*the lion* in the example in sentence 18). The guiding hypothesis was that processing of this noun phrase would be facilitated for participants who had access to the grammatical representation of the possibility that manner-of-motion verbs, which often occur as intransitives, may also occur with transitive behavior (i.e., with a direct object). Thus, the assumption was that such participants would read the second noun phrase faster than the participants who did not have access to the necessary grammatical representation. The results are displayed in Fig. 3.2, in which bar height corresponds with processing complexity (the higher the bar, the longer it took to read the critical region). An ANOVA indicated reliable differences between monolinguals and bilinguals, ($F_1(4,32)=6.75$, $p<0.01$, $F_2(4,28)=49.24$, $p<.001$);

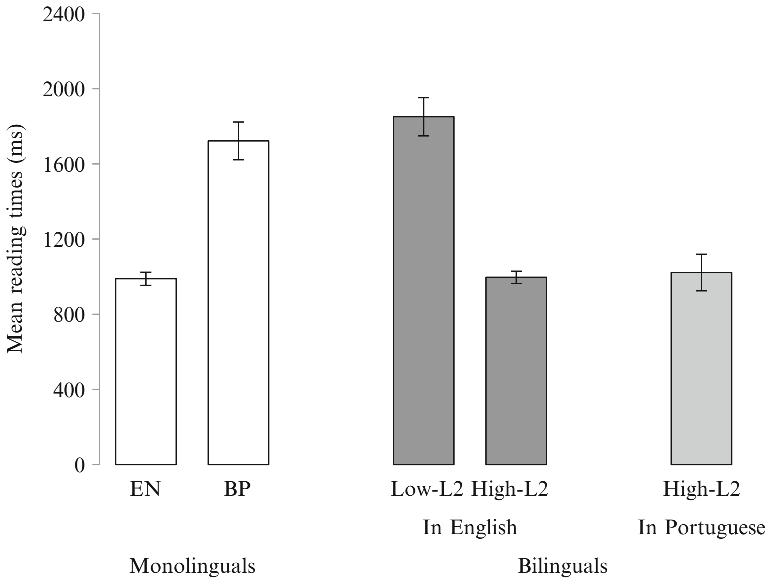


Fig. 3.2 Self-paced reading times (in ms) for the third frame (direct object noun phrase) for materials with induced movement constructions, for English (EN) and Brazilian Portuguese (BP) monolinguals (*white bars*) reading in their respective L1s; for low- and high-L2 proficiency Portuguese/English bilinguals reading in English (*dark gray bars*); and for high-L2 proficiency bilinguals reading in Portuguese (*light gray bar*). (Adapted from Souza, 2012)

post-hoc tests confirmed the differences observed in Fig. 3.2: English monolinguals and high L2-bilinguals (reading in English or Portuguese) do not differ from each other, but they differ from Portuguese monolinguals and from low-L2 bilinguals reading in English.

Using a very different task and a very different measure, the data from this experiment replicate the differences between the monolingual groups observed with the acceptability judgment task. Self-paced reading times for the critical region show that Brazilian Portuguese monolinguals and low-L2 proficiency bilinguals experience processing difficulties while reading induced movement alternations. In contrast, English monolinguals and high-L2 proficiency bilinguals are not experiencing the same kind of difficulty. The high-L2 proficiency bilinguals appear to be departing from the restrictions of their L1, exhibiting more tolerance for the innovative construction in their L1.

Momentary Lapses in Performance or Long-Term Reorganization of the Grammar?

The next logical question is to ask where this departure from L1 restrictions is coming from. Bilinguals exhibit a tolerance for the induced-movement with manner-of-motion verbs, but is this a long-lasting tolerance? That is, does it persist beyond the few hundred ms it takes to integrate the object of the verb during self-paced reading? The findings presented in Fig. 3.2 come from a procedure designed to measure on-line processing, during the rapid integration of words in the course of sentence processing. Increased reading times could simply mean momentary lapses in performance, or they might instead reflect a long-term reorganization of the grammar. In other words, this tolerance of structures that are ungrammatical in L1 might be brief or it might be sustained, in which case it could be indicative of a change in the argument structure representations in the bilinguals' L1.

This question required the design of a methodology that would allow exploring to what extent argument structure representations in L1 undergo long-term changes as a result of bilingualism. We developed a *sentence recall/sentence matching* paradigm that would capture behaviors reflecting speakers' internal competence—in this case, knowledge of subcategorization frames for verbs. Each trial began with a sentence displayed in the center of a computer screen, which participants were asked to read silently (with a time-out duration of 9000 ms). When they were done reading silently, a button press made a gray icon appear, signaling to the participant to prepare for oral recall. After 1000 ms, the gray icon turned red, signaling that the computer was recording audio. Participants then had to say aloud the sentence they just read. Immediately after a window of 9000 ms of time for recording, a new sentence appeared on the computer screen, and participants made a matching judgment on whether this sentence was the same as the sentence just read-and-said. There was feedback on every trial, indicating to participants whether their matching judgment was correct. The display was controlled by DMDX (Forster & Forster, 2003) running on a personal computer. Participants wore microphones and used the computer's keyboard to interact with the display.

To the participants, this was a sentence matching task (not a task about detecting ungrammaticality). This procedure allowed us some justification for presenting ungrammatical materials: we instructed participants that they would be seeing sentences that might seem a bit odd, but stressed that their job was simply to read and remember the sentences carefully, to see how accurately they could perform the matching task. They were not asked to react to the grammaticality of the materials at all.

This procedure yields a number of different measures of processing difficulty, all of which might reflect sensitivity to grammatical violations in the stimuli. None are strictly on-line measures, though all were collected in a speeded (and quite demanding) task. *Whole sentence reading times* reflect the mean amount of time it took participants to read the sentence silently when it appeared on the first screen. *Speech initiation times* are the amount of time it took participants to begin uttering the sentence when the recording icon appeared on the second screen. *Oral recall errors* are the number of words omitted or recalled inaccurately when participants uttered the sentence. (Note that additional data can be generated by performing acoustic analyses on the recorded utterances).

The data originated from three groups of participants: English monolinguals ($N=12$), high-L2 proficiency bilinguals ($N=13$), and low-L2 proficiency bilinguals ($N=11$). English monolinguals were recruited in New York City. Bilingual participants, also recruited in New York City, were all dominant speakers of their first language, Brazilian Portuguese, and learned English in adolescence or adulthood. Although English was their weaker language, they were all fluent speakers of the language. The task is, in fact, too demanding for bilinguals who are not very fluent in both languages. The two different proficiency groups in the bilinguals were formed based on proficiency in English, using VLT scores (Nation, 1990). As reported earlier: low-L2 bilinguals had VLT scores of 3, and high-L2 bilinguals had VLT scores of 5.

In this experiment, the critical set of materials consisted of items designed to instantiate three item types, illustrated below: six items in each list had change-of-state verbs with causative meanings (19a), another six contained the induced-movement construction used in the preceding experiments (19b), and a third set consisted of six pseudocausatives (ungrammatical causative-like sentences created by using intransitive verbs and adding a direct object; 19c).

19. (a) *As jovens dançarinas secaram suas saias no teatro*. [*The young dancers dried their dresses at the theater*.]
    (b) *A senhora elegante andou seu marido a um assento*. [*The elegant lady walked her husband to a seat*.]
    (c) *A mulher engraçada riu as crianças durante a festa*. [*The funny woman laughed the children at the party*.]

Pseudocausatives (19c) are ungrammatical in both languages, change-of-state causatives (19a) are grammatical in both languages. This design offers a way to measure, within participants, disruptions caused by induced-movement materials (19b), compared to the grammatical (19a) and ungrammatical (19c) "baselines." Filler items

($N = 18$) were created by using the lexical content from the target items in structures where the verbs all appeared as intransitives and were followed by a conjunct:

20. (a) *A saia de couro secou mas os sapatos demorarão mais*. [*The leather dress dried but the shoes took longer*.]
    (b) *Os cavalos andaram mas o fazendeiro ficou sentando*. [*The horses walked but the farmer remained seated*].
    (c) *A professora criativa riu mas o aluno falava sério*. [*The creative teacher laughed but the student spoke seriously*.]

Exactly half of the materials were presented as mismatching trials. The mismatching trials were restricted to the change-of-state targets and the fillers. Mismatches were located at different places in the sentences: the mismatch was early, in the middle, or at the end. Monolingual participants performed the task once, in their native language, and took 15–20 min to complete the procedure. Bilingual participants performed the task twice, within the same session, once in each language: Portuguese first, followed by English. For each task in each language, participants were pseudorandomly assigned to one of the two lists of the experiment, which counterbalanced lexically related target-filler pairs, so that translation-equivalent lexical content would not overlap. To encourage participants to "forget" the task in Portuguese before they began the task in English, they were asked to take 5 min to play a nonverbal video game (Pac-Man, 2011). A complete bilingual session took 50–60 min.

Figure 3.3 displays data for English monolinguals; for each graph, bar height corresponds with processing complexity: the higher the bar, the longer whole sentence reading times and speech initiation times, and the greater the number of errors. We performed paired *t*-tests by participants ($df_1 = 11$) and items ($df_2 = 10$); the graphs annotate comparisons that were marginal or significant in the analyses. As expected, the monolingual response patterns reflect the grammaticality of the induced motion

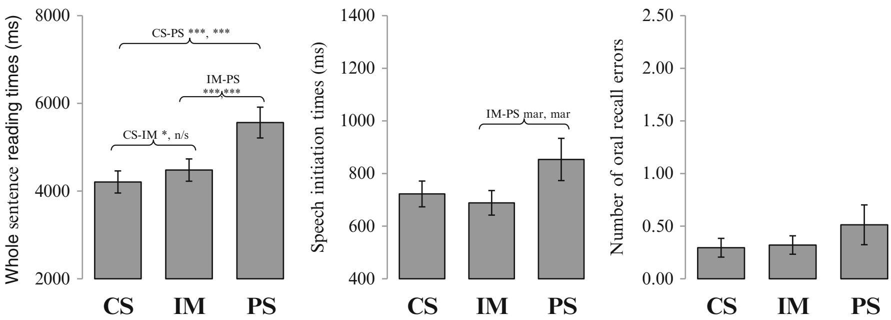


Fig. 3.3 Mean whole sentence reading times (*left* panel), speech initiation times (*center* panel), and number of oral recall errors (*right* panel) for English monolinguals performing the sentence recall/sentence matching task in English, for materials with change of state (CS), induced movement (IM), and pseudocausative (PS) constructions. *Note*. ***$p < .001$; **$p < .01$, *$p < .05$, marginal (mar) $.05 < p < .10$, n/s $p > .10$

construction, most clearly with whole sentence reading times, where induced motion materials are significantly different from the ungrammatical pseudocausative construction, $t_1(11) = 7.63$, $p < .001$; $t_2(10) = 5.67$, $p < .001$, but are significantly different from the grammatical change-of-state construction only by participants, $t_1(11) = 2.68$, $p < .05$; $t_2(10) = 1.46$, $p > .10$. Induced motion materials also had marginally lower speech initiation times than ungrammatical pseudocausatives, $t_1(11) = 1.82$, $p = .094$; $t_2(10) = 5.67$, $p < .001$, and did not differ from grammatical change-of-state materials in speech initiation times, $t_1(11) = 0.88$, $p > .10$; $t_2(10) = 1.88$, $p = 0.89$.

The bilingual data are displayed in Fig. 3.4 for low-L2 proficiency bilinguals and Fig. 3.5 for high-L2 proficiency bilinguals. Let us first consider low-L2 bilinguals, who are the participants we expect to differ most from English monolinguals in how they treat induced motion materials. As with the monolingual data, we performed *t*-tests by participants ($df_1 = 10$) and items ($df_2 = 10$), and the graphs annotate comparisons that were marginal or significant. The data patterns suggest low-L2 proficiency bilinguals do differ from their monolingual English counterparts, and not just in overall processing difficulty for their nondominant language (in all three

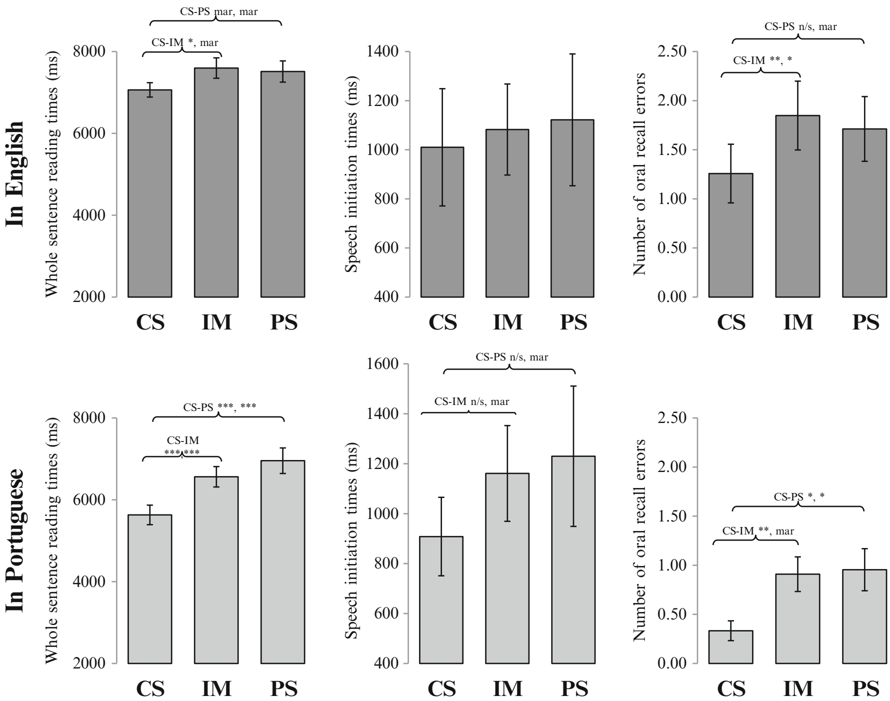


Fig. 3.4 Mean whole sentence reading times (*left* panel), speech initiation times (*center* panel), and number of oral recall errors (*right* panel) for low-L2 proficiency bilinguals performing the sentence recall/sentence matching task in English (*dark gray bars*) or Portuguese (*light gray bars*), for materials with change of state (CS), induced movement (IM), and pseudocausative (PS) constructions. *Note*. ***$p < .001$; **$p < .01$, *$p < .05$, marginal (mar) $.05 < p < .10$, n/s $p > .10$

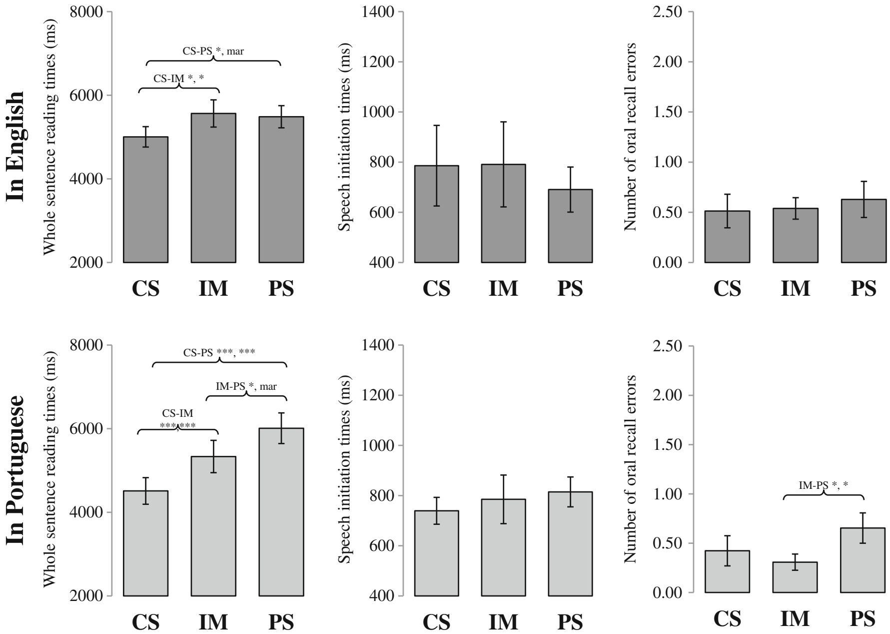


Fig. 3.5 Mean whole sentence reading times (*left* panel), speech initiation times (*center* panel), and number of oral recall errors (*right* panel) for low-L2 proficiency bilinguals performing the sentence recall/sentence matching task in English (*dark gray bars*) or Portuguese (*light gray bars*), for materials with change of state (CS), induced movement (IM), and pseudocausative (PS) constructions. *Note*. ***$p<.001$; **$p<.01$, *$p<.05$, marginal (mar) $.05<p<.10$, n/s $p>.10$

measures, these bilinguals perform much worse in English than in Portuguese). In the sentence reading times and oral recall errors of low-L2 bilinguals, in both Portuguese and English, low-L2 bilinguals have a response pattern we might expect from Portuguese monolinguals: induced motion materials pattern with ungrammatical pseudocausatives (for all three measures, $p>.10$ for the IM-PS comparison), rather than with the change-of-state materials, a pattern stable for reading times and oral recall errors and numerically in the right direction for speech initiation times in Portuguese (CS-IM comparison: reading times, $t_1(10)=4.36$, $p<.001$; $t_2(10)=4.40$, $p<.001$; oral recall errors: $t_1(10)=3.44$, $p<.01$; $t_2(10)=2.00$, $p=.074$).

The data for high-L2 proficiency bilinguals are displayed in Fig. 3.5. We have left this group for last, since they are the most interesting: they are unlike English monolinguals in English, and they are unlike their low-L2 proficiency bilingual counterparts in Portuguese. As with the other data, we performed *t*-tests by participants ($df_1=12$) and items ($df_2=10$), and the graphs annotate comparisons that were marginal or significant. High-L2 bilinguals performed in English in a non-native-like fashion: induced motion materials patterned with ungrammatical pseudocausatives (reading times: $p>.10$) rather than with the grammatical change-of-state, $t_1(12)=2.25$, $p<.05$; $t_2(10)=2.58$, $p<.05$. In Portuguese, in contrast, high-L2 bilinguals patterned after English monolinguals with oral recall errors: induced motion materials were easier than materials with pseudocausatives, $t_1(12)=2.19<.05$; $t_2(10)=2.24$, $p<.05$. Lastly, with sentence reading times, induced motion materials were an intermediate category: more costly than grammatical change-of-state causatives, $t_1(12)=5.18$, $p<.001$; $t_2(10)=3.18$, $p<.01$, but less costly than ungrammatical pseudocausatives, $t_1(12)=2.53$, $p<.05$; $t_2(10)=2.00$, $p=0.74$. The data suggest that induced motion verbs are undergoing some sort of change in the Portuguese of bilinguals with high proficiency in English, perhaps as a result of sustained exposure to English.

Summary: A Walk Across Language Boundaries

In this section, we have presented data about how bilinguals process argument structure when preferences for the bilingual's two languages are cross-linguistically different. We provided evidence (in the off-line magnitude estimation judgments reported in Fig. 3.1 and in the on-line self-paced reading data reported in Fig. 3.2) that monolinguals treat the induced motion alternation differently: it is grammatical in English, but ungrammatical in Portuguese.

Do bilinguals differ from their monolingual counterparts, and are those differences driven by performance mechanisms or competence repositories? The emerging picture is rather complex. With materials in English, even low-L2 bilinguals resemble English monolinguals in their off-line acceptability ratings (Fig. 3.1), but not in on-line self-paced reading (Fig. 3.2). The tolerance that high-L2 bilinguals exhibit in English seeps into Portuguese (Fig. 3.2).

The performance of bilinguals in our sentence recall/sentence matching task provides an even richer picture of the way bilinguals process argument structure. With materials in English, in both bilingual groups, we observe a divergence from the English monolingual group: induced motion materials pattern with ungrammatical pseudocausatives for bilinguals. The tolerance for these materials suggested by bilingual acceptability ratings data in English is not reflected in their immediate responses to the materials in the sentence recall/sentence matching task. Perhaps this is because the task is less subject to metalinguistic influences, or (more likely) because the acquisition of this property of their new language is still in flux.

Finally, with materials in Portuguese, a difference emerges between low-L2 bilinguals, whose responses are what we would expect them to be, were they Portuguese monolinguals, and high-L2 bilinguals, whose data patterns suggest their L1 is undergoing a change: developing a tolerance for induced motion constructions in Portuguese. These changes in L1 do not result in a compromised overall proficiency in L1 for these high-L2 bilinguals: they are still Portuguese dominant bilinguals. We take this to be evidence of change in the representations for these verbs in L1 for high-L2 proficiency bilinguals. Among the unresolved questions we will have to probe with future studies is the source for this change in argument structure realizations, which might be entirely happening within lexical projections or might instead reflect a tendency (stronger in bilinguals than in monolinguals) for lexical coercion into syntactic argument structure.

Results such as these converge with findings from a number of other investigations that demonstrate that bilinguals are not simply two monolingual systems combined into one. Representations in an L2 can be influenced by the L1 (as reflected in our bilinguals' performance in English). Furthermore, representation in the L1 can be influenced by the L2, particularly as proficiency in L2 increases. The emerging picture evokes the term *multicompetence*, proposed by Cook to describe the knowledge state of bilinguals (Cook, 1991, 2006). Multicompetence is the compound state of linguistic knowledge that results from the coexistence of one or more grammars in the bilingual mind (Cook, 1991), which can be construed as a state that does not correspond to the state of monolingual speakers of the bilingual's L1 or L2. Viewing bilinguals as multicompetent is one way of stressing that bilinguals are not two monolinguals within the same person. Perhaps one property of multicompetence is higher tolerance, on occasion, for constructions not licensed in L1 but licensed in L2.

Summary and Conclusions

This chapter provides an overview of some important considerations for beginning investigators—considerations that should help with formulating studies about both broad and narrow questions in the psycholinguistics of bilingualism. The process involves choosing a specific phenomenon, linking up empirical techniques to theoretical distinctions, and building on experimental findings incrementally.

We began the chapter by describing the theoretical underpinnings of the on-line vs. off-line distinction in language processing research. It is not possible to distinguish between on-line and off-line experimental paradigms without assuming a difference between what speakers know about a language (i.e., linguistic competence) and how they put that language to use in real time (i.e., linguistic performance). We also argue that on-line and off-line tasks do not strictly reflect performance and competence, respectively. Instead, any psycholinguistic technique relies on measuring some aspect of performance, and the investigation itself may be focused on developing a better understanding of the performance mechanisms or the competence repositories. We also described differences between speeded and unspeeded psycholinguistic tasks and stressed that the speeded vs. unspeeded distinction does not overlap with the on-line vs. off-line distinction: unspeeded tasks could be designed in such a way that they reflect on-line processing and speeded tasks may rely on off-line processing.

In surveying how some on-line and off-line techniques have been used to study bilinguals, we focused on the question of whether the bilingual cognitive architecture integrates or separates the bilingual's two languages. The range of research we considered (including investigations of lexical space, L1-to-L2 transfer, and cross-linguistic priming) included investigations that tap competence and performance using both on-line and off-line techniques. The most insightful investigations are conducted using multiple techniques and are deliberate in connecting the experimental methodology with the empirical questions being pursued: an investigation regarding lexical representations in bilinguals requires dramatically different approaches than an investigation about parsing strategies in bilinguals.

Finally, we illustrated the process of experimental design by drawing from some of our own recent work on how bilinguals process and represent argument structures that differ for monolinguals of their two languages. No single experimental paradigm could answer all of our questions directly. Instead, we used the data patterns revealed by one experiment to help formulate new research questions and to drive the design for the next. This incremental approach to experimental design is highly desirable in any kind of empirical investigation, but is especially important in studying bilinguals, whose variable linguistic backgrounds mean that the results of one experiment may not easily generalize with a new sample of participants.

List of Keywords

Acceptability judgments, Age of acquisition, Agent, Argument structure, Causative, Change-of-state verbs, Cross-linguistic priming, Cross-linguistic variation, Discrimination tasks, Dominant language, Elicited production, Language transfer, Grammaticality judgments, Induced-motion alternation, Lexical access, Lexical decision tasks, Lexical priming, Lexical representations, Lexical space, Linguistic competence, Linguistic performance, Magnitude estimation judgments, Manner-of-motion, Metalinguistic awareness, Moving window paradigm, Multicompetence, Off-line, On-line, Oral recall errors, Phonological transfer, Post-verbal adverbs, Prepositional phrase, Processing-based transfer, Pseudocausatives, Reaction time, Reading times, Self-paced reading, Sentence matching, Sentence recall, Speeded tasks, Syntax, Transfer, Transitivity, Unspeeded tasks, Verb argument, Verb phrases, Vocabulary Levels Test (VLT)

Review Questions

1. In this chapter, we described the parser as *the mechanism that decodes the syntactic properties of the sentence being processed*. Is the parser identical to the theoretical entity that linguists refer to as the *grammar*? Which facts presented in this chapter motivate the existence of the parser?
2. The studies reported in this chapter suggest bilingualism has effects on the first language of bilinguals. The reverse type of cross-linguistic influence (i.e., L1 influence on L2) can sometimes be described as a strategy L2 users employ in an attempt to overcome gaps in their knowledge of the second language. Considering that L1 influences on L2 may reflect conscious strategic behavior to cope with insufficient knowledge of a linguistic system, evaluate the contribution that studies like the ones reported in this chapter can bring to the question of whether bilingual language representations are integrated or separate.

3. Which specific contributions can *off-line* vs. *on-line* methods bring to research that tries to answer the question of whether bilinguals' language representations are integrated or separate?
4. In this chapter, we advocate linking empirical techniques to theoretical models of language representation or language processing. What are some of the liabilities of not making these connections?

Suggested Student Research Projects

1. One of the consequences of the notion of *performance* as distinct from *competence* is that lapses in performance are temporary. Consequently, ambiguous sentences such as example (3) in this chapter could be temporarily disruptive to the point of being perceived as ungrammatical by some readers (a performance issue), but such misperceptions would eventually be corrected (an effect of their competence). Test this hypothesis by designing and administering an unspeeded grammaticality judgment task. Let your materials consist of 6 sentences like sentence (3), 12 sentences that are clearly ungrammatical, and 18 sentences that are clearly grammatical. Create a randomized list for your sentences and collect grammaticality judgments from 10 to 15 speakers.
2. We could hypothesize that, if readers are given more time, they are more likely to overcome the misperception of ungrammaticality of an ambiguous sentence such as (3). Using the same materials used in Research Project #1, test this hypothesis by comparing the judgment data from Research Project #1 with judgment data collected through a version of the task in which readers are given a maximum time of 5,000 ms to produce their judgments. You might want to present your materials using software like DMDX to control the 5,000-ms time-out.
3. This chapter discussed one specific type of argument structure construction: the induced-movement alternation. Find another argument structure realization pattern in which English contrasts with a language other than English (you may want to develop this project with a speaker of that language). If you have access to L1 speakers of that language who are also L2 speakers of English, design a study to investigate whether they can acquire the English pattern.
4. As a follow-up to Research Project #3, design a study that examines the influence of L2 on L1 with your chosen argument structure realization pattern.

Related Internet Sites

DMDX display software: http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm
DMDX overview: http://www2.gsu.edu/~eslnxj/dmdx/usedmdx.html
Sentence parser: http://zzcad.com/parse.htm

Sentence processing script: http://step.psy.cmu.edu/scripts/Linguistics/Boland1990.html
Syntactic ambiguity script: http://step.psy.cmu.edu/scripts/Linguistics/MacDonald1993.html
Tasks library: http://www.millisecond.com/download/library/categories/psycholinguistics

Suggested Further Reading

Carreiras M., & Clifton, C. E. (Eds.) (2004). *The on-line study of sentence comprehension*: *Eyetracking*, *ERP and beyond*. Hove, England: Psychology Press.
Cook, V. J. (Ed.) (2003). *Effects of the second language on the first*. Clevedon, England: Multilingual Matters.
Garrod, S. (2006). Psycholinguistic research methods. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (2nd ed., Vol. 10, pp. 251–257). New York: Elsevier.
Grosjean, F. (2008). *Studying bilinguals*. Oxford, England: Oxford University Press.
Jarvis, S., & Pavlenko, A. (2007). *Crosslinguistic influence in language and cognition*. New York: Routledge.
Pickering, M. J., & Van Gompel, R. P. (2006). Syntactic parsing. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 455–503). London: Academic.

References

Adams, B. C., Clifton, C., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review, 5*(2), 265–270.
Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208.
Bard, E., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language, 72*(1), 32–68.
Brousseau, A.-M., & Ritter, E. (1992). A non-unified analysis of agentive verbs. In D. Bates (Ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics* (pp. 53–64). Stanford, CA: CSLI Publishing.
Cambrussi, M. F. (2009). *Alternância Causativa de Verbos Inergativos no Português Brasileiro [Causative alternation of inergative verbs in Brazilian Portuguese]*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
Cook, V. J. (1991). The poverty-of-the-stimulus argument and multi-competence. *Second Language Research, 7*(2), 103–117.
Cook, V. J. (2002). Background to the L2 user. In V. J. Cook (Ed.), *Portraits of the L2 user* (pp. 1–28). Clevedon, UK: Multilingual Matters.
Cook, V. J. (2006). Interlanguage, multi-competence and the problem of the "second" language. *Rivista di Psicolinguistica Applicata, 6*(3), 39–52.
Cruz-Ferreira, M., & Abraham, S. A. (2011). *The language of language: A linguistics course for starters*. North Charleston, SC: CreateSpace Independent Publishing Platform.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford, England: Oxford University Press.

Fernández, E. M. (2003). *Bilingual sentence processing: Relative clause attachment in English and Spanish*. Amsterdam: John Benjamins.

Fernández, E. M., & Cairns, H. S. (2011). *Fundamentals of psycholinguistics*. Malden, MA: Wiley-Blackwell.

Ferreira, F., & Tanenhaus, M. K. (2007). Introduction to the special issue on language-vision interaction. *Journal of Memory and Language, 57*(4), 455–459.

Forster, K., & Forster, J. (2003). {DMDX}: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers, 35*(1), 116–124.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. (2006). *Constructions at work*. New York: Oxford University Press.

Goodwin, C. J. (2003). Psychology's experimental foundations. In S. F. Davies (Ed.), *Handbook of research methods in experimental psychology* (pp. 1–23). Malden, MA: Blackwell.

Grosjean, F. (2008). *Studying bilinguals*. Oxford, England: Oxford University Press.

Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism, 16*(1), 11–21.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15*(6), 409–414.

Jarvis, S., & Pavlenko, A. (2007). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Juffs, A. (2000). An overview of the second language acquisition of links between verb semantics and morpho-syntax. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 197–227). Malden, MA: Blackwell.

Kutas, M., Federmeier, K. D., & Sereno, M. I. (1999). Current approaches to mapping language in electromagnetic space. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 359–392). Oxford, England: Oxford University Press.

LAPD Shot. (2013). *Crash blossoms*. Retrieved Feb 16, 2013, from http://www.crashblossoms.com/archives/825

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*(5), 791–824.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford, England: Oxford University Press.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 601–618). Hillsdale, NJ: Erlbaum.

Montrul, S. (2001). Agentive verbs of manner of motion in Spanish and English as second languages. *Studies in Second Language Acquisition, 23*(02), 171–206.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, England: Cambridge University Press.

Pac-Man. (2011). Jenkat Media, Inc. Retrieved from http://www.jenkatgames.com

Pickering, M. J., & Van Gompel, R. P. (2006). Syntactic parsing. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 455–503). London: Academic.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). Oxford, England: Oxford University Press.
Schachter, J. (1993). A new account of language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 32–46). Amsterdam: John Benjamins.
Sorace, A. (2010). Using magnitude estimation in developmental linguistic research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 57–72). Amsterdam: John Benjamins.
Souza, R. A. (2011). Argument structure in L2 acquisition: Language transfer revisited in a semantics and syntax perspective. *Ilha do Desterro, 60*, 153–187.
Souza, R. A. (2012). Two languages in one mind and the online processing of causatives with manner-of-motion verbs. *Revista Virtual de Estudos da Linguagem (ReVEL), 10*(6), 220–239.
Sprouse, J. (2011). A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language, 87*(2), 274–288.
Weinreich, U. (1953). *Languages in contact*. New York: Publications of the Linguistic Circle of New York.
Weskott, T., & Fanselow, G. (2011). On the informativity of different measures of linguistic acceptability. *Language, 87*(2), 249–273.
White, L. (2003). *Second language acquisition and Universal Grammar*. New York: Cambridge University Press.

Chapter 4
Rapid Serial Visual Presentation: Bilingual Lexical and Attentional Processing

Jennifer M. Martin and Jeanette Altarriba

Abstract This chapter examines the use of *Rapid Serial Visual Presentation* (*RSVP*) as a research method for studying reading and attention in bilinguals. Theoretical background and methodological considerations are provided for the most common ways in which RSVP is used: *lexical processing*, *repetition blindness* (*RB*), the *attentional blink* (*AB*), and *executive control*. The authors also describe and discuss relevant studies that have used bilingual participants, whether exclusively or in comparison to monolinguals. To date, there has been relatively little use of RSVP in bilingual research. However, this chapter provides rationale for its use as a well-controlled experimental method that is especially well-suited for use with bilinguals (whose reading speeds tend to vary a great deal). Suggestions for future research are also provided.

Introduction

Throughout this volume, the various methods used to study bilingual reading and related processes are presented along with discussions on how these methods inform theory and research in bilingualism. The focus of this chapter is on *rapid serial visual presentation* (RSVP). RSVP is a method in which letters, digits, or words are presented one at a time for a designated brief period of time. For example, Fig. 4.1, below, shows a sample experiment that uses an RSVP stream of eight items presented serially for 100 milliseconds (ms) with no pause in between (0 ms interstimulus interval; ISI). The boxes in the figure indicate what the participant would be viewing on the computer screen, proceeding chronologically from the upper left corner to the lower right corner. Stimuli may comprise a list of items or a full phrase or sentence. The greatest strength of this method is flexibility and control in manipulating

J.M. Martin, M.A. (✉) • J. Altarriba, Ph.D.
Department of Psychology, University at Albany, State University of New York,
Social Science Building-Room 399 1400 Washington Avenue, Albany, NY 12222, USA
e-mail: jmartin8@albany.edu; JAltarriba@albany.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_4

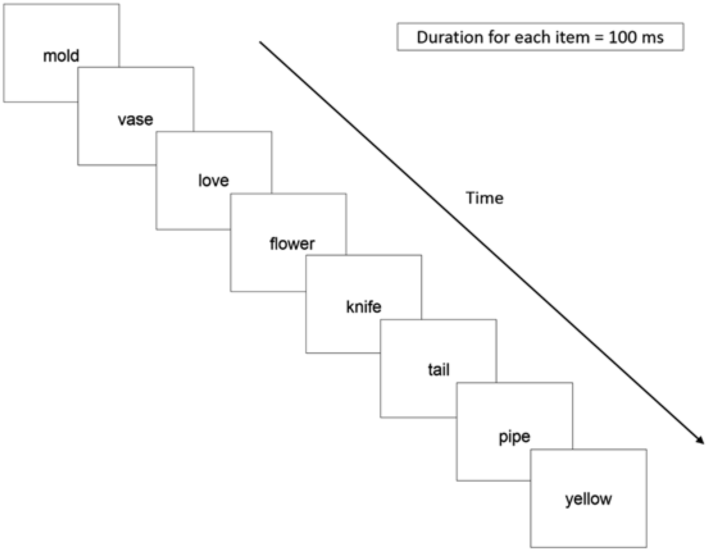


Fig. 4.1 Example of an RSVP Stream

the duration of presentation for each individual item in the series. If a full sentence or list of items were to be presented in its entirety, the relative time that the participant has to process each item could vary greatly. Reading performance can be measured in two main ways: time to read each word (or a whole passage), and comprehension of presented materials. In RSVP, time to read each word and each passage is constant for all participants so reading performance is measured by testing memory for and comprehension of presented materials. While the time it takes a participant to read a specific item can be measured using other methods, in many situations it may be preferable to directly control the amount of time the participant has to process each item. Other options could include eye-tracking (in which a camera tracks via eye movements where and how long the participant looks on the screen) or self-paced reading (in which stimuli are presented one at a time and the participant chooses when to advance to the next item). While many researchers have used one technique or another in isolation, researchers in the field of reading research have called for converging evidence amongst such techniques as RSVP and eye-tracking (Potter, 1984).

In an RSVP study, items are usually presented horizontally in the center of a computer screen. This arrangement is done given that regardless of whether the language being used is read from left to right, right to left, or vertically, most participants are faster to process words presented on the right (Smigasiewicz et al., 2010). It is best to control location and avoid a possible confounding effect on processing time by presenting all words at the center. Additionally, reading speed is fastest for horizontal text (Yu, Park, Gerold, & Legge, 2010), so this orientation is preferred over any type of rotated or vertical text. The amount of time allotted to process each item in RSVP could vary, but it is typically held constant at a short period selected for that experiment, often between 50 and 300 ms. Given the rapid nature of presentation, Potter (1984) contended that RSVP makes equal the time that the reader has to visually perceive a word for both skilled and unskilled readers, so any differences in comprehension are likely due to language processing at a more advanced stage than vision. This balance becomes important for research on bilingualism because bilinguals often vary considerably in their reading abilities and do not always reach the level of a native speaker. Additionally, even skilled bilinguals are often much slower to read in their second language (L2) as compared to their first language (L1; Favreau & Segalowitz, 1983). Using RSVP, an appropriate duration of presentation can be applied.

As previously mentioned, converging evidence gathered with eye-tracking methodology could neatly support processing conclusions from studies using RSVP. In order to facilitate this comparison within the empirical portions of this chapter, eye-tracking methodology shall be introduced here. Eye-tracking uses cameras to identify where an individual's eyes are directed and makes assumptions about what they are attending to, based on where they are looking. The main theory behind eye-tracking is called the *eye-mind assumption* (Just & Carpenter, 1980) and purports that what the individual is looking at reflects what that person is attending to (Lai et al., 2013; Rayner, 1998). In the case of reading, the individual's eye fixates (remains still) at a certain point or word, and then makes a saccade (rapid physical jump) to the next point of fixation. Eye-tracking studies follow these fixations and saccades and analyze where eyes are directed at each point in time, including initial fixations, regressions (returning to a point that was fixated on earlier), and total time, among other measures (Frenck-Mestre, 2005; Lai et al., 2013). Eye-tracking has been used primarily in monolingual populations, but there are also a growing set of studies using eye-tracking with bilinguals (Altarriba, Kroll, Sholl, & Rayner, 1996; Bartolotti & Marian, 2013; Frenck-Mestre & Pynte, 1997; Libben & Titone, 2009).

A final introductory point of interest is that RSVP has been gaining attention in media and other nonpsychology fields. Recent work in design and engineering of electronics has suggested using RSVP as a method of reading on-screen displays. Part of this work has been intended for computer screens (Beccue & Vila, 2004) and for small screens such as mobile phones, PDAs, or tablets (Öquist & Lundin, 2007). Some suggest that RSVP increases reading speed without cost, especially by cutting out "unnecessary" regressions (eye movements back to look at certain words for a second time; Taylor & Taylor, 1983), though others suggest that regressions are helpful rather than harmful or unnecessary (Schotter, Tran, & Rayner, 2014). The use of RSVP for small-screen reading is based on the limited space on the display, causing users to need to scroll frequently. RSVP is suggested as a way to avoid such hand movements by the user and let them passively watch the sentence presented at a comfortable pace. As opposed to its use as a research tool, using RSVP to improve small-screen reading would require that reading comprehension is maintained *and* that the user is subjectively happy with the experience. Some work has been supportive (Beccue & Vila, 2004; Rahman & Muter, 1999); however, several works have suggested that there are problems with this application of RSVP (Öquist & Lundin, 2007).

Öquist and Lundin (2007) asked participants to use a touch-screen mobile phone to read text selections of comparable length and difficulty using various small-screen reading formats, including RSVP. Öquist, and Lundin (2007) found that participants read text selections (counterbalanced across presentation techniques) *slower* via RSVP than when using the other small-screen formats (e.g., scrolling), though this format produced equally high comprehension as the others. Finally, the perceived task load or difficulty of RSVP was slightly greater than some of the other options. Öquist and Lundin (2007) asserted based on this test that perhaps other formats should be considered in place of RSVP for small-screen reading.

Beccue and Vila (2004) implemented a test of a customizable RSVP program on a computer screen. In this experiment, participants were allowed to adjust the color of the text and the background, the font, the size of the text, and the presentation rate. The RSVP group was compared with a control group that read the text presented on the screen all at the same time (traditional), using the default settings for the RSVP task (selected for commonness and ease of reading). The reading materials used by Beccue and Vila came from a standardized reading comprehension test used for grades 9–16 and were made up of a number of text selections with accompanying multiple choice comprehension questions. Participants read the passages using either traditional or RSVP presentation and answered the comprehension questions. Results showed faster reading speed using RSVP for the five shorter text selections used and slower reading for RSVP than for the control group for the two longer text selections, suggesting that RSVP is best suited for shorter passages. Further, no statistically significant differences were found for comprehension question accuracy for the RSVP and traditional presentation formats. However, the results of this study should be viewed with caution, as the participants in this study were university students largely in computer focused or computer-related fields and as such may have been more familiar with a variety of presentation formats that the typical reader would not have experienced. A more compelling argument could be made with support for these results in a more appropriate group of participants with typical computer experience. Finally, given that the main task in this study was reading in English and answering reading comprehension questions, it is odd that only about half of the participants were native speakers of English (50 % in the control group and 62 % in the experimental group). Reanalysis of these findings should compare the monolinguals and bilinguals in this sample, and further work should seek to replicate these findings with better measures of language history. Above all, notable researchers using this paradigm have suggested that overall, RSVP's best contribution is as a research tool, not a reading format in real life for most people, except perhaps for scanning a list for a certain item (Potter, 1984).

Before presenting research using RSVP as a research tool, it is important to address several limitations to using this method as a measure of reading and processing. A main criticism relates to the memory demands created by RSVP performance measures and the influence of memory on measuring processing. It is known that memory span for RSVP lists varies with presentation rate, with fewer items recalled at faster presentation rates (Potter, 1999). However, due to their greater degree of structure, basic sentences are easily recalled, even sentences with as many as 14 words are easily reported with high accuracy (Potter, 1984, 1999). A large majority of RSVP studies assess performance through different questions asked after the RSVP stream has been presented (various measures are discussed below). These accuracy questions take place after online processing of the sentence, so they may not be a completely direct measure of processing time (Mitchell, 1984). It is thought that using measures such as recall tasks allows for influence by other factors besides processing, such as memory storage and retrieval (Mitchell, 1984). Despite these limitations, RSVP allows researchers to examine processing time by seeing what types of mistakes are made when presentation rates are very fast. On its own, the knowledge that individuals *can* read very fast and occasionally make mistakes may be less theoretically interesting than studying *which* processes suffer at very quick presentation rates (Potter, 1999). The remainder of this chapter will present previous research using the RSVP method to study bilinguals (and monolinguals where important for understanding the bilingual work), along with recommendations for researchers and ideas for future explorations.

Lexical Processing

RSVP can be used to examine lexical (word-level) and sublexical (letter or word base) effects within sentences and lists. This method has been employed in monolingual research (e.g., Andrews & Bond, 2009) because of its helpfulness in controlling how long each participant has to process individual words and sentences. One way in which RSVP has been used to study bilingualism is through the examination of language-specific sublexical effects. Previous research in this area used bilinguals whose two languages have different language families. For instance, Wong, Qu, McGugin, and Gauthier (2011) asked Chinese-English bilinguals of varying proficiency to search for specific letters or characters in a series of RSVP items. They sought to determine if expertise in reading distractor letters (defined as other letters in the RSVP stream that were not to be reported later) interfered with a target detection task (responding to the presence of a designated target letter) as compared to pseudoletters. An example of one trial type would be highly proficient Chinese-English bilinguals searching for a specific Roman letter (e.g., *P*) amid Chinese character and pseudoletter distractors. In Fig. 4.2, participants are presented with two possible target letters (*P* and *S*). Once they advance the computer, an RSVP stream containing one of the possible target letters along with distractors, is presented for 100 ms per item. Here, the participants see a *P*, along with pseudoletters and Chinese characters. Finally, the same two possible target letters are presented, and the participants indicate via key press which of the two letters was present in the

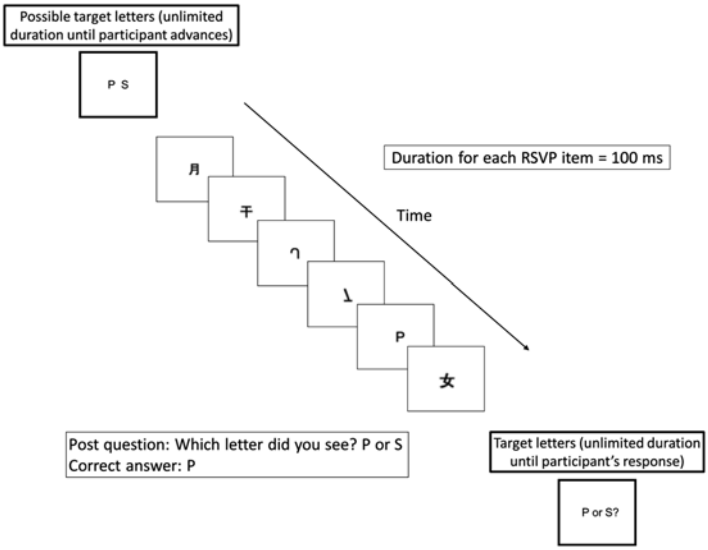


Fig. 4.2 Example of an RSVP series from Wong et al. (2011) presenting letters *P* and *S* and a target (*P*)

RSVP stream. They also asked monolingual English speakers to complete the same task. The two participant groups allowed Wong et al. (2011) to compare performance when searching for a familiar letter among all unfamiliar distractors (in the case of the English monolinguals) or for a familiar letter among mixed familiar and unfamiliar distractors (in the case of Chinese-English bilinguals). Other conditions in this experiment used each type of character (Roman, Chinese, and pseudoletter) as a target and a distractor, in counterbalanced blocks. Thus, the independent variables were language experience (English monolingual or Chinese-English bilingual), target type (Roman letters, Chinese characters, or pseudoletters), and distractor type (Roman letters, Chinese characters, or pseudoletters). The dependent variable analyzed was accuracy in target identification.

The results showed that for Chinese-English bilinguals, their accuracy at detecting a Roman letter declined when Chinese character distractors were present in the RSVP stream. This finding supported the notion that there was interference from Chinese distractors for the Chinese-English bilinguals when searching for a Roman letter. Surprisingly, when English monolinguals were asked to search for pseudoletter targets, they did not show a deficit if Roman letter distractors were present in the RSVP stream. This result is somewhat unexpected because the participants did not seem to be distracted by the Roman letters that they could comprehend any more than the Chinese characters that they could not understand. Wong et al. (2011) also repeated their experiment under a more difficult condition. In this case, participants were required to increase the load on their memory by rehearsing four digit sequences. Wong et al. obtained similar results under these conditions as they had when no additional memory task was required. The results of the follow-up experiment reduce the likelihood that the effects of searching for unfamiliar pseudoletters in Experiment 1 were produced because the task required too much focus and attention (leaving no resources to be distracted by the Roman letters). Wong et al. concluded that a common expert perceptual system for recognizing known letters is being used for both Chinese characters and Roman letters for the Chinese-English bilinguals. This means that they rejected the conclusion that recognizing symbols relies on working memory and support a model in which perceptual expertise can operate under memory load, but this expert perceptual system is used when each of a bilingual's known symbolic systems is being processed. However, the English monolinguals are only tapping this resource to process the known Roman letters. It was argued that while Roman letters hold no meaning in isolation, Chinese characters represent words, which could have an influence on this study. The authors make no further reference to this issue; however, replication of this work using languages with different alphabets, such as the Cyrillic alphabetic system and the Roman alphabetic system (whose symbols represent the same linguistic level; that of letters, not words) would add credence to the conclusions drawn here. Additionally, though Wong et al. measured the average presentation rate for each block at which acceptable accuracy could be maintained, it would be interesting to also compare reaction time for each response across condition. The time to choose the correct option from the two target letters may also be a useful measure of processing difficulty for this task.

Other research has used Hebrew-English bilinguals (Velan & Frost, 2011) as a test case of languages with different base morphologies (i.e., internal structures). Most English words have a base unit such as *drive* which is the base for other words such as *driving*, *driver*, *drove*, and *driven*. In contrast, Hebrew units of meaning are not orthographically near in distance. For example, the root *z.k.r* is used to create words that reference the concept of memory and is used to create *zikaron* ("a memory"), *mazkir* ("a secretary"), and *hizkir* ("reminded"), among others (Velan & Frost, 2011). These Hebrew words share the root concept of memory and the letters *z*, *k*, and *r*, but they appear in various positions relative to each other with different intervening letters.

Velan and Frost (2011) used Hebrew words with root patterns as described above, along with other Hebrew words that had been borrowed into Hebrew from languages with unified bases similar to English in an RSVP task. Thus, the authors created two groups of target words that either had Hebrew root patterns or English-like base words. Hebrew-English bilinguals completed a report task in which they viewed Hebrew or English sentences and repeated the entire sentence after the final word had been presented. The target word in half of the sentences had two letters transposed, but the participants were asked to report the word as if it had not been altered. Results were analyzed for accuracy in reporting the target word. The target words that were English showed high accuracy of report even for words with transposed letters, while the Hebrew target words with transposed letters showed a decrease in accuracy consistent with previous work (Velan & Frost, 2007). However, Hebrew words with English-like base roots were recalled with high accuracy similar to English words. These results indicated that the fully Hebrew words with root patterns instead of base roots did not produce similar effects as previous work with other Indo-European languages, such as English and French. Letter transposition more strongly impairs processing of words with root patterns like in Hebrew than words with base words like in English. It is important to know of such cross-language differences because many experiments study effects of orthography or spelling on other processes. Given this, future research that manipulates orthography should be extremely mindful of language families, particularly Semitic languages, and seek to create frameworks that can take into account languages with root patterns as well as those with base words. RSVP is a useful tool for studying sublexical processing of isolated letters (Wong et al., 2011) and of orthographic effects (Velan & Frost, 2007, 2011) because small differences can be found by tightly controlling the amount of time in which a participant can process each item. The limited time allows very small differences in processing time and accuracy across trial conditions or participant groups to be revealed.

RSVP is also an excellent medium for examining sentence context effects in word comprehension. Altarriba, Carlo, and Kroll (1992) conducted one of the first studies to examine the effects of language dominance and sentence context on bilinguals' word naming. The bilingual participants were Spanish dominant and proficient in English, their second and nondominant language. Altarriba et al. used sentences containing interlingual homographs, which are words that are spelled the same in two languages, but mean different things (e.g., *fin* is a fish part in English and means *end* in Spanish). These target words were embedded in sentences that were biased towards one of the interpretations of the homograph (e.g., *We knew the play had reached its FIN when we saw the curtain fall*, is biased towards the Spanish *end* interpretation of the homograph *fin*). Controls were created for both English and Spanish sentences that did not fit the meaning of the sentence in either language. Participants viewed each sentence presented via RSVP and named the target word (written in all capitals, e.g., *FIN*) aloud.

Altarriba et al. (1992) analyzed the time it took participants to name each word under the different contextual and language conditions. The homograph was either easily predictable by the sentence context or unexpected based on sentence context. Additionally, the meaning of the homograph either matched the language of the sentence context or did not. Results indicated that Spanish-English bilinguals were slower to name ambiguous homographs whenever they appeared in English (i.e., their nondominant language) than control words, regardless of whether the sentence biased the Spanish or English interpretation of the homograph. However, when naming words appeared in Spanish (i.e., their dominant language), participants were slower to name words that were biased towards the meaning of the homograph in English and showed no slowdowns for homographs consistent with the Spanish sentence context. The conclusions of this study suggest that bilinguals access the meaning of ambiguous words in both of their languages when working in their nondominant language. However, when bilinguals are reading in their dominant language, ambiguous words are only accessed in their dominant language unless the context strongly suggests a word meaning from their nondominant language.

This finding is noteworthy due to the uneven interference of the dominant and nondominant languages and implies that both language dominance and surrounding context are important for understanding how bilinguals access the meanings of words. The results taken together generally provide support for nonselective access of bilinguals' two languages; that is, both languages are activated even when they are only using one at a time. Nonselective access is a critical feature of an important connectionist model of bilingual word recognition, the Bilingual Interactive Activation Plus Model (BIA+; Dijkstra & Van Heuven, 2002). This model includes four levels of nodes: letter features, letters, words, and language (one node for each of a bilinguals' two languages). One example of the implications of this model is that because cognates share letters, they are activated whenever those letters are presented without regard for which language is the context until the final level of processing is reached (Basnight-Brown, 2014). In the case of interlingual homographs, at the final level of processing, the selection of the meaning of the word in just one language as a final interpretation takes some additional time (interference). It does seem, however, that even proficient bilinguals have more spillover from their dominant to their nondominant language. It should be noted that, while the participants were proficient in both languages, it may be a stronger argument if these findings were replicated in highly proficient speakers of both languages who are still dominant in one. This would provide further support for conclusions based on dominance separate from proficiency. Additionally, this effect could be larger or smaller when tested with two languages that share fewer homographs, so that homographs are less often expected. Further work in this area could test these empirical questions.

Based on the principle of converging evidence using multiple methods, Altarriba et al. (1996) advanced the question of sentence constraint and its interaction with lexical features by conducting a study testing the effects of sentence constraint and word frequency. The study investigated two possible hypotheses. The first is that sentence constraint is a conceptual variable and would act similarly in either of a bilingual's languages. The alternative hypothesis would support the notion that sentence constraint is also acting at a lexical level. This would mean that constraint could interfere with processing a word that represents the same concept, but in the context inappropriate language. Altarriba et al. conducted two experiments with the same sentence stimuli, employing eye-tracking in Experiment 1 and naming of a target word in a sentence presented via RSVP in Experiment 2. In each experiment, Altarriba et al. manipulated the language of the sentences (Spanish or English), language of the target word (Spanish or English), sentence constraint (low or high), and the frequency of the target word (low or high). For example, a low constraint sentence with a low constraint target word (underlined) would be *The market had a new variety of pumpkin/calabaza in the fall*. The dependent variables collected were eye-tracking data (Experiment 1) and naming of the target word (Experiment 2). The participants for each experiment were Spanish-English bilinguals (Spanish L1, highly proficient in both languages).

The results showed consistent patterns across the eye-tracking and time to name the word in RSVP. Beginning with eye-tracking: low-frequency Spanish words were initially fixated on for less time when they appeared in high-constraint sentences than in low-constraint sentences, and for high-frequency words the reverse was true. This implies that it is easier to read low frequency words when they are highly constrained by context. For the English sentences, the bilingual participants performed like monolinguals: they looked at low-frequency, low-constraint words for more time than they did at low-frequency, high-constraint sentences, high-frequency, low-constraint sentences, and high-frequency, high-constraint sentences. These differences also appeared in the data from Experiment 2 in which participants were slower to name the target words that had been fixated on for a longer time in Experiment 1. The consistency of the results of two tasks using two methods with the same sentences provides excellent converging evidence that the processing in RSVP tasks is similar to that in eye-tracking studies. This study provides theoretical support for the effect of frequency on word processing and the notion that it interacts with sentence constraint.

Schwartz and Kroll (2006) extended the findings of Altarriba et al. (1992) and Altarriba et al. (1996) by examining the effects of different sentences on word comprehension for Spanish-English bilinguals of varying proficiency. The sentences that they presented using RSVP were either of high constraint or low constraint. Low-constraint sentences have an element (a word, phrase, or clause) that could be easily and plausibly filled in by a number of different options. In contrast, a high-constraint sentence could only be completed by one or two plausible options. For example, the high-constraint sentence, *Leslie bit into the juicy red…* is likely to be completed with fewer options such as *apple* or *grape*. The low-constraint sentence, *Leslie bought a nice*, *red…* has many possible options to complete the sentence, such as *apple*, *coat*, *truck*, *ball*, *purse*, *bicycle*, *vase*, or *shirt*. Schwartz and Kroll constructed high- and low-constraint sentences in which a target word appeared partway through the sentence. These sentences were matched on number of words, syntactic complexity, and length of the word that immediately preceded the target word. The cloze probability of each word was normed on a set of separate participants. The probability of the blank being filled in by the target word was approximately .66 for the high-constraint sentences and .04 for the low-constraint sentences, and did not differ across the target word conditions. The target words included, (1) cognates (e.g., *band* vs. *banda*) or words that are orthographically and/or phonologically highly similar in the different target languages; (2) cognate controls of similar English word frequency and word length to the cognates (e.g., *pencil*); and (3) interlingual homographs similar to those used by Altarriba et al. (1992). Finally, the fourth type of word was composed of matched homograph controls of similar English word frequency and word length to each homograph (e.g., *frog*). Participants viewed the sentences presented via RSVP and named the target word (printed in red) out loud into a microphone (see Fig. 4.3 for an example). Experiment 1 tested Spanish-English bilingual participants who were highly proficient in both languages and used both on a daily basis. Experiment 2 used Spanish-English bilinguals (and some Spanish-English-Valenciano trilinguals) who were more proficient in their L1, Spanish, and of moderate proficiency in English.

Cognates were named faster than matched controls by participants in both experiments when sentences were of low constraint for the target word; however,

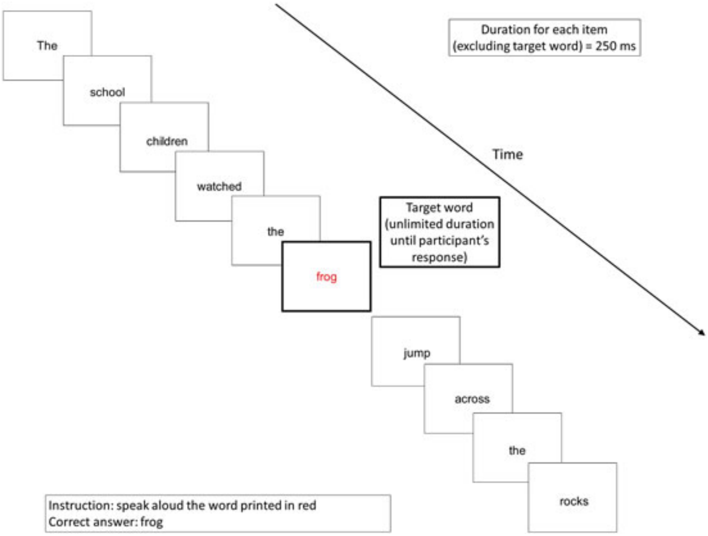


Fig. 4.3 Example from Schwartz and Kroll (2006) of a low-constraint control sentence containing the target word *frog* in *red* color

this effect did not appear for sentences that highly constrained the target word. Interlingual homographs did not differ from controls for the highly proficient bilinguals, though error rates for homographs were slightly higher than matched control words for the less proficient bilinguals in Experiment 2. This means that the degree of constraint of the sentence context affected the relationship between cognate status and naming time. The results were interpreted by Schwartz and Kroll (2006) to support the effects of linguistic context (in this case, sentence constraint) on bilinguals' lexical access as is posited within the BIA+ model proposed by Dijkstra and Van Heuven (2002) described above. The letters shared by the cognates are processed in both languages nonselectively and so the additional activation of both cognates at the word-level means that the final level of languages is reached more quickly. Selecting just one of the cognate word representations does not impair interpretation because cognates share semantic/conceptual representations. However, one would expect that interlingual homographs would be named more slowly because their letters are shared, but word meaning is not. Further work should examine this difference, perhaps by using multiple techniques such as RSVP and eye-tracking.

Van Hell and de Groot (2008, Experiment 3) sought to extend the study conducted by Schwartz and Kroll (2006) by using Dutch-English bilinguals and varying word type within the cognates, homographs, and controls. Words in these categories were divided into two further word types: abstract and concrete. Concrete words describe things that exist in the physical world (e.g., *chair* or *cat*), whereas abstract words describe things that do not exist physically (e.g., *hope* or *heritage*). The sentences used by Van Hell and de Groot (2008) were similar to those of Schwartz and Kroll (2006), including high- and low-constraint sentences. Sentences were presented via RSVP with dashes taking the place of the target word during

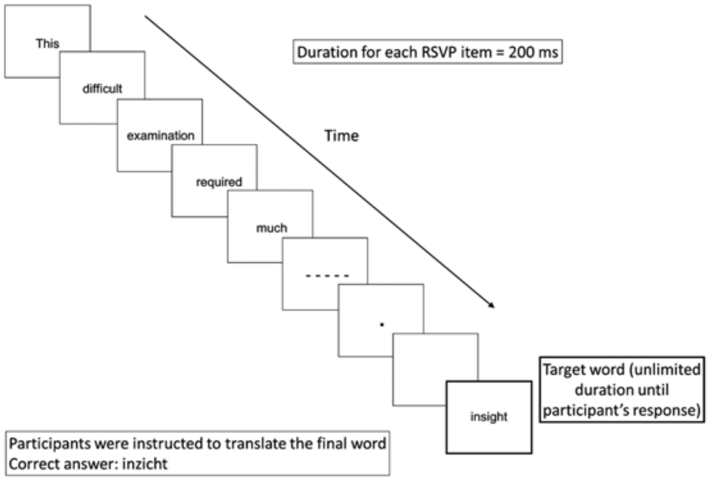


Fig. 4.4 Example of a low-constraint sentence from Van Hell and de Groot's (2008) translation task

presentation of the sentence, and then the missing target word was presented at the end for the participant to translate as quickly and as accurately as possible into a microphone (see Fig. 4.4 for an example). Van Hell and de Groot (2008) found that cognates were translated faster than controls when presented within a low constraint sentence only, and that while concrete words were translated faster than abstract words, this facilitation did not depend on whether the word was a cognate. Thus, the results of this study by Van Hell and de Groot (2008) using a translation paradigm support the conclusion that sentence context is important for recognizing (and translating) words, conceptually replicating the general findings of Altarriba et al. (1992), and Schwartz and Kroll (2006). However, the results from Van Hell and de Groot may not necessarily be fully consistent, as the translation method used may have involved strategic processing.

In this section, RSVP has been shown to effectively aid in the development of psychological research concerning lexical processing, both at the lexical and sublexical level. Some of these effects have converged with other methods with different strengths, such as eye-tracking. Future work should use RSVP to examine other possible influences on bilinguals' ability to access words in their various languages, such as the language of the sentence context, regular and unusual word spellings, or priming effects, in addition to working to provide more converging evidence with other methods.

Attention

In addition to being an avenue for studying lexical processing, the RSVP paradigm is also useful when examining attention. Cognitive psychologists use many different techniques to answer questions about how humans are able to attend to and take in information. When employing the RSVP methodology, an interested researcher will necessarily be focusing on visual attention. Huang and Pashler (2012) define visual attention as *…the selection of part of the available visual information, so that this part is admitted to consciousness* (p. 30). Thus, the information to be processed must first be "available" to the individual (seen in the case of visual attention) and then selected for out of all the available options. Scientists study the ways in which specific information is selected for and reaches consciousness while other information is missed.

Typically, in this area of research, participants are asked to read a series of letters, words, or a phrase or sentence that is presented sequentially using RSVP (see Fig. 4.1 for an example of a word series). Performance measures, such as those discussed below that probe the participants' accuracy at perceiving and comprehending the series, allow researchers to gain insight into their attentional processing. Patterns of errors can often be particularly illuminating, as shall be seen below through an in-depth look at two well-studied attentional effects: *repetition blindness* (RB) and the *attentional blink* (AB), and how they have been used to show possible differences between the attentional abilities of bilinguals and others.

Performance Measures

One way to measure performance in an RSVP task is to ask participants to reproduce the stimuli that had just been presented in that series. Using the example in Fig. 4.1, the participant would report, *mold*, *vase*, *love*, *flower*, *knife*, *tail*, *pipe*, *yellow* in the original presentation order. In the usual case, a microphone is used to record the participant's voice as they verbally repeat the sentence or list. A scoring sheet may be used by an experimenter following along during the task as a check in case of poor recording quality, or scoring may take place only at a later point in time using the audio tape. Performance is then considered in terms of how many items/words the participant remembered and also whether they were reported in the correct order.

Another performance measure is the report or the detection of a target item. The participant will respond yes/no as to whether an item was present in a series (see Fig. 4.5). The execution of a *detection task* has several variations. The targets may be specified to the participant before the trial begins so that they are aware of what they are to be looking for. For instance, they may be told that they will be asked whether or not there was a word with a certain characteristic (e.g., a word printed in a certain font color, an emotion word, an animal name, a color word [e.g., *yellow*] or a specific word [e.g., *red*, *happy*, or *elephant*]). In other cases, the participant may see the whole series before they are told what the target was and then are asked to respond. It is more common for the target(s) to be specified prior to presentation of the series, and this method tends to increase accuracy and be easier for participants. In cases in which the researcher wants to ensure that the participant processes all items (including distractors), they may want to use the technique of giving the target after the RSVP stream has been presented.

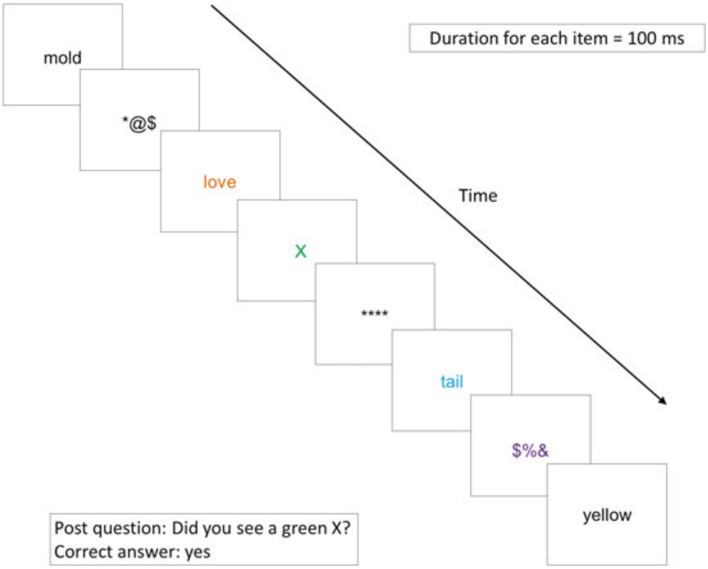


Fig. 4.5 An RSVP stream with a target detection post question regarding the target *X* in *green* color

Participants can also make responses about target identity rather than simply target presence (see Figs. 4.3 and 4.7 for examples). In this case, the participant would be asked about the identity of an item with a certain characteristic or in a certain position (e.g., *what word was printed in red*? or *what word came directly after the green X*?). Again, they may be told before the trial about what to be looking for. Researchers using RSVP to study attention should select the performance measure that best suits their research questions. It is important to note that it is certainly possible to use several of these tasks in combination, with several targets that may be similar or different. The empirical work in the following sections uses variations on these performance measures to examine important effects such as repetition blindness and the attentional blink.

Repetition Blindness

Repetition blindness (RB) refers to the effect in which participants show low accuracy in reporting the second instance of an item within one RSVP stream (Kanwisher, 1987). The effect was first shown where the second target was identical to the first. For example, when presented the sentence: *Sally wanted the <u>ball</u>, but the <u>ball</u> was lost*, participants would be less accurate at reporting the second instance of the word *<u>ball</u>*, as compared to the first instance (see Fig. 4.6). The second instance of the target word may appear 300 ms after the first one, for example. Control conditions

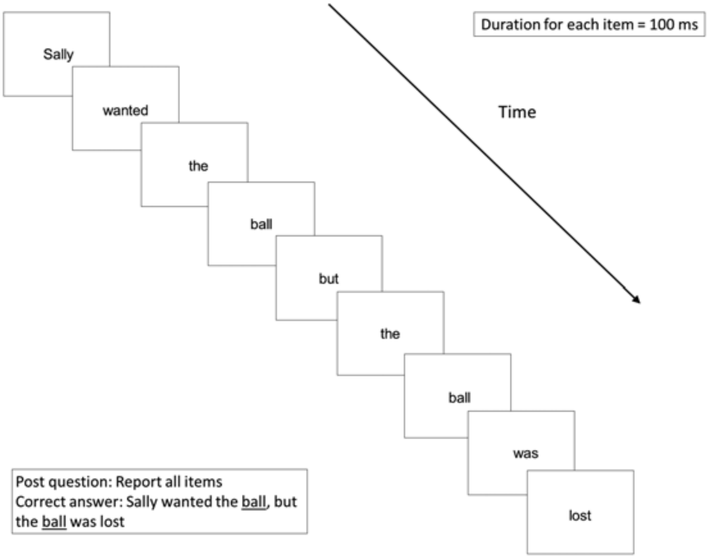


Fig. 4.6 An RSVP stream that may show repetition blindness for the second instance of the target word (*ball*)

without repeated words are also used as a comparison for accuracy. These control conditions often substitute a logically consistent word for one of the instances, such as: *Sally wanted the ball, but the bat was lost*.

RB is generally defined as the deficit in accuracy of reporting the target word when it is repeated, as compared to the control word substitution (MacKay, Hadley, & Schwartz, 2005). There are certain procedural requirements to produce a true RB effect. In terms of presentation limits, the RSVP stream must be presented for no more than 200 ms per item as slower presentation rates reduce or eliminate the effect (Coltheart & Langdon, 2003). Additionally, the second target must appear less than 500 ms after the first target. However, limited presentation time or delay between targets is not the only procedural element required to produce an RB effect. Whittlesea and Masson (2005) conducted a study with several different types of distractors between targets. They used presentation durations that had previously produced RB, but instead of using words as distractors, they used conditions with symbol strings (e.g., *&^%*&^), and with all fillers being the word *white*. The timing of each item in the RSVP stream remained constant across all these conditions, but performance was much better in the symbol strings and *white* filler conditions than when the fillers were random words. With distractors that lack meaning (i.e., symbol strings or repetitions of the same word), noticing the repeated words was too easy. These results imply that meaningful distractors are necessary to obtain an RB effect.

Once the effect was established, additional work sought to find the limits of how far it would generalize. Several theories variously emphasized whether repetition blindness would be experienced when the two targets matched on lexical, phonological, orthographic, or semantic characteristics. Evidence for orthographic-based repetition blindness comes from studies that use targets that are orthographically similar, but phonologically and semantically dissimilar (Bavelier, Prasada, & Segui, 1994; Kanwisher, 1991). Stimuli from such studies could substitute one or more letters that change the phonology of the word (e.g., *change-charge*, *rod-rid*) or use subparts of the word that are phonologically distinct when presented in isolation (e.g., *chunk-hunk*, *with-wit*). However, pronounceable nonwords (e.g., *gerb* or *sath*) that are orthographically plausible do not show RB and in fact show priming effects (Coltheart & Langdon, 2003). Thus, while orthographic RB is a replicable effect, it may require real lexical stimuli. In a study demonstrating RB effects for Chinese characters, Yeh and Li (2004) also showed RB for nonidentical characters that shared sublexical repeated components (i.e., a system of lines in the same location and orientation in two different characters that do not constitute a complete character on their own). RB effects have also been obtained for word pairs that match phonologically, but not orthographically (Bavelier & Potter, 1992). This effect has been shown in several forms (e.g., *6-six*, *I-eye*, *fancy-FANCY*). An interesting way to further test the RB effect is to use a cross-modal type task, similar to that used in priming work (e.g., Stewart & Heredia, 2002; see also Chaps. 6 and 10), to present one of the repetitions in the opposite mode (i.e., a spoken word inserted in a printed RSVP sentence). Obtaining RB even when the repeated word was presented in two different modes would provide evidence that physical form is less important than phonological or conceptual similarity, though Soto-Faraco and Spence (2002) did not find evidence of cross-modal RB. This methodology could also be applied to bilingual research. In summary, RB effects are found with stimuli that occur less than 500 ms apart, are separated by meaningful distractors, and can be found with repeated orthographic elements.

Some semantic representations, such as emotionally arousing words like taboo or sexual words, may capture attention more so than neutral words and change patterns of RB. When using taboo stimuli, MacKay et al. (2005) found that if the second target was a taboo word, there was a reduction in RB; however, if instead, the first target was the taboo word, there was an increase in the effect. The results of this experiment mean that regardless of the position in which they appeared, taboo words drew the attention of the participants. This heightened attention increased the processing of that word and decreased the processing of other surrounding words. Silvert, Naveteur, Honoré, Sequiera, and Boucart (2004) examined whether a category united by an affective semantic category like negative emotional words (e.g., *war* or *murder*) would produce a greater repetition blindness effect than a neutral semantic category like animals, or neutral words not related by a semantic association. Silvert et al. (2004) supported their hypothesis that negative targets would produce a greater difference between streams that contain targets and controls. This suggests that RB is decreased for negative emotional words in comparison to neutral words.

Knickerbocker and Altarriba (2013) more clearly specified RB effects for emotional stimuli by showing that words that describe an emotional state (e.g., *love*, *hate*) created significantly greater repetition blindness than either neutral words or emotion-laden words that evoke an emotion while not describing it (e.g., *puppies*, *cancer*). These results imply that emotion and emotion-laden words are separate categories of words that capture attention differently. Other work using emotional words should thus be conscious of this distinction. In a different semantic manipulation, Arnell, Shapiro, and Sorenson (1999) showed that under a variety of task conditions, individuals consistently showed a reduced RB effect when their own name was the target as compared to conditions where another name was the repeated target. The results here suggest that increased meaning and familiarity for one's own name in comparison to other common names can capture attention even when repeated. Research investigating semantic repetition blindness using monolinguals supports the notion that some types of semantic items such as words that describe emotion states, words that evoke emotions, taboo or sexual words, or something as simple as one's own name actually reduce the RB effect.

However, the most interesting examinations of semantically based RB show the valuable use of bilingual participants in research. In a number of studies, proficient bilinguals were used to test semantic RB with translation equivalents (e.g., *horses* and its Spanish translation *caballos*) instead of the within-language synonyms that had been previously used (e.g., *mad* and its synonym *angry*). MacKay and Miller (1994) used cross-language translations as a stronger test of semantic RB. They argued that although within-language synonyms like those that Kanwisher and Potter (1990) had used did not produce RB, cross-language translation equivalents would produce RB. MacKay and Miller (1994) theorized that synonyms are never exactly equivalent in meaning and may have restricted finding a true semantic RB effect. Thus, MacKay and Miller used sentences that mixed Spanish and English and contained several types of target pairs: exact matches (e.g., *duck-duck*), translation matches (e.g., *duck-pato*; *pato* is the Spanish translation of *duck*), and non-matched cross-language pairs (e.g., *duck-vaso*; *vaso* is the Spanish translation of cup or glass). An example of a sentence with a translation-matched word was, *They saw horses, but caballos were prohibitir* [sic] *to enter*, with *They saw sheep, but caballos were prohibitir* [sic] *to enter* serving as a control sentence. The sentences were mixed in terms of language and presented at 70 or 90 ms per word. The target pairs were separated by one to two words. The performance measure used here was full report and showed an RB effect for the translation pairs. The authors concluded that cross-language translation pairs can produce semantic RB. Cross-language RB would support the notion that words have very similar concepts even when translated. This conclusion is based on common explanations for RB that differentiate between "type" concepts and "token" instances (Kanwisher, 1987). When words are repeated in an RSVP stream, the concept (type) of the word is noticed and remembered, but individuals do not remember two separate instances (tokens) of that concept appearing. If cross-language translation pairs produced RB, it can be concluded that the translations were understood to be the same concept, even though their physical spelling is different. Data supporting cross-language RB would imply a strong conceptual similarity for translations across different languages.

Altarriba and Soltano (1996), however, argued that MacKay and Miller's (1994) results were driven by methodological issues. Their main criticisms targeted the low overall performance by the participants and argued that it was driven by the nongrammatical mixing of languages in the sentence stimuli. Proficient bilinguals do engage in language-mixing or code-switching (Heredia & Altarriba, 2001); however, true code-switching has consistent rules. A language switch often occurs at the boundary between two phrases or subparts of the sentence that are relatively independent. For example, a proficient bilingual might say, *My car is broken*, *entonces monto en bicicleta* (which translates as *My car is broken*, *so I ride a bicycle*). The example is a sentence that retains grammaticality in the appropriate language: English in the first phrase and Spanish in the second phrase. MacKay and Miller's (1994) materials often switched languages within a phrase instead of between phrases, and also violated grammaticality within a phrase. For example, *They saw horses*, *but caballos were prohibitir to enter*, includes an incorrect spelling (present tense of prohibit in Spanish is *prohibir*, not *prohibitir* which in fact is not a real Spanish word) and also changes languages at locations other than phrase boundaries. An appropriate correction could be, *They saw horses*, *pero caballos eran prohibidos a entrar*. Additionally, *caballos were prohibitir to enter* translates as *horses were to prohibit to enter*. A more appropriate translation would use "prohibidos" to replace the English "prohibited." These violations of commonly followed language-mixing conventions and of Spanish grammar may have been difficult for participants to follow and led to the overall low accuracy.

Given these critiques, Altarriba and Soltano (1996, Experiment 1) created sentences with more naturalistic switch points, but retained the same word pair types as MacKay and Miller (1994) in addition to new cognate (e.g., *imagination-imaginación*) and noncognate translations. Altarriba and Soltano (1996) also used a full report as a performance measure. There was no repetition blindness effect obtained across languages, in contrast to the within-language identical pairs. In Altarriba and Soltano's Experiment 2, lists of three mixed-language words were presented interspersed with symbols in an RSVP sequence in order to create a condition with lower working memory load. The translation pairs with the word lists that were noncognates actually showed increased recall, or facilitation, which was unexpected based on MacKay and Miller (1994). These pairs showed 13 % better recall than unrelated mixed-language pairs, implying that reading the translation of a word in another known language *facilitates* attending to it later in an RSVP stream (cf. Sánchez-Casas, Davis, & García-Albea, 1992).

MacKay, Abrams, Pedroza, and Miller (1996) disputed some of Altarriba and Soltano's (1996) results, based on several procedural elements (see MacKay et al., 1996). These procedural choices included the unequal word length between the first and second targets and the choice to use base recall rates instead of conditional recall. Additionally, language switches were predictable because languages were switched at phrase boundaries. This predictability could be potentially problematic if it increased the salience of translations in the sentences used in Experiment 1 and reduced repetition blindness. Mackay et al. later attempted to address some of these issues specific to Experiment 2 with improved procedures (MacKay, James, &

Abrams, 2002). An additional manipulation included first targets with either many possible translations or few possible translations. MacKay, James, and Abrams (2002) found cross-language facilitation (improved performance) for the pairs that included a first target with many possible translations, but this effect did not extend to cross-language pairs in which the first target had only one possible translation. They concluded based on their conceptual replication of Altarriba and Soltano (1996) that cross-language semantic repetition blindness rarely occurs and was possibly caused in MacKay and Miller (1994) by ungrammatical language mixing in their sentences.

The specific procedural choices made by each set of authors were instrumental in narrowing down the restriction of semantic blindness effects to salient words (i.e., emotion words or one's own name, as noted above) and not merely typical word equivalents such as within-language synonyms or cross-language translations. The evidence so far does not support a general semantic RB across languages and implies that, though translation pairs share some conceptual basis, they are still maintained as separate instances in one's mental representations (unlike exact repetitions within one language). While word translations are based on similar conceptual representations, the different physical and lexical properties of the words seem to limit the RB tendency to lump two instances under the same conceptual type at the expense of storing two separate instances.

The theory and work presented in this section dealing with RB effects in monolinguals and bilinguals highlight similarities between word pairs. Within monolinguals, repetition blindness can occur for words with similar orthography and vary in relative size based on semantic factors. Bilingual research using RSVP methodology in this manner allows for strong tests of semantic RB effects based on shared concepts for the two targets, but different orthographies (i.e., different languages). Studies to date conclude that semantic blindness does not tend to occur in cross-language translation equivalents in sentences and that facilitation is actually shown when the words appear in lists. Future research should examine whether different semantic relationships, such as names or emotion words that have been shown to reduce repetition blindness in some monolingual contexts, are possible to extend in a bilingual context, in either sentences or lists.

The Attentional Blink

Another predominant use of RSVP in attention research is the *attentional blink* (AB) paradigm. In this method, a series of stimuli (e.g., words, letters, digits, or a mix of both) are presented quickly one at a time. In Fig. 4.7, one can see an example of possible AB stimuli. In the figure, *lag* describes the relative delay of the second target (T2) after the first target (T1). The performance measure in these experiments is a target monitoring task. In this task, participants are asked to detect or identify two specific targets within the series. When the second target (e.g., *knife* in Fig. 4.7) is presented closely after the first (*love*), participants are less accurate at detecting

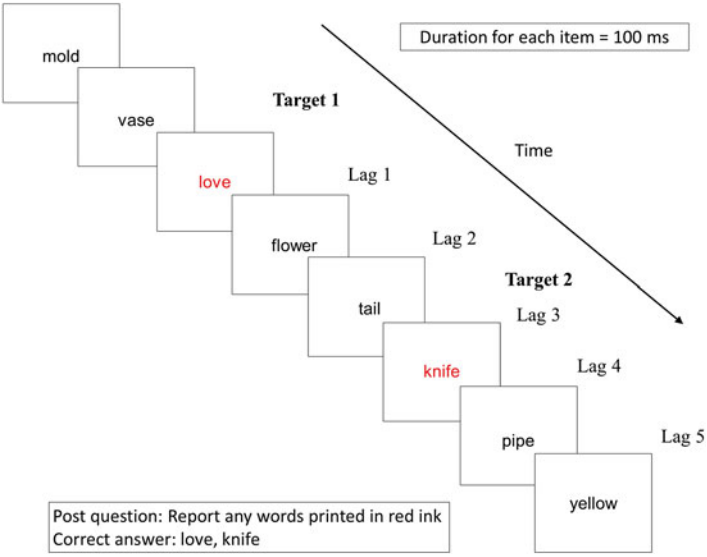


Fig. 4.7 An RSVP stream for the attentional blink effect with targets 1 and 2 (*love*, *knife*) in *red* and a 300 ms interval between targets (Lag 3)

and reporting the second target, which is referred to as the AB. The name references a metaphorical eye blink in attention, causing something to be missed.

AB is revealed when Target 2 (T2) and Target 1 (T1) accuracy are compared across trials with various time delays between T1 and T2. There are specific temporal limits to the AB. The relative position in the RSVP stream, and hence the time delay, is referred to as lag. Lag is defined in reference to the first target. That is, there is a lag of two items between T1 and T2. Most previous research has used a presentation rate of 100 ms per item in the RSVP stream, though the rate can vary slightly (MacLean & Arnell, 2012). Thus, an item presented at Lag 1 is the next item after the first target (T1) and appears 100 ms after T1. An item at Lag 2 is two items after T1 and appears 200 ms after T1, etc. When the first target appears (T1), if the second target (T2) appears more than 200 ms, but less than 500 ms after the T1, individuals are less accurate at reporting T2 (see Fig. 4.8). Historically, this effect was first reported by Broadbent and Broadbent (1987), and the term was coined by Raymond, Shapiro, and Arnell (1992). Additionally, the effect has been consistently replicated (see Dux & Marois, 2009 for an extensive review). AB has been used to study attentional processes and conscious perception.

AB is an attentional deficiency that is not merely a limit of sensation, but one of conscious perception. The attentional nature of this effect has been shown through carefully controlled conditions, such as in Raymond et al. (1992). In this study, an RSVP task was presented with two targets within the stream, as is typical in the AB paradigm; however, participants were given instructions to *ignore* the first target (T1). The results revealed that participants showed high accuracy for reporting the

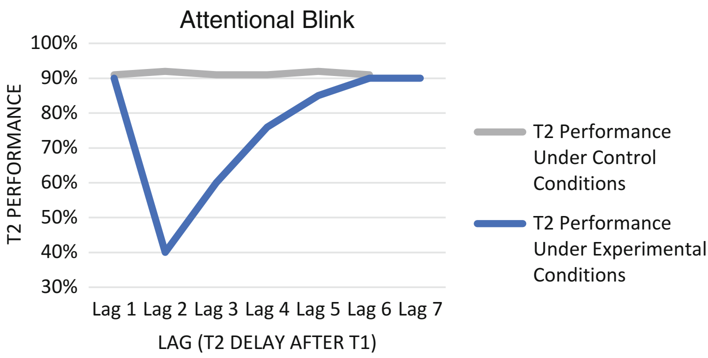


Fig. 4.8 Theoretical time course of AB. Adapted from "A conceptual and methodological framework for measuring and modulating the attentional blink," by M. H. MacLean and K. M. Arnell, 2012, *Attention*, *Perception*, & *Psychophysics*, *74*, p. 1082, Copyright 2012, by Springer

second target (T2). When an identical series was presented to other participants with the instructions to attend to and report the first target as well, accuracy was severely reduced for T2. All elements remained the same from control to experimental trials, with the exception of whether the T1 was purposefully attended to or not. If AB were based only in sensory limits, T2 would still have reduced accuracy whenever T1 was present, even if T1 was ignored. However, empirical results from Raymond et al. (1992) showed that T2 accuracy was excellent when T1 was ignored, so it is possible to conclude that the AB is caused by attentional limits (limitation of the ability to direct attention to the second target soon after the first), not by sensory limits (if sensation were limited, it would be impossible for participants to ever see the second target regardless of task instructions).

Theoretically, a general selective attention model is most useful. When the first target (T1) appears, the individual must attend to it; however, this selection takes up much of the available attentional resources from later targets. Additionally, the process of inhibition (of further distractors) magnifies this effect. The participant prioritizes processing the target and blocks subsequent stimuli from entering further processing. A model that is only based on limited-resource notions of attention cannot account for situations in which AB is diminished through experimental manipulations.

Olivers, van der Stigchel, and Hulleman (2007) constructed an experimental condition in which more than two targets appeared in succession (see Fig. 4.9 for an example with four consecutive targets) and resulted in strong attenuation of AB. The same amount of resources would have been dedicated to the first target; however, since the next items in the RSVP task were targets instead of distractors, the subsequent targets were not inhibited. Also, Di Lollo, Kawahara, Shahab Ghorashi, and Enns (2005) defined their target as a string of three digits (each presented one at a time) and also showed reduced AB as compared to trials in which the string was interrupted by a distractor (see Fig. 4.10 for an example with the target defined as three specific digits presented consecutively). This result suggests that it

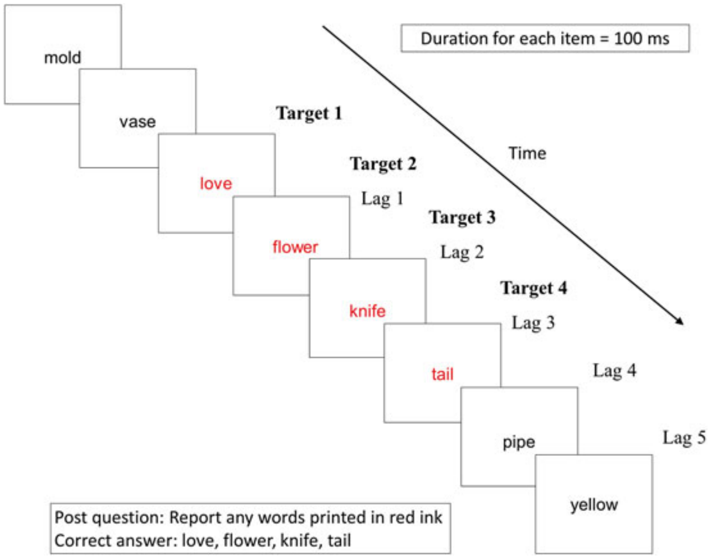


Fig. 4.9 An RSVP stream from Olivers, van der Stigchel, and Hulleman (2007) measuring AB for consecutive targets, *love*, *flower*, *knife*, *and tail* (in *red* color)

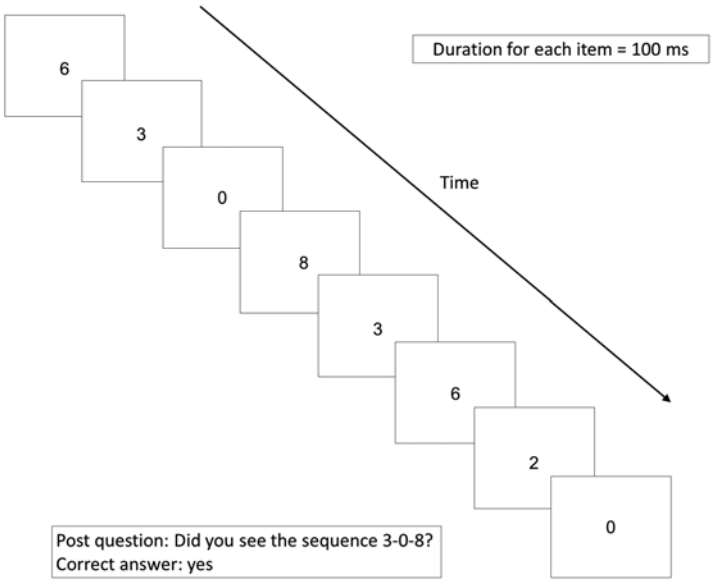


Fig. 4.10 Example from Di Lollo et al. (2005) of an RSVP stream of digits with the target defined as a three digit sequence (*3-0-8*)

is certainly possible to continue attending to multiple items in an RSVP task, but this only occurs when the items are targets or part of a target sequence. To explain such data, one must also include the selective and inhibitory processes that are initiated in reaction to the instructions to report certain stimuli only (the targets). If another target occurs within the time window of these processes (i.e., from 500 to 600 ms), it is less likely to be reported.

The AB effect is found by comparing T2 report performance at various delays after presentation of T1. At least two targets are necessary along with distractor items; however, it is also possible to use more than two targets. In order to analyze whether AB has occurred, there must be at least two trial types that vary the T2 delay from T1. A true AB occurs when performance diminishes at short lags, but accuracy increases to near control levels at longer lags (average T2 accuracy is insufficient to show whether AB has occurred). Researchers using the AB paradigm should choose one lag that is within the time window of AB impairment (200–500 ms) and one lag that is somewhat longer for comparison. The required performance measure task for Target 1 and Target 2 do not have to be the same although it is preferable that they be the same if comparison between T1 and T2 performance overall is desired (MacLean & Arnell, 2012). These basic requirements established through work with monolinguals are critical for bilingual research, as well. To summarize, the presentation rate ought to be approximately 100 ms per item, T1 and T2 should occur with a short lag and be compared to a long lag, and analysis must be based on the relative decrease in accuracy for T2 at short lags.

Experimenters use the AB paradigm for within-subjects designs comparing different word types as stimuli or between-subjects designs to examine individual differences (Willems, Wierda, van Viegen, & Martens, 2013) such as age (Georgiou-Karistianis et al., 2007; Lahar, Isaak, & McArthur, 2001), and video game experience (Green & Bavelier, 2003) just to name a few. Of special interest here is the case of bilinguals as compared to monolinguals or comparing bilinguals of varying proficiency.

In the context of studying effects of bilingualism, researchers may treat bilingualism as a dichotomous variable (having two separate categories) and compare participants who are roughly equal in all other ways except language experience. Alternatively, one can measure bilingual experience in some numerical way (e.g., proficiency score, total years of use, age at first use of L2, etc.) and correlate the bilingualism score with a measure of AB. If the first, dichotomous conception of bilingualism is preferred, the appropriate measure of whether the AB is different for the two groups is a two-factor design and analysis should look for an interaction between lag effects on T2 performance and group. A main effect of group would merely indicate a difference in overall T2 performance without necessarily changing the AB effect. If a more numerical rather than dichotomous measure of bilingualism is preferred, correlational methods are more appropriate for analysis. Correlations or regressions may be a more appropriate choice in many cases, as speakers can range from entirely monolingual (i.e., no ability to speak or comprehend any parts of any other language other than the native language) to multilingual with high fluency and proficiency in all languages spoken (Polinsky & Kagan, 2007; see MacLean and Arnell, 2012, for a discussion of statistically appropriate measures of the attentional blink).

In relation to bilingualism, Colzato et al. (2008) compared monolinguals with bilinguals in a traditional AB task and established that bilinguals were found to have a *larger* attentional blink than monolinguals (see discussion below). Based on Colzato et al.'s (2008) findings, Khare, Verma, Kar, Srinivasan, and Brysbaert (2012) sought to clarify the change in AB magnitude as a function of the time course of bilingual proficiency. Khare et al. (2012) used Hindi-English bilinguals of varying English proficiency living in India, as participants and found that higher proficiency bilinguals showed a larger relative AB at short lags when compared to lower proficiency bilinguals. These results support Colzato et al.'s (2008) previous conclusions that bilingual proficiency increases AB and extend them by showing that increased AB in bilinguals is not related to age of acquisition. AB effect research in bilinguals should continue to explore the limitations and basis for differences from monolinguals. We return to this issue in the "Executive Control" section.

Emotional AB

An important manipulation of stimuli characteristics in previous research in the AB paradigm has been to systematically vary the emotional valence (positive or negative) and arousal (how intense the word is or how physiologically aroused it makes the reader feel) of the target and distractor words. This line of work has produced interesting results referred to as the *emotional attentional blink* (EAB; see McHugo, Olatunji, & Zald, 2013 for a review). Highly arousing words, particularly taboo or sexual words, have shown results consistent with stimulus-driven attentional capture, as a variety of studies have found that participants attend to emotional words regardless of their relevance to the task's goal. In stimulus-driven attentional capture, features of the stimulus itself (such as valence or physical attributes like color or font) draw an individual's attention without higher-level intention to direct their focus to that feature of the stimulus. In this case, the valence and arousal of a distractor is not relevant to identifying targets, but the emotional distractor is attended to regardless. In some instances in previous work, the emotional words appeared as distractor words and caused target words to be missed (e.g., Arnell, Killman, & Fijavz, 2007) and in others, the emotional words appeared as targets (specifically, T2s defined by ink color) and reduced attentional blink effects (e.g., Anderson, 2005). In both types of results, the emotionally arousing stimuli captured attention despite task irrelevance.

Several researchers have chosen the EAB task to study bilingual emotion and language processing. Specifically, previous research has examined the effect of bilingual experience on EAB (Colbeck & Bowers, 2012). Colbeck and Bowers (2012) used Arnell et al.'s (2007) methodology using taboo/sexual words as distractor words prior to the presentation of an unrelated target word defined by ink color. Results with monolinguals had shown increased AB in conditions that featured the taboo/sexual distractors as compared to neutral control words (Arnell et al., 2007). Participants in Colbeck and Bowers' (2012) study were English monolinguals or

Chinese-English bilinguals of high-language proficiency. They found that while all of their bilingual participants had shown high proficiency, the EAB was reduced when compared to the English monolinguals. The authors concluded that length of bilingual experience, rather than proficiency, predicts the attention-capturing effects of emotionally arousing words, implying that a late-acquired L2 has less emotional strength than an L1. Clearly, additional research is required with bilingual participants of varying proficiencies, years of experience, and number of known languages. One possible confound in Colbeck and Bowers (2012) is based on the years of experience and age of acquisition among their bilingual participants. The bilingual participants not only learned English at a later age than the monolinguals but also have been speaking English for fewer years. Thus, while Colbeck and Bowers' conclusions were based on the length of *bilingual* experience, the participants also had less experience with the test language (English). Additional participants could include proficient bilinguals whose first language is English (equal English experience to the monolinguals and greater bilingual experience), or those who had been bilingual from a young age (equal English experience to the monolinguals and much greater bilingual experience).

Executive Control

While linguistic and attentional factors can be assessed using the RSVP paradigm, this task has also been used by researchers to evaluate executive control. Executive control is part of working memory and Baddeley's model of working memory is the most commonly used in cognitive psychology. It defines working memory as *the system necessary for holding and manipulating information while performing a wide range of tasks…* (Baddeley & Della Sala, 1996, p. 1398; see also Baddeley & Hitch, 1974). Executive control can also be conceived of as the ability to intentionally direct attention and keep relevant information to the task at hand readily available (Carlson & Wang, 2007). Thus, when we refer to information being *in working memory*, it refers to information that is readily available. Inhibitory control refers to *the ability to control potentially interfering thought processes and actions…such as suppressing a dominant response in accordance with rules* (Carlson & Wang, 2007, p. 489). Inhibitory control is important in situations in which distractors are highly salient and responses to them need to be inhibited in order to increase accuracy in responses to correct targets. Individuals select certain information from the environment to attend to (essentially at the expense of other less important or distracting information) and keep that chosen information active and ready to be used in current and future processing.

Working memory and inhibitory control are involved in efficient processing in the RSVP task. Participants must keep their attention narrowly focused on the briefly presented words, for if attention strays, they will not read words that were presented while their attention wandered. Additionally, working memory is required to keep all previously read words in the stream active in order to construct a coherent

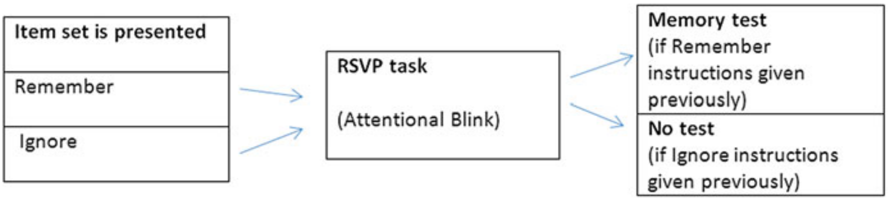


Fig. 4.11 Procedure for Akyürek and Hommel (2006)

sentence out of the individually presented stimuli. Due to working memory and inhibitory control being so closely involved, the RSVP paradigm has been used to study executive control in both monolinguals and bilinguals.

Several studies (Akyürek & Hommel, 2006; Arnell, Stokes, MacLean, & Gicante, 2010) have used RSVP methods and other executive control tasks to narrow in on the contribution of executive control abilities to RSVP performance. These studies, generally testing monolingual participants, have tested the same participants on several types of tasks and give a sense of the relative contributions of executive control and reading abilities on an individual's performance on RSVP tasks. Akyürek and Hommel (2006) used a dual task to test the role of working memory in RSVP processing and particularly the AB effect, previously examined in this chapter. Figure 4.11 shows the procedure for this study. Initially, all of their participants were presented with a set of items prior to the RSVP task, but varied in whether they had been instructed to remember the initial set of items or ignore them (the far left stage of Fig. 4.11). All participants then completed the same RSVP task measuring the AB (the center stage of Fig. 4.11). Finally, those instructed to "remember" the initial set of items completed a memory test after the RSVP stream. Thus, half of their participants completed the RSVP task with additional memory load and half did not.

Akyürek and Hommel (2006) found a main effect of memory demands on identification of targets in RSVP, with increasing interference (i.e., decreased accuracy in target identification performance) with a larger number of things to remember, but they did not see a change in AB magnitude (i.e., the relative performance deficit based on delay between the targets). While the mean accuracy was worse when participants had to remember more items, it did not change the relative difference in performance at various delays between the first and second target. This means it is overall more difficult to perform the target identification task when one is also holding items in working memory than without memory load. However, it is no more *relatively* difficult to identify the second target at a given lag from the first target under the working memory load. These results led the authors to conclude that working memory storage capacity is related to general RSVP performance, but it does not change the specific time-based attentional blink effect.

Arnell et al. (2010), however, tested participants on a similar RSVP task along with three different working memory tasks requiring an increasing degree of active manipulation of the items in working memory. The first was the *forward-digit span* in which a set of digits was presented and after a short period of time, the participant was asked to report the digits in the same order in which they were presented. For example, the participant may have been given the series, *2-7-8-5-3-3-9*, and to answer correctly, they would respond by repeating the original digits in the same order of presentation. The second was the *backward-digit span*, which is similar in procedure to the forward-digit span, however, the participants were asked to report the previously presented digits in the reverse temporal order from which they were presented (e.g., *9-3-3-5-8-7-2*). Finally, the *Operational Span* (O-Span) task was used, in which the participants were given one letter to remember after completing one math problem (Turner & Engle, 1989). For example, the participant may have been given the multiplication 8×3 and after they correctly responded *24*, they received the letter *g* for later recall. Next, they may have solved the equation $87 - 34$, and received the letter *p* for later recall. At the end of the series, the participant was asked to recall the series of letters they were asked to remember, in the order of presentation (e.g., *g*, *p*). O-Span is a strong measure of working memory because participants cannot engage in strategies such as rehearsal, as they are also solving math problems typically with minimal accuracy. Arnell et al. (2010) found a similar effect of memory load on RSVP performance, but only on the working memory test requiring the most executive control. O-Span was predictive of the attentional blink magnitude. That is, these results suggest that executive control is important to the attentional blink task beyond simple memory abilities. Thus, it is clear that the relationship between the RSVP paradigm and executive control is fairly complex.

This element of executive control in the RSVP paradigm is an important tool in examining questions of differences between monolinguals' and bilinguals' executive control capabilities. A body of research has revealed that bilinguals' lifelong practice with actively selecting one language and inhibiting another increases their general inhibitory control capacity (e.g., Prior & MacWhinney, 2010). Abutalebi and Green (2007) argued that the proposed bilingual advantage occurs because bilinguals manage both languages and select one to actively use at a time. This active selection of one language and inhibition of the other is analogous to an exercise for bilinguals' general executive control skills. Essentially, selection and inhibition of language transfers to some other related, but nonlinguistic tasks.

These differences have been examined using a variety of tasks including the *Simon task* (Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Luk, 2008; Bialystok & Martin, 2004; Simon & Ruddell, 1967), in which participants are presented with one of two possible target stimuli, and are instructed to press a particular key when each stimulus is presented (see Fig. 4.12). For example, if the target stimulus is a *z*, participants may be required to press the *z* key on the left side of the keyboard, but when a symbol such as the / (forward slash) appears, they must press the / key that is on the right side of the keyboard. The target stimuli can appear in different physical locations, either on the left side or right side of the screen. Inhibitory control is required in order to inhibit the distracting information that is given by a target appearing in the spatial location *opposite* of the correct response key, as the individual's inclination is to press the key that is in the *same* spatial location as the target. Individuals who are better able to exercise inhibitory control not only make fewer errors, but also provide the correct response quicker

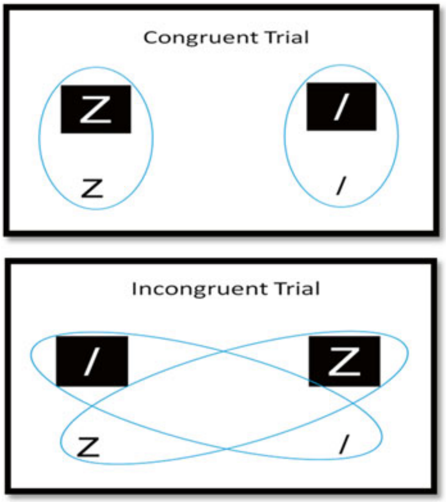


Fig. 4.12 Congruent and incongruent trials from a version of the Simon task (Adapted from Simon & Ruddell, 1967)

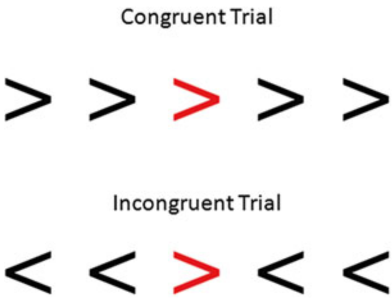


Fig. 4.13 Congruent and incongruent trials from the Flanker task (Adapted from Eriksen & Eriksen, 1974)

when the target appears in the opposite location of the correct key. Those results that use the Simon task to support bilingual advantages (Bialystok et al., 2004, 2008) have found quicker average reaction times for proficient bilinguals when compared to monolinguals or low-proficiency bilinguals.

The *flanker task* has also been used in several versions to investigate possible bilingual advantages (Costa, Hernández, & Sebastián-Gallés, 2008; Luk, De Sa, & Bialystok, 2011). Generally, in this task, participants should attend to a central target stimulus that has one or more “flanking” distractor stimuli (Eriksen & Eriksen, 1974). Inhibitory control is required to suppress the irrelevant information provided by the distractor stimuli and respond only with the response associated with the central target stimulus. For example, arrows are commonly used as both the target and distractor stimuli (see Fig. 4.13). The target stimulus is a centrally located red arrow. Typically, the correct response is to press the left arrow key when the red arrow points to the left, and to press the right arrow key when the red arrow points to the right. This arrow may be presented by itself, or with distractors to either side flanking it. When flanks are present, they are similar arrows; however, they are black in color and point in a congruent or incongruent direction to the target arrow. In this case, the participant should ignore the black distractor arrows and respond according to the direction of the red arrow just as when the red arrow is presented alone during other trials. Previous research shows that participants are slower to respond when the distractor stimuli point in a direction incongruent with the target stimulus, thus creating what is called the flanker effect. In several studies (Costa et al., 2008; Luk et al., 2011), bilinguals have demonstrated a smaller flanker effect than monolinguals, or in the case of Luk et al. (2011), earlier bilinguals showed a smaller deficit than late bilinguals.

However, these positive results are not without controversy. Such differences can be difficult to replicate in multiple executive control measures (Paap & Greenberg, 2013) or do not extend to all tasks (Bialystok & Martin, 2004; Hernández, Martin, Barceló, & Costa, 2013). For example, Bialystok and Martin (2004) found that bilingual children outperformed monolingual children on a *card sorting task* that asked children to sort cards on perceptual features (e.g., by color), then change which perceptual feature is used to divide the cards (e.g., shape), but this advantage did not extend to semantic features (e.g., natural categories like animals or fruits). Others find consistent effects, but posit different explanations than inhibitory control differences (e.g., Colzato et al., 2008). The RSVP task may be a helpful tool to investigate these questions because the task can be customized with different executive control and linguistic demands and contribute to a full understanding of bilingual executive control abilities.

One way in which RSVP can be manipulated to test bilinguals' executive control abilities is to use either low or high-constraint sentences, as defined above. Linck, Hoshino, and Kroll (2008) sought to compare Spanish-English bilinguals' performance on an RSVP task with their performance on the Simon task. Linck et al. asked participants to name target words that were either highly constrained by the sentence context, or less predictable from the sentence context of a sentence presented through RSVP. Linck et al. examined the latencies to name the target word in a sentence that was entirely in the participant's L2 in order to test whether sentence constraint had an impact on the interference from the bilinguals' more dominant L1. The target words were either cognates of English and Spanish, or noncognates. Linck et al. found an interaction between cognate status and sentence constraint such that cognates were only more quickly named than noncognate targets in low-constraint sentences (results were controlled for baseline language proficiency). This result suggests that high constraint in sentence context minimizes the interference or at least the effect of interference from the L1 to L2 processing. Interestingly, the effect of cognate status could be predicted by participants' scores on a reading span working memory test, but not by the magnitude of the Simon effect. These relationships suggest that reading working memory is more related to interactions with cognate status than executive control abilities.

Colzato et al. (2008) used several tasks including RSVP to test whether the proposed bilingual advantage in executive control is driven by active inhibitory systems that are in place to keep down unwanted responses, or a more general ability to focus on goal-related information instead. They employed an AB task along with inhibitory control tasks to assess monolinguals' and bilinguals' executive control abilities. They found no difference in the tasks tapping largely inhibitory abilities, and a *larger* AB in the RSVP task. This finding may be counterintuitive based on research showing improved executive control in bilinguals; however, this result is congruent when bilinguals' inhibitory control advantage is viewed in terms of selecting stimuli relevant to their current goal rather than avoiding irrelevant stimuli. In this experiment, the participants' goal was to accurately report target stimuli in the RSVP task. Instead of actively inhibiting task-irrelevant stimuli (i.e., the RSVP distractors), bilinguals were better at focusing on the goal task (i.e., target report). Bilinguals' focus on goal-relevant information is posited to explain why bilinguals performed "poorer" in terms of the AB effect: they had more effectively devoted more attention to the first target at the expense of the second target. Thus, the bilinguals may have been focusing on the goal of reporting the target more effectively than monolinguals, thereby decreasing AB. In other executive control tasks such as the flanker task, the same prioritization of goal-relevant stimuli is advantageous, allowing participants to report the direction of the target arrow without being distracted by the flanking arrows. Further research utilizing the RSVP paradigm will assist in clarifying the distinctions between executive control abilities of monolinguals and bilinguals of varying proficiency.

These few examples of studies examining bilinguals' executive control abilities using the RSVP task show its potential for future increased use as a different measure of executive control using language as stimuli, as opposed to the shape stimuli used in the flanker task. Researchers interested in using this paradigm in this way can customize the sentence type, target word type, number of target words, modality of response, language blocking, and a number of other elements in order to fine-tune the task specifications to allow for desired comparisons.

Limitations of RSVP

In addition to those limitations created by high working memory demands, single word presentation does not allow readers to make regressions to reread words previously presented. A final note that compares RSVP to eye-tracking addresses whether RSVP's lack of opportunity for regressions is problematic. In RSVP, presenting words one at a time prevents the reader from going back to reaffirm the identity or interpretation of any words. However, in natural reading, individuals often make regressions. Schotter et al. (2014) used a technique called the *trailing mask paradigm* to empirically test the importance of regressions to comprehension. In this task, participants try to read normally, but once they have moved their fixation past a word, that and all previous words are masked. For example, if the participant was reading the sentence *Bobby walked right past the block on the floor of his messy room*, they would read normally from left to right, and *Bobby* would be masked first. At the point at which they had read up to *block*, the sentence would look like this: *XXXXX XXXXXXX XXXXX XXXX XXX block on the floor of his messy room*.

This technique ensures that, although the participant reads left to right in a natural way, they cannot make regressions (or at least they would be uninformative as the words have been replaced with masks) to earlier parts of the sentence. Some of the sentences were garden-path sentences with ambiguous meaning, such as the example above. Others were not misleading, such as when *block* is replaced with *truck* in the above example to create the unambiguous sentences *Bobby walked right past the truck on the floor of his messy room*. In this experiment, Schotter et al. (2014) manipulated the ambiguity of the sentence (misleading garden-path sentences or not misleading sentences) and the format of sentence presentation (traditional or trailing mask), and measured the participants' accuracy for comprehension questions about the sentences.

Schotter et al. (2014) found that participants' comprehension test answers were more accurate for unambiguous sentences (a main effect of sentence ambiguity), and less accurate in the trailing mask condition (a main effect of presentation format). The interaction of sentence ambiguity and presentation format approached significance ($p = .06$). The authors argued that the statistical null hypothesis for this effect should be retained and that an interaction between sentence ambiguity and presentation format does not exist. The absence of an interaction would support their claims that the ability to make regressions aids comprehension for *all* types of sentences (not just ambiguous sentences). However, given that the interaction very nearly reached significance, further replications would increase the persuasiveness of their claims. Schotter et al. also showed that in the traditional presentation format condition, the sentences in which regressions were made had equal comprehension performance to those that did not. The results of this study imply that regressions may compensate for poor understanding, and thus, the inability to make regressions in RSVP could make it challenging to read material any more difficult than short sentences.

Summary and Conclusions

Within this chapter, an overview of the RSVP methodology has been provided. It has included definitions and explanations of concepts related to this methodology, notes for use in research, and previous work with bilinguals that has employed this paradigm. Specifically, RSVP is most commonly used in bilingual research to study lexical processing, attentional processes including the AB and RB, and executive control. Future research might examine more closely differences in lexical processing and attentional processes as they vary with proficiency and age of acquisition. The initial work that has begun using RSVP paradigms with bilingual participants has often been limited to comparing bilinguals and monolinguals (e.g., Colzato et al., 2008; Wong et al., 2011), though some studies have done very well at teasing these characteristics apart (Khare et al., 2012) or have focused exclusively on bilinguals (e.g., Altarriba et al., 1992, 1996). Another focus of future research should include manipulations of word types and contexts for studies using bilingual participants. While emotional words are one popular type of semantic manipulation, other characteristics such as animate/inanimate objects, word class (i.e., noun, verb, adjective, etc.), age of acquisition for specific words, or phonological or orthographic regularity would also be intriguing to investigate. This paradigm is a tightly-controlled, interesting way to investigate how bilinguals perceive, read, and attend to verbal stimuli, though it does have some limitations including the effects of memory on performance measures and the lack of regression capability. While to date it has only been used in a limited number of studies with bilinguals, investigators in the field should carefully consider whether RSVP may add to their options of laboratory techniques.

Keywords and Concepts

Attention, Attentional blink (AB), Backward-digit span tasks, Bilingual advantage, Bilingualism, Card sorting task, Cognate, Context effects, Item detection, Emotion, Executive control, Flanker task, Forward-digit span task, Goal-related focus, Homograph, Item identification, Inhibition, Lag, Lexical processing, Negative emotional words, Operational span task (O-Span), Perception, Rapid Serial Visual Presentation (RSVP), Reading, Repetition blindness (RB), Selection, Semantic features, Simon task, Sublexical elements, Target monitoring task, Translation, Valence, Visual attention, Word naming, Working memory

Review Questions

1. Brainstorm a list of advantages and disadvantages for using RSVP instead of a method like eye-tracking or self-paced reading. Include such factors as cost, generalizability, control, and naturalism among what you consider, along with any others that you think of yourself.
2. What other types of words may impact bilingual reading using RSVP? Specifically, given the minimal repetition blindness across languages for most words, might specific types of words induce an RB even across languages?
3. What can individual differences in executive control abilities such as those displayed by bilinguals on some tasks tell us about how these processes work in all populations?
4. Consider recent media attention on RSVP as an everyday reading method for "more efficient" reading. What are some possible advantages and disadvantages to using RSVP in this way? What impact could screen size and sentence difficulty have on reading performance?
5. What does bilinguals' good performance on high-constraint sentences vs. low-constraint sentences say about how they are accessing their two languages?

Suggested Student Research Projects

1. Use the RSVP applet (See Related Internet Sites below) to create an RSVP experiment. Enter words or letters into the applet. Have a friend or a classmate participate in the experiment, and then ask them to report back to you what they saw. Run the experiment once more, but this time, ask a different participant to locate just one or two specific words or letters. Repeat the experiment several times with different participants, using different items. How accurate are participants at perceiving the target items? Do you see differences based on whether participants are reporting the whole sequence or just one item? Ask them about their experience.
2. Translate some of the words from your lists from the first project into another language (or ask a proficient speaker to do this) for a new set of mixed-language lists. Find a bilingual participant who speaks both languages used in your stimuli and replicate the identification task described in the first project. How accurate are they at reporting targets written in the same or different language as the rest of the target words in the list? Are they better at reporting target words from their dominant language (ask your participant which language they feel more comfortable using)? If you can, repeat the experiment with another participant who speaks the same two languages, but has the opposite dominant language as your first participant.
3. Previous research has shown that taboo words (i.e., expletives) used as distractors create a larger attentional blink and that this effect is greater for the bilinguals' more proficient language. This may be considered an emotional attentional blink for negatively valenced, high arousal words. Design a study that tests whether the emotional attentional blink generalizes to *positively valenced* emotional stimuli. Consider what type of target stimuli and what type of distractors to use. How would you test if the same pattern of results is found for more and less proficient bilinguals?

Related Internet Sites

Attentional Blink: http://www.scholarpedia.org/article/Attentional_blink

Attentional Blink Experiment: http://psych.hanover.edu/javatest/cle/cognition/cognition/attentionalblink2_instructions.html

Reading and RSVP: http://jhenderson.org/vclab/Blog/Entries/2014/3/7_Am_I_Reading_This_Right.html

RSVP: http://en.wikipedia.org/wiki/Rapid_Serial_Visual_Presentation

RSVP applet: http://www.mscottreynolds.com/MyRSVP.html

RSVP script for E-Prime: http://step.psy.cmu.edu/scripts/Attention/Shapiro1994.html

Speed reading: http://www.huffingtonpost.com/2014/02/27/spritz-reading_n_4865756.html

Suggested Reading

Colzato, L. S., Spapé, M. M. A., Pannebakker, M. M., & Hommel, B. (2007). Working memory and the attentional blink: Blink size is predicted by individual differences in operation span. *Psychonomic Bulletin and Review*, *14*(6), 1051–1057.

Dux, P. E., & Marois, R. (2009). How humans search for targets through time: A review of data and theory from the attentional blink. *Attention*, *Perception*, *&* *Psychophysics*, *71*(8), 1683–1700.

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology*, *4*, 1–13.

McHugo, M., Olatunji, B. O., & Zald, D. H. (2013). The emotional attentional blink: What we know so far. *Frontiers in Human Neuroscience*, *7*, 1–9.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275.

Akyürek, E. G., & Hommel, B. (2006). Memory operations in rapid serial visual presentation. *European Journal of Cognitive Psychology, 18*(4), 520–536.

Altarriba, J., Carlo, M. S., & Kroll, J. F. (1992, May). *The multiple access of meaning in processing cross-language ambiguity*. Paper presented at the 64th annual meeting of the Midwestern Psychological Association, Chicago, IL.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition, 24*(4), 477–492.

Altarriba, J., & Soltano, E. (1996). Repetition blindness and bilingual memory: Token individuation for translation equivalents. *Memory and Cognition, 24*(6), 700–711.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134*(2), 258–281.

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing, 22*(6), 687–711.

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: Target misses follow attention capture by arousing distractors in RSVP. *Emotion, 7*(3), 465–477.

Arnell, K. M., Shapiro, K. L., & Sorenson, R. E. (1999). Reduced repetition blindness for one's own name. *Visual Cognition, 6*(6), 609–635.

Arnell, K. M., Stokes, K. A., MacLean, M. H., & Gicante, C. (2010). Executive control processes of working memory predict attentional blink magnitude over and above storage capacity. *Psychological Research, 74*(1), 1–11.

Baddeley, A. D., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351*(1346), 1397–1404.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–89). New York: Academic Press.

Bartolotti, J., & Marian, V. (2013). Language learning and control in monolinguals and bilinguals. *Cognitive Science, 36*(6), 1129–1147.

Basnight-Brown, D. M. (2014). Models of lexical access and bilingualism. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of bilingual memory* (pp. 85–107). New York: Springer.

Bavelier, D., & Potter, M. C. (1992). Visual and phonological codes in repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 134–147.

Bavelier, D., Prasada, S., & Segui, J. (1994). Repetition blindness between words: Nature of the orthographic and phonological representations involved. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1437–1455.

Beccue, B., & Vila, J. (2004). Assessing the impact of rapid serial visual presentation (RSVP): A reading technique. In V. Larios, F. F. Ramos, & H. Unger (Eds.), *Advanced distributed systems: Third international school and symposium (ISSADS 2004)* (Vol. 3061, pp. 42–53). Guadalajara, Mexico: Springer.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290–303.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859–873.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339.

Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics, 42*(2), 105–113.

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development, 22*(4), 489–510.

Colbeck, K. L., & Bowers, J. S. (2012). Blinded by taboo words in L1 but not L2. *Emotion, 12*(2), 217–222.

Coltheart, V., & Langdon, R. (2003). Repetition blindness for words yet repetition advantage for nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(2), 171–185.

Colzato, L. S., Bajo, M. T., van der Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 302–312.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59–86.

Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research, 69*(3), 191–200.

Dijkstra, A. F. J., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197.

Dux, P. E., & Marois, R. (2009). How humans search for targets through time: A review of data and theory from the attentional blink. *Attention, Perception, & Psychophysics, 71*(8), 1683–1700.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition, 11*(6), 565–574.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21*(2), 175–198.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology, 50A*(1), 119–148.

Georgiou-Karistianis, N., Tang, J., Vardy, Y., Sheppard, D., Evans, N., Wilson, M., et al. (2007). Progressive age-related changes in the attentional blink paradigm. *Aging, Neuropsychology, and Cognition*, *14*(3), 213–226.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*(5), 164–168.

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language, 69*(3), 257–276.

Huang, L., & Pashler, H. (2012). Boolean map approach to visual attention. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (2nd ed., pp. 29–46). New York: Guilford.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27*(2), 117–143.

Kanwisher, N. G. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 404–421.

Kanwisher, N. G., & Potter, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 30–47.

Khare, V., Verma, A., Kar, B., Srinivasan, N., & Brysbaert, M. (2012). Bilingualism and the increase attentional blink effect: Evidence that the differences between monolinguals and bilinguals generalizes to different levels of second language proficiency. *Psychological Research, 77*(6), 728–737.

Knickerbocker, H., & Altarriba, J. (2013). Differential repetition blindness with emotion and emotion-laden word types. *Visual Cognition, 21*(5), 599–627.

Lahar, C. J., Isaak, M. I., & McArthur, A. D. (2001). Age differences in the magnitude of the attentional blink. *Aging, Neuropsychology, and Cognition, 8*(2), 149–159.

Lai, M. L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T. C., Lee, S. W. Y., et al. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*(88), 90–115.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 381–390.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon, 3*(3), 349–374.

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control. *Bilingualism: Language and Cognition, 14*(4), 588–595.

MacKay, D. G., Abrams, L., Pedroza, M. J., & Miller, M. D. (1996). Cross-language facilitation, semantic blindness, and the relation between language and memory: A reply to Altarriba and Soltano. *Memory and Cognition, 24*(6), 712–718.

MacKay, D. G., Hadley, C. B., & Schwartz, J. H. (2005). Relations between emotion, illusory word perception, and orthographic repetition blindness: Tests of binding theory. *The Quarterly Journal of Experimental Psychology, 58A*(8), 1514–1533.

MacKay, D., James, L. E., & Abrams, L. (2002). Cross-language facilitation, repetition blindness, and the relation between language and memory: Replications of Altarriba and Soltano (1996) and support for a new theory. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 89–109). Amsterdam: Elsevier Science.

MacKay, D. G., & Miller, M. D. (1994). Semantic blindness: Repeated concepts are difficult to encode and recall under time pressure. *Psychological Science, 5*(1), 52–55.

MacLean, M. H., & Arnell, K. M. (2012). A conceptual and methodological framework for measuring and modulating the attentional blink. *Attention, Perception, & Psychophysics, 74*(6), 1080–1097.

McHugo, M., Olatunji, B. O., & Zald, D. H. (2013). The emotional attentional blink: What we know so far. *Frontiers in Human Neuroscience, 7*, 1–9.

Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 69–89). Hillsdale, NJ: Bradford/MIT Press.

Olivers, C. N. L., van der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological Research, 71*(2), 126–139.

Öquist, G., & Lundin, K. (2007). Eye movement study of reading text on a mobile phone using paging, scrolling, leading, and RSVP. In M. Ojala & T. Ylianttila (Eds.), *Proceedings from the 6th International Conference on Mobile and Ubiquitous Multimedia (MUM 2007)* (pp. 176–183), Oulu, Finland.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258.

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the ‘wild’ and in the classroom. *Language and Linguistics Compass, 1*(5), 368–395.

Potter, M. C. (1984). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. E. Kieras & M. A. (Eds.), *New methods in reading comprehension research* (pp. 91–118). Hillsdale, NJ: Bradford: MIT Press

Potter, M. C. (1999). Understanding sentences and scenes. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief stimuli* (pp. 13–46). London, England: Bradford: MIT Press.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition, 13*(2), 253–262.

Rahman, T., & Muter, P. (1999). Designing an interface to optimize reading with small display windows. *Human Factors, 41*(1), 106–117.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849–860.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/noncognate distinction. *European Journal of Cognitive Psychology, 4*(4), 293–310.

Schotter, E. R., Tran, R., & Rayner, K. (2014). Don’t believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science, 25*(6), 1–9.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*(2), 197–212.

Silvert, L., Naveteur, J., Honoré, J., Sequiera, H., & Boucart, M. (2004). Emotional stimuli in rapid serial visual presentation. *Visual Cognition, 11*(4), 433–460.

Simon, J. R., & Ruddell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*(3), 300–304.

Smigasiewicz, K., Shalgi, S., Hsieh, S., Möller, F., Jaffe, S., Chang, C. C., et al. (2010). Left visual-field advantage in the dual-stream RSVP task and reading-direction: A study in three nations. *Neuropsychologia, 48*(10), 2852–2860.

Soto-Faraco, S., & Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *The Quarterly Journal of Experimental Psychology, 55A*(1), 23–40.

Stewart, M. T., & Heredia, R. R. (2002). Comprehending spoken metaphoric reference: A real-time analysis. *Experimental Psychology, 49*(1), 34–44.

Taylor, I., & Taylor, M. M. (1983). *The psychology of reading*. New York: Academic.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127–154.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*(3), 431–451.

Velan, H., & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review, 14*(5), 913–918.

Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition, 118*(2), 141–156.

Whittlesea, B. W. A., & Masson, M. E. J. (2005). Repetition blindness in rapid lists: Activation and inhibition versus construction and attribution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 54–67.

Willems, C., Wierda, S. M., van Viegen, E., & Martens, S. (2013). Individual differences in the attentional blink: The temporal profile of blinkers and non-blinkers. *PLoS One, 8*(6), 1–16.
Wong, A. C. N., Qu, Z., McGugin, R. W., & Gauthier, I. (2011). Interference in character processing reflects common perceptual expertise across writing systems. *Journal of Vision, 11*(1), 1–13.
Yeh, S. L., & Li, J. L. (2004). Sublexical processing in visual recognition of Chinese characters: Evidence from repetition blindness for subcharacter components. *Brain and Language, 88*(1), 47–53.
Yu, D., Park, H., Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text. *Journal of Vision, 10*(2), 1–17.

Chapter 5
Bilingual Reading: The Visual Moving Window

Roberto R. Heredia, Belem G. López, Omar García, Wualú A. Altamira, and Patricia G. González

Abstract This chapter critically reviews the self-paced *visual moving window* (VMW) technique and its variants as used in the bilingual reading literature. In the first section, we provide a general overview of some of the variables known to compromise the validity of an experimental task, or an experiment in general. In the second section, we review bilingual reading experiments investigating the effects of code-switched or mixed-language sentences (e.g., *Andrea dropped the LETTER/CARTA in the mailbox*) as a function of context, word frequency, grammatical gender (masculine, feminine), and cognates (words with overlapping orthographical and meaning across languages) vs. homographs (words with overlapping orthographical representations but different meaning across languages). Finally, task strengths and weaknesses are discussed. We conclude by suggesting directions for future research and how this task can be used in conjunction with other tasks to explore bilingual sentence processing.

Introduction

In this chapter, we provide a critical overview of the *visual moving window's* (VMW; Just, Carpenter, & Woolley, 1982; see also Chap. 3) functionality and its variants, as used in bilingual reading research. In the VMW task, participants

R.R. Heredia, Ph.D. (✉) • W.A. Altamira
Department of Psychology and Communication, Texas A&M International University, 5201 University Boulevard, Laredo, TX 78041-1900, USA
e-mail: rheredia@tamiu.edu; wualualtamira@dusty.tamiu.edu

B.G. López
Department of Psychology, Texas A&M University, College Station, USA
e-mail: bglopez09@neo.tamu.edu

O. García • P.G. González, Ph.D.
Cognitive Neuroscience Laboratory, Texas A&M International University, 5201 University Boulevard, Laredo, TX 78041-1900, USA
e-mail: omaromega@hotmail.com; p.g.gonzalez22@gmail.com

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_5

typically read sentences word-by-word or phrase-by-phrase on a computer screen, from left-to-right. Before continuing our discussion about the VMW and bilingual reading, we would like to address several methodological issues requiring special attention. One important aspect that we would like to emphasize is that although the VMW task and its variants are typically considered *online tasks* in the sense that they measure tacit knowledge (i.e., *implicit knowledge*) that is beyond the participant's awareness, no psycholinguistic task is perfect or immune to methodological negligence. The validity of these tasks, as with any other psycholinguistic task, can be readily compromised by such issues as *participant effects* (e.g., demand characteristics), *experimenter effects* (e.g., experimenter expectancies), or choice of experimental stimuli (e.g., highly predictable sentences likely to engage strategic processing or guessing), to name just a few. Another important methodological issue that we would like to further stress is that reaction time (RT) or reading time as a dependent variable can be a necessary but not a sufficient enough condition to classify a psycholinguistic task as online (e.g., García, Cieślicka, & Heredia, 2015; Swinney, Love, Walenski, & Smith, 2007). As an example, consider a psycholinguistic task in which bilingual participants rate sentences of the type (1a) *The <u>PAN</u> is fresh* in relation to whether the sentence is likely to be semantically meaningful on a 1–7 scale (1 = non semantically plausible, 7 = semantically plausible). Depending on the experimental instructions (bilingual vs. monolingual) and other language factors (proficiency/dominance), participants may or may not provide high semantic plausibility ratings. What is interesting about this sentence is that if interpreted in English, it may be somewhat anomalous because cooking utensils such as *PANS* are usually associated with heat, but it would also be possible for a *PAN* to be fresh in the sense of being cold. However, from a bilingual language perspective, where *PAN* could also be *BREAD* in Spanish, the sentence would be rated as highly semantically plausible. In either case, the decision process to assign ratings to the sentence would certainly require multiple mental strategies such as problem solving, as participants entertain the different possibilities. These largely decision/strategic processes would be most clearly revealed in relatively longer reading times, if the rating task is timed (as in measuring the time taken to rate each sentence). The same issue would apply, for example, if a global reading measurement were used, where reading times for the entire sentence are used, as opposed to utilizing a localized task capable and sensitive enough to pinpoint the cognitive process being unfolded as reading takes place.

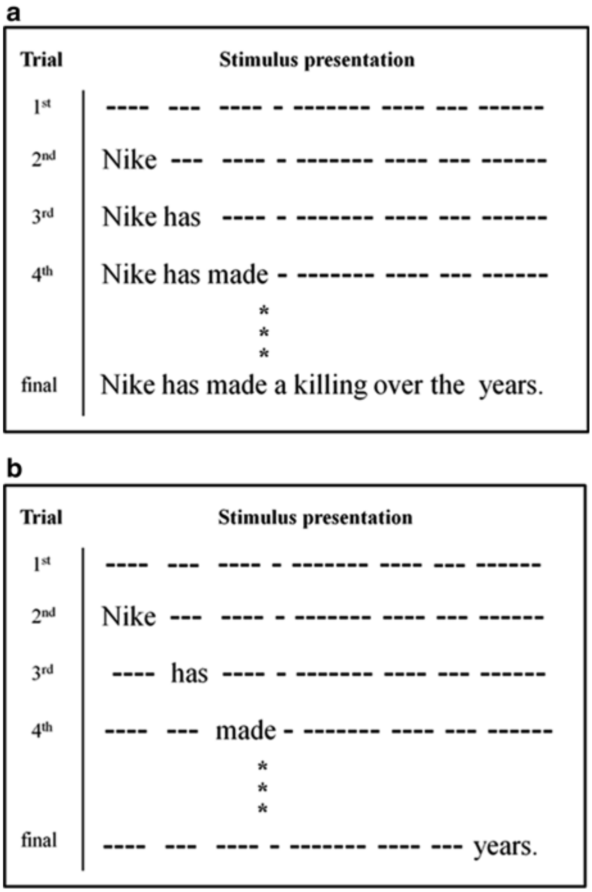
Further, consider the results reported by Cook (1990, 1997) in which Japanese English bilinguals took approximately 7600 ms to infer that the reflexive pronoun or reflexive anaphoric reference (e.g., *himself*) was directly connected to an explicitly mentioned earlier noun (*doctor*) as in sentence, (2a) *The skier said the doctor helps HIMSELF*. In light of the findings that word recognition is a fast process taking approximately less than 400 ms or less than 200 ms when estimates are taken from reading (Rayner, Pollatsek, Ashby, & Clifton, 2012, p. 53), relatively long RTs or reading times produced by a reading task should be taken with caution since they may be indicative of other cognitive processes (e.g., conscious or strategic processes) being measured. In turn, it may very well be the case that these types of reading tasks are simply not sensitive to the time course of language processing (García et al., 2015; Heredia & Muñoz, 2015; Hillert, 2002; Swinney & Osterhout, 1990; Swinney et al., 2007; see also, Bolger & Zapata, 2011). Indeed, mental and language processing efficiency is exemplified by Lieutenant Commander Data as he reflects about his newly learned human ability to feel emotions and the number of seconds he spent "thinking" of joining the enemy to retain these human qualities: …*And for a time*, *I was tempted. How long a time*? Asks the captain. *Zero point sixty-eight seconds* [680 ms], *sir. For an android*, *that is nearly an eternity*, responds Lieutenant Commander Data (Herman & Frakes, 1996).

Task Description

The VMW is a self-paced reading task in which participants are in direct control of the pace or speed with which they move from word-to-word or phrase-to-phrase as they read a given sentence. Stimulus words or phrases are presented on a computer screen one at a time or in segments (see Fig. 5.1), and from left-to-right. The participants' task is to simply read each word or phrase and press a key or button as fast as possible after reading the presented stimulus. To see the next stimulus word or phrase, participants must continuously press a button and repeat the behavior consistently throughout the experiment. For example, depending on the theoretical question at hand, sentence (3a) *Nike has made a killing over the years* could be presented word-by-word or phrase-by-phrase to assess reading time differences between the reading of the idiom's compositionality (i.e., word-by-word) or holistically as a chunk or phrase (see for example, Cieślicka, 2013, 2015; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Titone, Columbus, Whitford, Mercier, & Libben, 2015). In addition to its literal interpretation, the idiomatic expression, *made a killing*, in sentence (3a) could also be interpreted in terms of its intended meaning that "Nike has made a lot of money over the years." As an additional point, note that an otherwise *incidental task* (a term we borrow from the memory domain to describe a task in which participants are unaware of the nature of the experiment) can be confounded—and in turn, elicit strategic processing—by simply altering the font size or color of the critical target.

The VMW is a chronometric paradigm since it includes reading time course estimations as a function of its button-pressing behavioral task (Carpenter, 1984, p. 9). Reading times are recorded and interpreted as the time taken to read or process a given stimulus. This task measures the time between a word onset presentation and the subsequent button pressing. It is assumed that the button pressing indicates that the word was coded, analyzed, and understood by the reader (Mitchell, 1984).

Fig. 5.1 (**a**, **b**) A cumulative presentation: Word segments remain on screen after each button press (**a**) and a noncumulative presentation: Characters are replaced by *dashes* after each button press (**b**)



Typically, an increase in reading times or longer reading times reflects difficulty in comprehension, relative to a control target (Katz & Ferretti, 2001).

As argued by Aaronson and Ferres (1984), the VMW is most accurate for single word or word-by-word presentation because the obtained reading times allow for fine grained data analyses, as opposed to global or general measures of reading. As it is the case with most self-paced reading tasks, the VMW is sensitive to individual differences since participants are allowed to proceed at their own pace; the participant's reading times reflect momentary processing of a word as the participant is exposed to it (e.g., showing differences for phrase boundaries and key content words); and its unique setup encourages and shows regular reading strategies and habits (Aaronson & Ferres, 1984).

VMW Variants

Variations of this task include the traditional VMW, discussed in the previous section, and the stationary moving window (SMW). The traditional VMW can be further subcategorized into cumulative (Fig. 5.1a) and noncumulative (Fig. 5.1b) characteristics, which also differentiate the various versions of this experimental technique. In the cumulative VMW (e.g., Katz & Ferretti, 2001), words appear on the computer screen without disappearing, even after pressing the button to advance to the next word, unlike the noncumulative presentation, in which nonspace characters (i.e., individual letters of a word) are replaced by dashes (e.g., ---- - ------- for *made a killing*; note that every dash represents a character), and every button press reveals the first word of the sentence while replacing the previous one with dashes. In the SMW or center noncumulative presentation (see Fig. 5.2), stimuli appear in the middle of the screen in the same location for each word (Just et al., 1982; Mitchell, 2004). The SMW is similar to the *rapid serial visual presentation* (RSVP) task; however, in RSVP, word presentation is set at a predetermined duration rate. To assure active reading and sentence comprehension, participants answer simple true/false questions at random throughout the study. Correct responses or *accuracy rates* are recorded and analyzed separately to determine the speed-accuracy tradeoff and

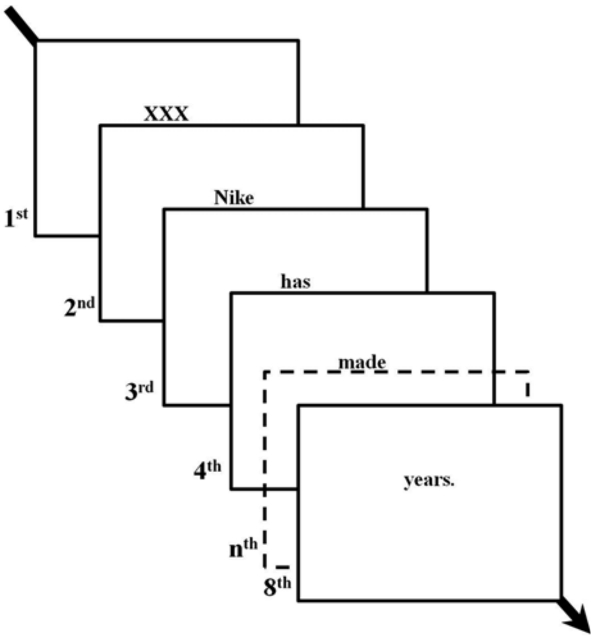


Fig. 5.2 Progression of a center noncumulative or SMW stimulus word per window (segment) as participants press a button continuously

assess possible relationships between accuracy or error rates and reading times. Another variant of the VMW that has been underutilized in the bilingual language processing is the *auditory moving window* (AMW; Ferreira & Henderson, 1990; Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). The AMW task is similar to the VMW; the exception being that words or phrases are presented auditorily (see discussion below for a bilingual version of this task).

Bilingual Reading and the VMW

In one of the first bilingual reading studies, Altarriba, Kroll, Sholl, and Rayner (1996) looked at the influence of contextual information and word frequency on the recognition of mixed or code-switched language sentences. In two experiments, Spanish-English bilinguals participated in an eye-tracking experiment in which their eye movements were recorded as they read (see Chaps. 3, 4 and 8), or a *rapid serial visual presentation task* (RSVP; see Chap. 4), in which sentences were presented word-by-word at a rate of 100 ms per word, in the center of a computer screen. In the RSVP, the participants' task was to name a critical target in uppercase letters that appeared in the middle of the sentence and in the center of the computer screen. Sentences were either all in English (see Table 5.1, sentence A. *He needed to put a stamp on the LETTER …*) or mixed language (sentence A. *He needed to put a stamp on the CARTA …*). At issue was whether the preceding context (high vs. low constraint) and word frequency (where high-frequency words are read faster than low-frequency words) had an effect on bilingual reading. In the high-constraint contextual condition, the preceding context was biased towards the critical word; for the low-constraint contextual condition, no biasing information followed the critical target. Additionally, targets were either high- or low-frequency words.

Findings were interpreted in terms of Schwanenflugel and LaCount's (1988) *Feature Restriction Model*. According to this model, sentence constraint determines

Table 5.1 Sample sentences used in Altarriba et al. (1996), and Heredia et al. (2003)

A. High-constraint context × high-frequency target
He needed to put a stamp on the LETTER/CARTA before he mailed it
B. Low-constraint context × high-frequency target
Andrea dropped the LETTER/CARTA in the mailbox at the corner
C. High-constraint context × low-frequency target
Pete took the warm cake out of the OVEN/HORNO and put it on the table
D. Low-constraint context × low-frequency target
We went to the store to buy a new OVEN/HORNO for our kitchen

the number of semantic featural descriptions generated during the reading process; a high-constraint context generates a higher number of restrictions, thereby limiting the activation of potential word candidates to only those that semantically match the expected completion of a sentence. Thus, for sentence (A) in Table 5.1, the preceding context would trigger a relatively large number of features associated with *stamps* (e.g., *letter*, *post office*, *mailman*, *ink*, *paper*), thus limiting the number of possible candidates to only a few. Candidates with semantic features matching those generated by the sentence would be responded to faster (Altarriba et al., p. 486). Low-constraint context, on the other hand, would generate fewer feature restrictions, thereby increasing the likelihood of generating multiple candidates during the reading process. Thus, for sentence (B) in Table 5.1, almost anything (e.g., *ball*, *glasses*, *box*, *bottles*) would make the sentence semantically felicitous. As predicted by the model, in both tasks and for the mixed-language conditions only, low-frequency target words benefited from the high-constraint contextual condition; bilinguals were faster to name low-frequency words under high than under low-constraint conditions. The reverse was true for high-frequency words; bilinguals were faster in naming target words under low-constraint contextual conditions. High-constraint contextual conditions produced higher reading times. Although Altarriba et al. report a set of impressive findings in which both the eye-tracking and the RSVP tasks behave very similarly, it is not clear why the monolingual sentences did not show any contextual effects, as did the reading of the mixed-language sentences (i.e., English sentences with embedded Spanish word targets).

Heredia, Martínez, Clark, and Moreno (2003, Experiment 1) extended Altarriba et al.'s (1996) study to include a more homogenous and representative sample of active bilinguals in a "quasi-pure" bilingual community where Spanish and English are typically mixed during the communicative process (e.g., Heredia & Altarriba, 2001). To further generalize the results to other experimental tasks, Heredia et al. utilized the self-paced SMW task, described in Fig. 5.2. As in the original study, high-proficiency Spanish-English bilinguals read sentences word-by-word presented in the center of the screen. Experimental sentences were taken from Altarriba et al. (1996). The overall analysis conformed to a 2(Critical Word Target: Spanish vs. English) × 2(Word Frequency: Low vs. High) × 2(Context: Low vs. High Constraint) design. The results of the three-way interaction are summarized in Table 5.2.

Similar to Altarriba et al. (1996), word frequency and context did not have an effect on reading English sentences. However, in the mixed-language condition, high-frequency Spanish words were read faster under low-constraint than under high-constraint contextual conditions. The opposite was true; low-frequency Spanish targets were read faster under high-constraint than under low-constraint context. Thus Heredia et al.'s results mirror Altarriba and colleagues' findings using a different bilingual population and a different experimental task.

The three-way interaction is further qualified by the three significant interactions described in Fig. 5.3a–c. Although context did not have an effect on reading English targets, Spanish targets were read faster under high-constraint contextual conditions (Fig. 5.3a). Moreover, Spanish word targets exhibited the word frequency effect (Fig. 5.3b), where high-frequency words are read faster

Table 5.2 Mean (*M*) reading times (in ms) and standard errors (SE) to target language (Spanish vs. English) as a function of word frequency (low vs. high) and contextual conditions (high constraint vs. low constraint)

Language	Word frequency and contextual condition	*M*	SE
Spanish	High-frequency high-constraint context	488	21
	High-frequency low-constraint context	451	20
	Low-frequency high-constraint context	484	27
	Low-frequency low-constraint context	575	21
English	High-frequency high-constraint context	384	10
	High-frequency low-constraint context	385	10
	Low-frequency high-constraint context	402	13
	Low-frequency low-constraint context	402	11

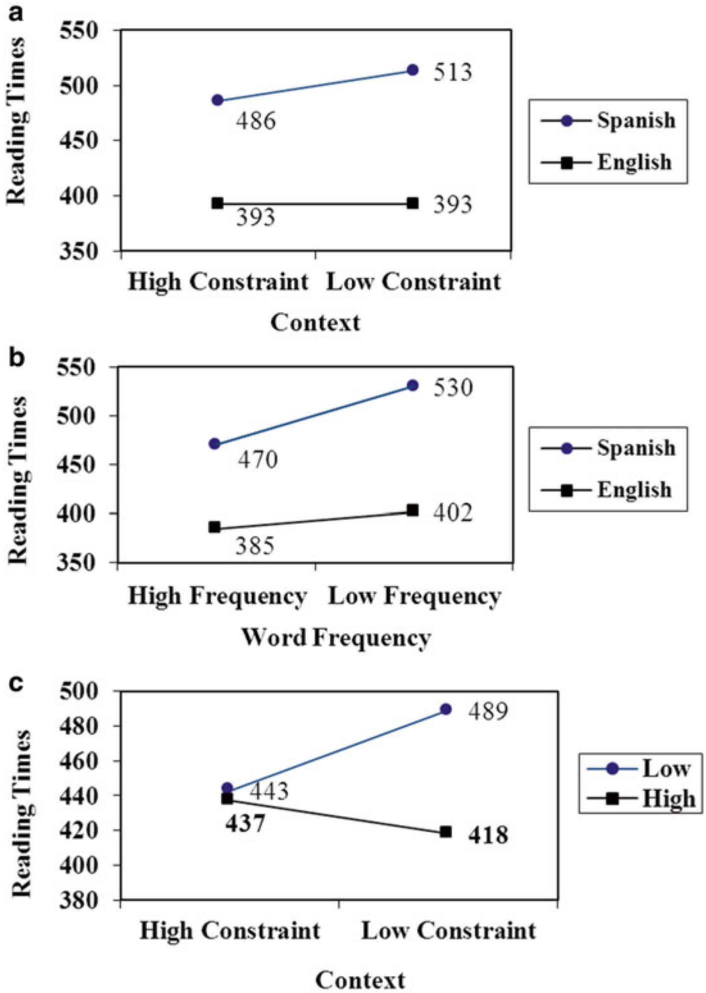


Fig. 5.3 (**a**–**c**) Interactions in Heredia et al.'s (2003) Experiment 1: Context vs. target language (**a**); word frequency vs. target language (**b**); and context vs. word frequency (**c**)

than low-frequency words (see for example, Rayner et al., 2012). English word targets followed the same pattern; however, the 17 ms reading facilitation for high-frequency words did not reach significance. In relation to word frequency and contextual constraint (Fig. 5.3c), high-frequency words were read faster under low-constraint contextual conditions. However, high-constraint contexts did not have an effect on word frequency.

In a second experiment, Heredia et al. (2003) used the SMW to further explore the effects of Spanish grammatical gender (*el* vs. *la*) in the comprehension of monolingual and mixed-language sentences (see Chap. 7; see also Dussias, 1997, 2001). At issue was whether the addition of grammatical marking to an otherwise English sentence would significantly slow down the reading process relative to a monolingual and a code-switched sentence without Spanish grammatical marking. As some research has reported, Spanish-dominant bilinguals assign English nouns the gender of the Spanish translation equivalent (e.g., Liceras, Fernández Fuertes, Perales, Pérez-Tattam, & Spradlin, 2008; Valenzuela et al., 2012). Because of the Spanish gender classification system in which nouns that end in *–o* (e.g., *perro* as in "dog") are masculine and nouns that end in *–a* (e.g., *casa* as in "home") are feminine, bilinguals would conceive of the English translation equivalents as masculine (*el dog*) and feminine (*la home*), respectively. Although it is the case that the English language contains a system noun classification to categorize certain nouns as masculine (e.g., *actor*, *buck*) or feminine (e.g., *actress*, *doe*), abstract (e.g., *idea*) and inanimate objects (e.g., *table*), are less likely to be part of this classification. However, bilinguals whose L1 is Spanish, would be more likely to assign grammatical gender to both animate (*el actor*, *el buck*) and inanimate objects alike (e.g., *la idea*, *la table*).

As shown in Table 5.3, sentences in Heredia et al.'s study were classified in terms of whether the target word in English matched (Congruent-match) or mismatched (Incongruent-match) the grammatical gender of the Spanish Equivalent. The monolingual English sentence was included to serve as the baseline. For the mixed-language condition, the Spanish correct (Congruent-match) and incorrect conditions (Incongruent-mismatch) were included, as well as a sentence containing a code-switched Spanish target without its Spanish-English marking (labeled Bilingual). Participants in this experiment were Spanish-English bilinguals from the same bilingual population as those from Experiment 1.

Table 5.3 Sample sentences used in Heredia et al.'s (2003) Experiment 2

English target + Spanish grammatical gender
We took a walk in LA CITY before we drove back (**Congruent-match**)
We took a walk in the CITY before we drove back (**Monolingual**)
We took a walk in EL CITY before we drove back (**Incongruent-mismatch**)
Spanish targets + Spanish grammatical gender
We took a walk in LA CIUDAD before we drove back (**Congruent-match**)
We took a walk in the CIUDAD before we drove back (**Bilingual**)
We took a walk in EL CIUDAD before we drove back (**Incongruent-mismatch**)

Table 5.4 Mean (*M*) reading times (in ms) and standard errors (SE) to target language (English vs. Spanish) as a function of word frequency (low vs. high) and Spanish definite article

Language	Word frequency + Spanish article (La vs. El)	*M*	SE
English	High frequency + congruent-match	483	25
	High frequency + monolingual	408	12
	High frequency + incongruent-mismatch	526	25
	Low frequency + congruent-match	479	18
	Low frequency + monolingual	439	18
	Low frequency + incongruent-mismatch	469	19
Spanish	High frequency + congruent-match	582	24
	High frequency + bilingual	518	26
	High frequency + incongruent-match	594	30
	Low frequency + congruent-match	600	31
	Low frequency + bilingual	543	29
	Low frequency + incongruent-mismatch	636	40

The results of a three-way interaction are described in Table 5.4. For the English high-frequency word condition, the Monolingual sentence was read faster followed by the Congruent-match (e.g., *LA CITY*). The Incongruent-mismatch target word took longer to read than the Monolingual target; however, its reading time did not differ from the Congruent-match. Although the low-frequency targets showed a similar pattern as the high frequency words, no reliable reading differences were exhibited among the three grammatical gender manipulations. Of greater interest, however, was the finding that the Spanish high- and low-frequency word conditions followed a pattern similar to the one exhibited by the English target conditions. In both high- and low-frequency word conditions, Spanish targets without grammatical gender markings were responded to faster than when the markings were present (see Chap. 7 for further details).

How do bilinguals classify and integrate grammatical gender information during the reading/language comprehension process? Although some studies (see for example, Dussias, 1997, 2001) are beginning to address these issues, future research may need to include other gender classification instances in which, for example, the ending of a noun is not associated with a particular gender, (e.g., *el barril*$_{\text{the-barrel-masc}}$), the noun accepts both determiners (e.g., *el azúcar*$_{\text{the-sugar-masc}}$ vs. *la azúcar*$_{\text{the-sugar-fem}}$), or exceptions to the *–o*/*–a* grammatical rule in which masculine (e.g., *la mano*$_{\text{the-hand-fem}}$), and feminine nouns (*el agua*$_{\text{the-water-fem}}$) take the opposite determiner. More interesting, however, would be the case for compound words (e.g., *peanutbutter*) that follow a left-headed compound structure in the bilingual's L1 (Spanish), and a right-headed structure in the L2 (English). Thus, during code-switching, Spanish-dominant bilinguals might be more likely to assign the masculine to the compound word following the Spanish structure, where *peanut* (*el cacahuate* $_{\text{the-peanut-masc}}$) is the head of the compound (e.g., *el peanutbutter*); whereas English-dominant bilinguals, might consider *butter* (*la mantequilla* $_{\text{the-butter-fem}}$) the head of the compound and assign the feminine article (e.g., *la peanutbutter*).

More recently, Heredia, Cieślicka, and García (2010a, 2010b) utilized the SMW task to investigate the processing of interlingual homographs and cognates across languages. Briefly, interlingual homographs (henceforth, homographs) are words between languages with similar or identical orthographical representations, but whose meaning is nonequivalent (e.g., *taller*="of great height" in English vs. *taller*="workshop" or "car shop" in Spanish); cognates, on the other hand, share overlapping orthographic (and phonological in some instances) and semantic representations across languages (e.g., *hospital* is spelled and means the same in Spanish and English). At issue is whether bilingual lexical access is *selective*, in the sense that during the comprehension of homographs, for example, retrieval is limited to one language candidate (or meaning) of the language that is contextually appropriate. Thus, for the highly biasing context in the homographic condition (see Table 5.5), the critical target *TALLER* would only activate the English meaning and not the "workshop" or "car shop" Spanish associated meaning. The opposing view, the *Nonselective Hypothesis*, poses a cognitive structure in which bilingual lexical access is exhaustive and, during the disambiguating process, multiple meanings of the homograph are activated simultaneously (e.g., Degani & Tokowicz, 2010; Libben & Titone, 2009; Schwartz & Kroll, 2006; Titone, Libben, Mercier, Whitford, & Pivneva, 2011; see also Chaps. 2, 4 and 8).

Spanish-English and English-Spanish bilinguals who were highly proficient in both Spanish and English participated in the self-paced SMW reading experiment. Participants were presented with English sentences, as described in Table 5.5, in which the preceding context biased the meaning of the critical homograph toward the English meaning. For the nonbiased contextual condition, the preceding context was neutral toward the meaning of the homograph. The cognate sentences followed the same format, except that for the contextually biased condition, the preceding context was biased towards both the Spanish and English meanings because cognates are similar in meaning across languages. The results of two-way interactions are summarized in Fig. 5.4.

Figure 5.4a shows that prior biasing context did not have an effect on the reading of the homograph. The 14 ms reading facilitation difference for the homograph, relative to the matched control, was not significantly different. However, for the nonbiasing contextual condition, the experimental matched control was actually faster than the homograph. That is, bilinguals were 23 ms slower in reading the homograph, relative to the matched control. This effect of *homograph inhibition* or *homograph interference* has been interpreted as lexical competition in which two potential meanings of an ambiguous word compete for activation (Libben & Titone, 2009; Titone et al., 2011). Thus, Heredia et al.'s (2010a, 2010b) results are suggestive of the *homograph interference effect*, and supportive of the claim that bilingual lexical access is nonselective, but only for the unbiased contextual condition. Figure 5.4b shows, again, that biasing context did not have an effect on the recognition of cognates. However, the nonbiasing contextual condition revealed the general *cognate facilitation effect*, whereby cognates are responded to faster than noncognates. The cognate effect has been generally taken as supporting bilingual nonselectivity (Libben & Titone, 2009; Schwartz & Kroll, 2006; see also Chaps. 2, 4, and 8).

Table 5.5 Sample sentences used in Heredia et al.'s (2010a, 2010b) SMW experiment

	Biasing context (1–5 rating scale)
Homograph/control	*The fact that Ernie has longer legs does not make him TALLER/BUSIER than the rest of the basketball team* (**biasing context = 3.9**)
Cognate/control	*The office secretaries believed the new DIRECTOR/MARRIAGE was the worst and awful alternative* (**biasing context = 4.0**)
	Nonbiasing context
Homograph/control	*Sometimes being TALLER/BUSIER is not necessarily better* (**biasing context = 2.1**)
Cognate/control	*All of them thought the DIRECTOR/MARRIAGE was the worst and awful alternative* (**biasing context = 1.7**)

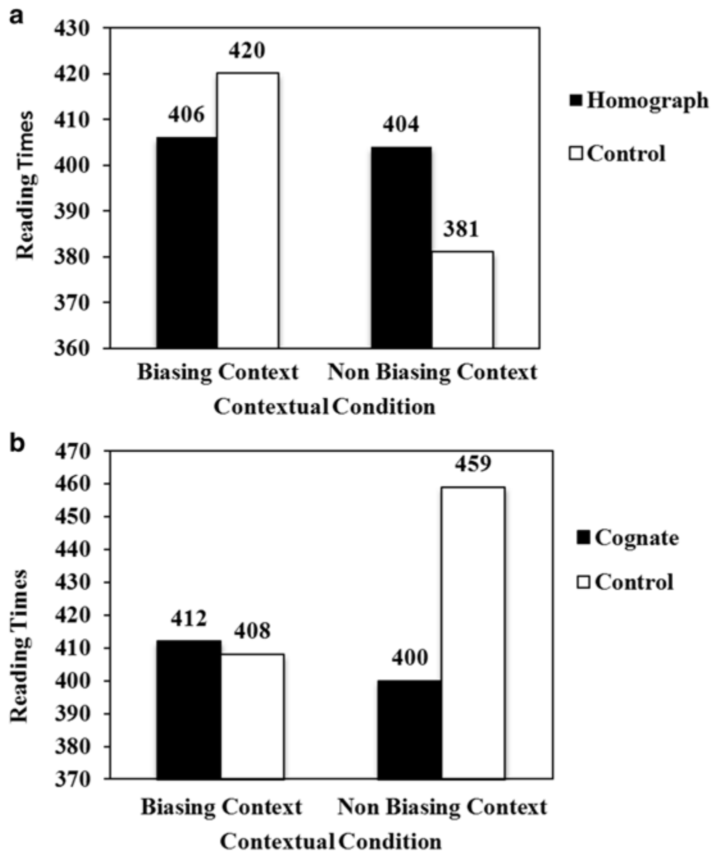


Fig. 5.4 Mean reading times for interlingual homographs (**a**) and cognates as a function of contextual condition (**b**)

In summary, we have provided an overview of some of our work utilizing the SMW. The data reviewed using this task, so far, reveal comparable findings as those reported using such tasks as RSVP, and to a certain extent eye-tracking (see for example, Altarriba et al., 1996; Libben & Titone, 2009; Titone et al., 2011, Chap. 8). Moreover, the results produced by the SMW using cross-language/bilingual materials are congruent with previous findings showing word frequency effects (Altarriba et al., 1996; Ferreira, Anes, & Horine, 1996), contextual effects (Altarriba et al., 1996; Libben & Titone, 2009; Schwanenflugel & LaCount, 1988; Titone et al., 2011), code-switching, homograph interference, and cognate interference effects. Additionally, the results exhibited by the SMW in Heredia et al. (2003) are also consistent with earlier work utilizing the AMW (Heredia, Stewart, & Cregut, 1997; Heredia & Vaid, 2002; see also Ferreira, Anes, & Horine, 1996). Briefly, Heredia et al. presented Spanish English bilinguals with spoken sentences such as (4a) ↓*Erika* ↓*estuvo*↓*buscando a la*↓*TEACHER*↓ *pero nunca*↓*la encontró*↓ (translation: "Erika was looking for the TEACHER but she never found her"; note: segments are depicted with the symbol "↓"). Sentences were taken from Altarriba et al. (1996) and translated into Spanish. However, in addition to varying word frequency and preceding context, Heredia et al. manipulated whether the target word (English in this case) embedded in a Spanish sentence was pronounced according to its standard English pronunciation (the *code-switch condition*) or as a *language borrower*, in which the English target word was pronounced with a strong Spanish accent. The general idea behind this manipulation was to simulate an English speaker with a very heavy Spanish accent (cf. *ship* |ʃɪp| vs. *sheep* |ʃip| for "standard English," and *ship* |ʃip| vs. *sheep* |ʃip| for a Spanish accented English pronunciation). Times between button presses or interresponse times (IRTs) were recorded for code-switch or language borrower targets. Analyses were performed on the IRTs and difference times (DTs). DTs are computed by subtracting the IRT minus the duration of the segment during the digitizing of the word. For example, if the IRT for the target *TEACHER* in sentence (4a) was 802 ms, and the time taken to record the target (duration time) was 493 ms, the DT would be 309 ms (see for example, Heredia & Stewart, 2002).

Like the results from Altarriba et al. (1996) and Heredia et al. (2003), findings with the AMW revealed (1) a robust effect of constraint context, where targets under the constraint context condition produced longer IRTs and DTs; (2) IRTs were shorter for high-frequency targets. However, this effect was marginally significant and nonreliable for DTs. (3) Code-switch targets were faster than language borrowers, but only for IRTs. (4) Of greater interest was the interaction of context by type of target (i.e., code-switch vs. borrower). Bilinguals were equally fast in their responses to both targets under low-constraint context conditions. Under high-constraint context conditions, code-switch targets were responded to faster, for both IRTs and DTs. (5) The interaction of target type by word frequency showed equivalent IRTs and DTs for both target types in the high-word-frequency condition. However, code-switch targets revealed faster IRTs and DTs than language borrowers. In summary, the VMW and its derivatives are clearly underutilized in the bilingual reading literature. However, current work in the related field of second-language reading is beginning to use this technique in addressing broader linguistic and syntactic processing issues such as syntactic ambiguity (Hopp, 2006; Jackson & Roberts, 2010), syntactic creativity (Dussias & Piñar, 2010; Dussias & Scaltz, 2008; Schulz, 2011), and second language processing (Dekydtspotter & Renaud, 2009; Marinis, 2003).

Advantages and Disadvantages

The VMW and its variants have several advantages over online gaze-contingency and other reading tasks that are worth noting. One particular aspect making this task attractive for researchers in bilingual reading is that it allows the investigator to have direct control over the amount of text read by participants during stimulus presentations. Sentences can be presented from one to three words at a time or in larger fragment segments depending on the working hypothesis. Moreover, the VMW exerts additional experimental control by preventing readers from looking to previous text after reading, thus forcing the reader to attend to each individual linguistic segment (Papadopoulou, 2005). Also, since the VMW relies on visual presentations, it does not suffer from increased processing load and language comprehension associated with auditory tasks such as its variant the AMW and the *cross-modal lexical priming task* (see Ferreira, Anes, & Horine, 1996; see also Chap. 6).

Other qualities associated with the VMW task worth highlighting include: (1) its word-by-word presentation modality runs smoothly and is not extremely unnatural or cognitively difficult for participants (Ferreira, Henderson, et al., 1996), thus enabling readers to read passages even if the presentation is not "normal reading"; (2) it allows participants to complete the task at their own pace and in their primary modality of language use (i.e., visual/written or spoken presentations); (3) it is reliable and sensitive to identify known psycholinguistic effects found in the mainstream reading literature (e.g., context, word frequency; e.g., Aaronson & Ferres, 1984; Ferreira, Henderson, et al., 1996; Kinnunen & Vauras, 2010; see also Witzel, Witzel, & Forster, 2012); and (4) results produced by this task are comparable to those of eye-tracking and both reveal similar effects in terms of reading time, consistency and accuracy (Kinnunen & Vauras, 2010). However, the eye-tracking technique is able to examine both early and late reading processes using multiple eye movement indexes (see Chaps. 4 and 8), so it is unclear and perhaps difficult to discern whether the VMW is capable of measuring early vs. late stages of the reading process. It may very well be the case that the VMW is most sensitive to late reading processes that reflect higher order processes such as semantic integration (cf. Witzel et al., 2012). Overall, the VMW is a straightforward, noninvasive procedure, and highly affordable since no specialized equipment or software is needed to construct and run experimental trials (Aaronson & Ferres, 1984). Finally, it pairs well with other behavioral methodologies (e.g., event-related potentials measuring electrical activity on the cerebral cortex in real time; see Chaps. 11 and 12) to study reading effects (Ditman, Holcomb, & Kuperberg, 2007).

By and large, the VMW is a highly reliable task producing highly replicable results; however, there are some important experimental artifacts associated with this task that bilingual researchers might want to take into consideration. (1) The VMW and its variants have been criticized as unnatural due to the single word segmentation and the required constant button pressing (Aaronson & Ferres, 1984). As a consequence, this task might not represent an accurate measure of the time taken by participants to read a target word or phrase since they might become fatigued as they press a button constantly for a considerable amount of time (Rouet & Passerault, 1999). (2) The way the VMW is set up creates some issues, as well. A notable artifact associated with this task, and noted in the literature, is that it might be open for strategy development. In the course of reading a sentence word-by-word, participants might develop a predictable button-pressing rhythm which might possibly invalidate reading times and their eventual analysis and interpretation (Witzel et al., 2012). To avoid these possible confounding effects, it is suggested that investigators adopt a "tapper" strategy, where they look at button press reading time variability and eliminate those participants that do not vary in their button presses (Witzel et al., 2012). (3) When investigating linguistic and contextual variables, reading time is a sensitive measure to be used, yet when sentences are presented word-by-word or phrase-by-phrase, processing may become difficult for participants, thus resulting in longer reading times (Potter, 1984). Longer reading times may not necessarily translate into a cognitive cost; alternatively, these longer reading times may be a result of the structural nature of the task or other possible participant strategies. To avoid these potential artifacts, it is critical for experimenters to consider the type of linguistic stimuli to be used, especially when investigating cross-linguistic or bilingual effects (Rayner, 1993). (4) This task is open to *spillover effects*, in which previous stimulus' (word's) properties and features transfer to subsequent stimuli (Shvartsman, Lewis, & Singh, 2014). A spillover effect shows incomplete processing of a word segment, which, in turn, carries over to the next word segment. These effects can be minimized by 3-words-at-a-time displays (i.e., participants see three different words at a time which they need to read and process before they move to the next triplet) or by varying the lengths of previous word displays (Mitchell, 1984).

Other possible artifacts related to the VMW task include emphasizing distinctive features of the critical stimuli (e.g., using uppercase characters and/or using different font colors) that encourage strategic processing (cf. García et al., 2015), and differences across languages (Rayner, 1993). English readers have been found to focus about four letters to the left of a fixation. The direction in which a language is read has an effect on the effective span of vision when reading. Right-to-left read languages have larger spans than those languages that are read-left-to-right. Languages with more dense orthographies (e.g., Polish, a language with many diacritics) have also been shown to have smaller spans. In short, bilingual reading researchers must consider language and orthographic density as they implement the VMW or its variants.

Conclusions and Final Thoughts

In this chapter, we reviewed the VMW paradigm and its variants, the SMW and AMW, as well as their usage in investigating issues related to the comprehension of code-switched/bilingual relative to monolingual sentences. We also reviewed data from studies looking at the effects of grammatical gender, as well as ambiguity resolution in bilingual reading. The primary goal of this chapter has been to present researchers and bilingualists with specific examples of how the VMW has successfully been utilized in the bilingual reading literature. The data so far suggest that the VMW is a reliable and robust reading task sensitive to cross-language or bilingual effects (e.g., code-switching, grammatical gender, homograph, and cognate), as well as other effects (e.g., word frequency, and context effects) found in the mainstream psycholinguistic literature. As with any experimental paradigm, there are strengths (e.g., direct control over stimulus presentation, affordability, experimental reliability, and validity) and weaknesses (e.g., spillover effects, response strategies, unnatural sentence reading segmentation) associated with this task. The decision as to whether the VMW is the appropriate technique for one's project would ultimately depend on the researcher's question of interest, given the various reading techniques described in this volume. Whether a bilingual version of the VMW is sensitive to early and late reading processes remains to be seen. How is the VMW different or similar to other tasks such as RSVP (Chap. 4), eye-tracking (Chap. 8), and the maze task (Forster, 2010; Forster, Guerrera, & Elliot, 2009; Qiao, Shen, & Forster, 2012; Witzel et al., 2012; see also Heredia, Altamira, Cieślicka, & García, 2012, for a bilingual version of the maze task), in regards to bilingual reading? Clearly, more studies of the type reported by Altarriba et al. (1996; RSVP vs. eye-tracking), discussed in this chapter, and Witzel et al. (2012; eye-tracking vs. AMW, vs. the maze—a sentence integration task) are needed to fully understand task demands and the extent to which the VMW measures early or late reading processes in bilingual reading. Further studies, for example, may consider replicating Altarriba et al.'s original study employing both a VMW and eye-tracking paradigm. In addition to including a homogenous bilingual population as the one described by Heredia et al. (2003), other important variables that might prove theoretically interesting would include language dominance and bilingual directionality where Spanish and English are the L1 and L2, respectively, and vice versa. In conclusion, we hope that the information presented in this chapter will provide the bilingual researcher with a general understanding of the advantages and limitations of the VMW in the bilingual reading domain. It is hoped that our discussion presented here serves as a starting point to increase our limited understanding of the VMW and bilingual reading comprehension.

List of Keywords

Animate, Auditory moving window (AMW), Borrower, Code-switching, Cognate facilitation effect, Contextual effects, Cumulative visual moving window, Difference times (DTs), Experimenter effects, Eye-tracking, Feature Restriction Model, Grammatical gender, High constraint context, Homograph inhibition, Homograph interference effect, Implicit knowledge, Inanimate, Interresponse times (IRTs), Noncumulative visual moving window, Nonselective Hypothesis, Online tasks, Participant effects, Proficiency, Reflexive anaphoric reference, Researcher effects, Selective Hypothesis, Spillover effects, Stationary moving window (SMW), Tacit knowledge, Visual moving window (VMW), Word frequency effect

Review Questions

1. Why might a researcher decide to utilize a moving window paradigm over other paradigms such as RSVP?
2. What are the advantages and disadvantages to using a cumulative rather than a noncumulative visual moving window?
3. In this chapter, the effects of *word frequency*, *cognate status*, and *sentence context* are discussed. What other variables in sentence processing may also be interesting to study using the VMW paradigm?
4. What are the pros and cons in employing the VMW task to study bilingual or second-language reading processes? What is your overall view of the VMW task after reading this chapter?
5. How would you react, after reading this chapter, if you heard a conversation in which a researcher claims that the VMW produces similar results to eye-tracking?

Suggested Student Research Projects

1. In this project, you will build a "low tech" VMW. Find the article by Altarriba et al. (1996) from the references below and select ten mixed-language and ten monolingual sentences. To simplify the study, select low-constraint context sentences with high-frequency words only. Since context is not an issue in your study, you can try to shorten the sentences as much as possible (e.g., *The boy saw a CARTA/LETTER on the street*). Type the sentences on cardstock paper so that it does not bend easily. It would probably work better if the paper were cut in half. Type one sentence per sheet of paper. The font should be large enough so that each sentence fits on the paper. Now, take another blank cardstock sheet of paper (the thicker the better) and cut an aperture (or window) large enough to be able to read your sentences one word at a time. Now you have a "window" that you can move throughout the sentence. Modify it so that it can slide back and forth throughout the sentence with ease. You now have the original "visual moving window." As it is always the case, refine your creation and prepare your experimental stimuli, and review Altarriba et al.'s method section so that you become very familiar with the overall materials and procedures.

2. Now that you have a "low tech" VMW, find a stopwatch to measure (in seconds or milliseconds) the amount of time it takes to read each sentence. As you execute your experiment, you need to keep track of the target words for each sentence so that you know when to start and stop the timing. There are two ways in which you can run the experiment. One way is to measure the reading time up to the target word (monolingual or mixed). Note that this method will produce large reading times. So it is important that your sentences are of the same length, and your target words are identical in length or have the same number of characters. A second possibility is to start your stopwatch immediately after the participant utters the word before the target and stop the timer the moment s/he pronounces the target word. So in this experiment, you are testing whether mixed-language (e.g., *The boy saw the CARTA…*) takes longer to read than monolingual sentences (e.g., The boy saw the LETTER…). When you are ready to run the experiment, please discuss the details with your professor and obtain Institutional Review Board (IRB) approval from your college or university. What do you expect to find? After you ran your experiment, did you find that mixed-language sentences took longer to read? Can you explain why?
3. In this project you will investigate the grammatical gender interpretation of compound words by bilinguals during code-switching. All nouns in Spanish have grammatical gender, with obligatory gender marking on preceding articles (e.g., *la casa* and *el perro*, the feminine and masculine forms of the determiner "the," respectively). As you already know, compound words (e.g., *peanut butter*) follow a left-headed compound structure in the bilingual's L1 (Spanish), and a right-headed structure in the L2 (English). Therefore, one would expect that the correct gender in Spanish for the English compound word *peanut butter* would be feminine if one were to translate into Spanish (e.g., *la mantequilla de cacahuate*). Now think of the possibility in which you prevent participants from translating these compound words by pressuring them to respond as fast and as accurately as possible. So, for this project, you can vary presentation time. One group may be assigned to a slow presentation in which s/he is exposed to the compound word for up to 5 s, and the other to a fast presentation (one or fewer seconds). The general idea here is that the group in the slow presentation condition would have enough time to possibly generate the Spanish translation and provide the correct response; whereas the group in the fast presentation condition will not and will resort to using the linguistic rule that is appropriate for his/her L1. For the purpose of this study, select a group of bilinguals whose L1 is Spanish and L2 is English or vice versa, but do not mix them. As you become more comfortable with this study, you can expand it to look at both types of bilinguals simultaneously. Now choose about 20 English compound words that are highly frequent (e.g., *peanutbutter*: "mantequilla de cacahuate," *spiderman*: "hombre araña"). The simplest way to run the experiment is to use powerpoint or libreoffice (www.libreoffice.org)/openoffice (www.openoffice.org), since this presentation software would allow you to manipulate the timing. Another possibility is to head over to OpenSesame (http://www.cogsci.nl/software/opensesame), a free/open source experiment builder which would allow you to more precisely manipulate stimulus presentation at the

millisecond level, and record both response times (i.e., how long participants take to respond), as well as the number of errors they make as they respond. Notice that with the presentation software you would only be able to look at the number of errors or accuracy. Visit the OpenSesame webpage and follow the instructions on how to construct a lexical decision task that requires participants to determine if a presented word is a legal (e.g., *top*: PRESS THE YES BUTTON) or a nonlegal (*tup*: PRESS THE NO BUTTON) word in English. For this project you could simply substitute the YES/NO BUTTONS for MASCULINE/FEMININE.

What would you expect to find? One possibility, under the fast condition, is that Spanish-dominant (Spanish-English) bilinguals might be more likely to assign the masculine (e.g., *el peanutbutter*) to the compound word following the Spanish structure, where *peanut* ("el cacahuate") is the head of the compound. Conversely, English-dominant bilinguals might consider *butter* ("la mantequilla") the head of the compound and assign the feminine article (e.g., *la peanutbutter*).

Were you able to find differences between the two timing conditions? Were you able to run the experiment using both techniques? If you did, were you able to obtain similar error or accuracy patterns between the two tasks?

Related Internet Sites

Auditory moving window: https://en.wikipedia.org/wiki/Auditory_moving-window
E-prime moving window script: http://step.psy.cmu.edu/scripts/Linguistics/Just1982.html
Experimental materials: http://www.tamiu.edu/~rheredia/materials.html
Experimental stimuli: http://www3.nd.edu/~memory/research.php
Gaze-Contingent Paradigm: https://en.wikipedia.org/wiki/Gaze-contingency_paradigm
Maze Task: http://www.u.arizona.edu/~kforster/MAZE/index.htm
PsyScope moving window scripts: http://psy.ck.sissa.it/psy_cmu_edu/scripts/index.html
Reading task: https://wiki.brown.edu/confluence/display/kertzlab/Self-Paced+Reading+Task

Suggested Further Reading

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, 17(1), 137–153.

Becker, C. A. (1979). Semantic context and word frequency effects in word recognition. *Journal of Experimental Psychology*: *Human Perception and Performance*, *5*, 252–259.

Lederberg, A. R., & Morales, C. (1985). Code switching by bilinguals: Evidence against a third grammar. *Journal of Psycholinguistic Research*, *14*, 113–136.

Macnamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, *10*, 480–487.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a speech model: The effect on lexical access. *Memory & Cognition*, *12*, 380–386.

Waters, G. S., & Caplan, D. (2004). Verbal working memory and on-line syntactic processing: Evidence from self-paced listening. *The Quarterly Journal of Experimental Psychology*, *57*(1), 129–163.

References

Aaronson, D., & Ferres, S. (1984). The word-by-word reading paradigm: An experimental and theoretical approach. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 31–68). Hillsdale, NJ: Lawrence Erlbaum.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Bolger, P., & Zapata, G. (2011). Semantic categories and context in L2 vocabulary learning. *Language Learning, 61*(2), 614–646.

Carpenter, P. A. (1984). The influences of methodologies on psycholinguistic research: A regression to the Whorfian Hypothesis. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 3–11). Hillsdale, NJ: Lawrence Erlbaum.

Cieślicka, A. B. (2013). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research, 42*, 475–503.

Cieślicka, A. B. (2015). Idiom acquisition and processing by second/foreign language learners. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 208–244). New York: Cambridge Press.

Cook, V. J. (1990). Timed comprehension of binding in advanced L2 learners of English. *Language Learning, 40*(4), 557–599.

Cook, V. J. (1997). The consequences of bilingualism for cognitive processing. In A. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 270–300). Mahwah, NJ: Lawrence Erlbaum.

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. *The Quarterly Journal of Experimental Psychology, 63*(7), 1266–1303.

Dekydtspotter, L., & Renaud, C. (2009). On the contrastive analysis of features in second language acquisition: Uninterpretable gender on past participles in English-French processing. *Second Language Research, 25*(2), 255–267.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology, 44*, 927–935.

Dussias, P. E. (1997). Sentence matching and the functional head constraint in Spanish/English codeswitching. *Spanish Applied Linguistics, 1*, 114–150.

Dussias, P. E. (2001). Psycholinguistic complexity in codeswitching. *International Journal of Bilingualism, 5*, 87–100.

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of *wh-gaps* by Chinese-English second language speakers. *Second Language Research, 26*(4), 443–472.

Dussias, P. E., & Scaltz, T. R. C. (2008). Spanish-English L2 speakers use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica, 128*, 501–513.

Ferreira, F., Anes, M. D., & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research, 25*, 273–290.

Ferreira, F., & Henderson, J. M. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 555–568.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 324–335.

Forster, K. I. (2010). Using a maze task to track lexical and sentence processing. *The Mental Lexicon, 5*, 347–357.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods, 41*, 163–171.

García, O., Cieślicka, A. B., & Heredia, R. R. (2015). Nonliteral language processing and methodological considerations. In R. R. Heredia & A. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 117–168). New York: Cambridge University Press.

Heredia, R. R., Altamira, W. A., Cieślicka, A. B., & García, O. (2012, November). *Bilingual lexical access: Interlingual homographs & the grammaticality task*. Poster presented at the 53rd annual meeting of the Psychonomic Society, Minneapolis, MN.

Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*, 164–168.

Heredia, R. R., Cieślicka, A. B., & García, O. (2010a, March). B*ilingual exhaustive language activation: Homograph and cognate processing*. Poster Presented at the 53rd Annual CUNY Conference on Human Sentence Processing. New York: University of New York.

Heredia, R. R., Cieślicka, A. B., & García, O. (2010b, November). *Bilingual sentence processing: Multiple language activation*. Paper presented at the 51st annual meeting of the Psychonomic Society, St. Louis, MO.

Heredia, R. R., Martínez, K. M., Clark, D., & Moreno, M. (2003, September). *Processing code-switched sentences in Spanish-English bilinguals: Context and word frequency effects*. Paper presented at the XIII Congress of the European Society for Cognitive Psychology, Granada, Spain.

Heredia, R. R., & Muñoz, M. E. (2015). Metaphoric reference: A real-time analysis. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 89–116). New York: Cambridge Press.

Heredia, R. R., & Stewart, M. T. (2002). On-line methods in bilingual spoken language research. *Advances in Psychology, 134*, 7–28.

Heredia, R. R., Stewart, M. T., & Cregut, I. (1997, November). *Bilingual on-line sentence processing: Frequency and context effects in code switching.* Poster presented at the 38th annual meeting of the Psychonomic Society, Philadelphia, PA.

Heredia, R. R., & Vaid, J., (2002, March). *Processing code-switched sentences: Effects of semantic constraint, guest word phonology and guest word frequency*. Poster presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York, NY.

Herman, R. (Producer), & Frakes, J. (Director). (1996). *Star trek: First contact* [Motion picture]. Los Angeles: Paramount Pictures.

Hillert, D. (2002). Cross-linguistic aspects of anaphor resolution. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 237–248). Amsterdam: North-Holland/Elsevier Science.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22*(3), 369–397.

Jackson, C. N., & Roberts, L. (2010). Animacy affects the processing of subject-object ambiguities in the second language: Evidence from self-paced reading with German second language learners of Dutch. *Applied Psycholinguistics, 31*, 671–691.

Just, M., Carpenter, P., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology, 111*(2), 228–238.

Katz, A. N., & Ferretti, T. R. (2001). Moment-by-moment reading of proverbs in literal and nonliteral contexts. *Metaphor and Symbol, 16*, 193–221.

Kinnunen, R., & Vauras, M. (2010). Tracking on-line metacognition: Monitoring and regulating comprehension in reading. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 209–229). New York: Springer.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 381–390.

Liceras, J. M., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua, 118*(6), 827–851.

Marinis, T. (2003). Psycholinguistic techniques in second language acquisition research. *Second Language Research, 19*(2), 144–161.

Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 69–89). Hillsdale, NJ: Lawrence Erlbaum.

Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eye-tracking, ERP and beyond* (pp. 15–32). New York: Psychology Press.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research, 21*(2), 98–120.

Potter, M. C. (1984). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 91–118). Hillsdale, NJ: Lawrence Erlbaum.

Qiao, X., Shen, L., & Forster, K. I. (2012). Relative clause processing in Mandarin: Evidence from the maze task. *Language and Cognitive Processes, 27*, 611–630.

Rayner, K. (1993). Eye movements in reading: Recent developments. *Current Directions in Psychological Science, 2*(3), 81–85.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading* (2nd ed.). New York: Psychology Press.

Rouet, J., & Passerault, J. (1999). Analyzing learner-hypermedia interaction: An overview of online methods. *Instructional Science, 27*, 201–219.

Schulz, B. (2011). Syntactic creativity in second language English: Wh-scope marking in Japanese-English interlanguage. *Second Language Research, 27*(3), 313–341.

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(2), 344–354.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.

Shvartsman, M., Lewis, R. L., & Singh, S. (2014). *Computationally rational saccadic control: An explanation of spillover effects based on sampling from noisy perception and memory* [Data file]. Retrieved from http://www-personal.umich.edu/~rickl/pubs/shvartsman-et-al-2014-cmcl.pdf

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*, 251–272.

Swinney, D., Love, T., Walenski, M., & Smith, E. E. (2007). Conceptual combination during sentence comprehension: Evidence for compositional processes. *Psychological Science, 18*(5), 397–400.

Swinney, D., & Osterhout, L. (1990). Inference generation during auditory language comprehension. In A. Graesser & G. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25, pp. 17–33). New York: Academic.

Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 171–207). New York: Cambridge University Press.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1412–1431.

Valenzuela, E., Faure, A., Ramírez-Trujillo, A. P., Barski, E., Pangtay, Y., & Diez, A. (2012). Gender and heritage Spanish bilingual grammars: A study of code-mixed determiner phrases and copula constructions. *Hispania, 95*(3), 481–494.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparison of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research, 41*, 105–128.

Chapter 6
Priming and Online Multiple Language Activation

Anna B. Cieślicka and Roberto R. Heredia

Abstract This chapter discusses cross-language activation in the course of processing language by bilingual speakers. We first discuss the *cross-modal lexical priming paradigm* (CMLP), a powerful tool to explore online multiple language activation. We next provide an overview of research concerning multiple language activation in the course of bilingual lexical processing. Finally, we present results of four experiments examining the effects of context in connected speech on cross-language priming in Spanish-English bilinguals. Participants in Experiment 1 listened to sentences in Spanish, their first language, and named Spanish and English targets, related or unrelated to a critical prime within the sentence. Experiment 2 was similar to Experiment 1, except that prior context was biased toward the critical prime. Experiments 3–4 were identical to Experiments 1–2, respectively, but with sentences in English, their second language. Comparable cross-language priming was observed for Experiments 1–2. Likewise, Experiments 3–4 showed similar priming patterns. However, the priming effect was significantly higher for the L2–L1 language direction. Results are discussed in terms of language dominance mechanisms and the *Revised Hierarchical Model* of bilingual memory representation.

Introduction

This chapter is motivated by the observation that, in some bilingual communities, bilinguals mix their two languages simultaneously in the course of spoken interaction (cf. Heredia & Stewart, 2002; Hummel, 1993; see also Heredia & Altarriba, 2001). To illustrate, consider sentences (1a–c) below.

A.B. Cieślicka, Ph.D. (✉) • R.R. Heredia, Ph.D.
Department of Psychology and Communication, Texas A&M International University, 5201 University Boulevard, Laredo, TX 78041-1900, USA
e-mail: anna.cieslicka@tamiu.edu; rheredia@tamiu.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_6

(1a) *It is difficult to admit that a WAR sometimes brings more profits than losses.*
(1b) *It is difficult to admit that a GUERRA sometimes brings more profits than losses.*
(1c) *Soldiers are trained for combat and GUERRA and that is why we invest in them.*

These sentences exemplify three important issues addressed in the current chapter. First, the use of both languages by a bilingual person in the same sentence indicates that the two languages must be simultaneously activated in the course of bilingual language processing. Second, notice that, unlike the monolingual English sentence (1a), Spanish words are embedded in sentences (1b–c). In these sentences the Spanish word *guerra* replaces the English word *war*. This linguistic phenomenon is known as code-switching. It is prevalent amongst bilinguals and it occurs automatically (Grosjean, 1988; Heredia & Altarriba, 2001; Li, 1996). Third, in sentence (1b) the preceding context does not influence or provide any information about the Spanish target word. In contrast, prior context in sentence (1c) is biased towards the Spanish code-switched word. This leads us to the following questions that we are discussing in this chapter: (1) Do both languages of a bilingual person become automatically simultaneously activated in the course of language processing? (2) How do bilinguals process and access information across languages during the processing of code-switched sentences? (3) What are the effects of context in the comprehension of code-switched sentences such as (1b–c) above? This chapter addresses each of these questions in turn. It starts with the description of the *cross-modal lexical priming task* (CMLP), the methodological paradigm that has been most frequently employed to address online multiple language activation. It then provides a brief overview of research into activation of languages in the course of bilingual processing and focuses specifically on the most influential model of bilingual lexical representation, the *Revised Hierarchical Model*. Next, research into code-switched sentence processing and the role of context are briefly examined. We then present the study with Spanish-English bilinguals that we conducted to look at the effects of context in connected speech on cross-language priming.

Priming and the Cross-Modal Lexical Priming Paradigm

The CMLP paradigm is a variation of the priming paradigm which combines auditory and visual modes of stimulus presentation. Priming paradigms have had a long tradition and have been used extensively in psycholinguistic research to investigate semantic memory (Meyer, Schvaneveldt, & Ruddy, 1975; Warren, 1977). Priming as such has been defined as a *facilitative effect of the presentation of a word on the identification or classification of a related word* (Masson, 1995, p. 3). A number of techniques have been developed to assess the priming effect.

The most basic of them are the *lexical decision task* and the *naming task*. In the lexical decision task, participants are presented with a string of letters on a computer screen and are asked to quickly decide if the string of letters (i.e., a word) is a legitimate word in a given language or a nonword. In the naming task, participants are to simply name a presented word. The time taken to make the word/nonword decision is called the *lexical decision time* or *reaction time* (RT), whereas the time taken to name a letter string is called the *naming latency*. Both of them are affected by different types of experimental manipulation. For example, presenting another stimulus, called the *prime* (e.g., *cat*) prior to the target (e.g., *dog*) will affect how quickly the target word is named in the naming task and recognized as a legitimate word in the lexical decision task. Priming experiments carried out with primes semantically and associatively related to the target show a decrease in RT in a lexical decision task and a shorter naming latency in the naming task (for a review, see Neely, 1991). Decrease in reaction time to the target caused by the earlier presentation of a prime is known as a *positive priming effect* (Jiang & Forster, 2001).

In the CMLP paradigm, participants are simultaneously involved in a passive and active task. The passive task consists in attending to spoken sentences presented auditorily via headphones. At some point during the auditory presentation, a visual target appears on the computer screen and participants perform an active lexical decision (i.e., decide, as quickly and as accurately as possible, if a displayed probe or target is a word or a nonword). The probes for lexical decision are presented at various points throughout the sentence, depending on the experimental focus. For example, during the auditory presentation of the sentence, *My diabetic cat is not at all bothered by the daily* shots[*1], *as he has been getting them for over a year now*, the word *gun* is displayed visually at the offset of the word *shots* (depicted by the subscript [*1]), and the participant makes a lexical decision on that word. The assumption behind the CMLP technique is that facilitation of a lexical decision will be demonstrated for those visual targets whose meanings have been primed by the auditorily presented input. Thus, if a participant's lexical decision to the visual target *gun* is facilitated, in that it is shorter than the lexical decision to its matched control word (e.g., *nun*), then it can be concluded that a weapon-related meaning of the word *shot* has been automatically activated, even if it is contextually inappropriate. Because of this ability to detect automatic activation of the different senses of lexically ambiguous words like *shot*, the CMLP paradigm has been extensively used in lexical ambiguity research to address the question of multiple access during the comprehension of ambiguous words (e.g., Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Simpson, 1981; Swinney, 1979; Swinney, Love, Walenski, & Smith, 2007; Tabossi, 1988; Tanenhaus & Donnenworth-Nolan, 1984), as well as to explore the mechanisms underlying figurative language processing (see, for example, Blasko & Connine, 1993; Cacciari & Tabossi, 1988; Cieślicka, 2006, 2007; Hillert & Swinney, 2001; Tabossi & Zardon, 1993; Titone & Connine, 1994; Van de Voort & Vonk, 1995). The paradigm has been also employed to examine the effects of context in bilingual language processing (Cieślicka 2006); Heredia & Muñoz, 2015; Heredia & Stewart, 2002).

One of the strengths of the CMLP paradigm is that it does not draw participants' attention to the presence of ambiguities in the experimental material. In addition, it prevents the development of anticipatory strategies by participants, as it is not predictable in terms of the point at which the visual target appears. For those reasons, the CMLP is viewed as a highly reliable experimental tool, sensitive to lexical access and processing (Onifer & Swinney, 1981; see also Garcia et al., 2015, for an extensive discussion of the CMLP paradigm).

However, the paradigm is not free from criticism, as some reservations have been raised against its use to tap online multiple activation. For example, it has been noted that multiple access demonstrated in the ambiguity studies employing the CMLP may reflect *backward priming*, defined as *temporal overlap in the processing of two words, … . [which] can be thought of as "mutual priming" analogous to that which occurs between simultaneously presented words* (Van Petten & Kutas, 1987, p. 191; see also Burgess, Tanenhaus, & Seidenberg, 1989). Under this mechanism, the priming of targets (e.g., *gun*) related to contextually inappropriate meanings of ambiguous words (such as [insulin] *shots* in the earlier example) results not from multiple access but from the backward priming effect, whereby the subsequent presentation of a target related to the unbiased meaning of an ambiguous word evokes activation of this previously irrelevant meaning (i.e., *gun shots*). This newly activated meaning is hence processed concurrently with its related target, leading to shortened response latency for this target. Different efforts to eliminate the backward priming effect were undertaken in studies employing the CMLP paradigm, but they failed to yield conclusive results (Glucksberg, Kreuz, & Rho, 1986).

In addition, the CMLP paradigm has been challenged on methodological grounds. Since it is a cross-modal task, consisting of an auditorily presented context which includes an ambiguous prime and a visually presented target which requires a lexical decision response, it places severe attentional demands on the participant (Sereno, 1995). Consequently, being required to constantly switch between the modes, the participant may resort to the strategy of preserving only the last one or two words of the auditory context in his or her articulatory rehearsal. Should this be the case, s/he would be responding to the visual target based on very limited contextual information and so the task might actually reflect context-free priming.

Despite those reservations regarding the use of the CMLP to explore the processing of lexical ambiguities, it is still considered a highly reliable tool to investigate online aspects of bilingual lexical access and figurative language processing (e.g., Heredia & Stewart, 2002). The paradigm can reliably reflect online processes, without being susceptible to backward priming, provided the primes are embedded in a sentential context (Cacciari & Tabossi, 1988). It has also been suggested that, while the CMLP employing the naming task can indeed be compromised by the backward priming effect, this is not true of the lexical decision task used in combination with the auditorily presented input (Cacciari & Tabossi, 1988).

Multiple Language Activation in Bilingual Lexical Processing

The question of multiple language activation in bilingual lexical processing has a long research tradition in the bilingual literature. While under the *language selective view*, only one language is activated and accessed at a time (e.g., Gerard & Scarborough, 1989), according to the *language nonselective access view*, all the languages known to an individual are activated simultaneously (Beauvillain & Grainger, 1987; Dijkstra, Grainger, & van Heuven, 1999). In one of the studies addressing this controversy, Gerard and Scarborough (1989) presented Spanish-English bilinguals with *interlingual homographs* (words which share spelling but have different meanings across languages, for example *SIN*, which means *without* in Spanish and denotes something morally unacceptable in English) and asked them to make a lexical decision on the presented targets, either in Spanish or English blocks. The results turned out to reflect the homographic words' frequency in the language of the response (i.e., frequency of Spanish words in the Spanish block and of English words in the English block), thus suggesting that participants were accessing each of their two lexicons selectively when they performed a monolingual task (see also Scarborough, Gerard, & Cortese, 1984).

Dijkstra, Van Jaarsveld and Ten Brinkee (1998) extended Gerard and Scarborough's (1989) study with a group of Dutch-English bilinguals. Experiment 1 replicated the findings of Gerard and Scarborough, in that no frequency effect was obtained in RTs to the English and Dutch readings of the interlanguage homographs used in the study. In Experiment 2, Dutch stimuli were added to the set of English targets, thus requiring the participants to respond "NO" in the English lexical decision task. This manipulation induced strong inhibition to homographs as compared to English controls. In addition, the frequency of the English or Dutch readings of the homographs had a significant effect on response latencies, unlike in the previous experiment. In Experiment 3, stimulus lists for the lexical decision included both Dutch and English items and the participants were instructed to respond "YES" to words in either language. The results showed facilitation of interlingual homographs as compared to monolingual control items and a strong frequency effect. Overall, Dijkstra et al. (1998) interpreted these results as evidence for nonselective language access which is sensitive to task demands and stimulus list composition, the claim further corroborated in a series of experiments by Dijkstra, Timmermans, and Schriefers (2000).

More specifically, Dijsktra et al. (2000) modified their methodology, task demands, and the specifics of instructions, which they viewed as essential factors influencing bilingual lexical processing. In Experiment 1, they presented highly proficient Dutch speakers of English with a list of Dutch-English non-cognate homographs and Dutch and English control items matching the homographs in terms of word frequency and length. The participants were instructed to perform a language decision task (i.e., to press one button when an English word was shown and another one if a Dutch word appeared on the screen). The results revealed that participants' RTs were slower and they opted less often for the English language decision when they saw interlingual homographs. Moreover, evidence of nonselective lexical access was obtained, as the participants reacted to the highest frequency reading of the homograph. In cases when both readings were of comparably low frequency, participants opted for the Dutch reading, which was interpreted as a compensation strategy when dealing with their weaker language (English).

Experiment 2 further manipulated task instructions, as this time Dutch-English bilinguals responded to the same list of stimuli but were instructed to respond only to English words (the so called go/no-go task, which requires participants to respond only if a stimulus from a particular language is presented). Slower RTs and a higher miss rate were recorded in response to those homographs whose Dutch reading had a higher frequency than the English one. Finally, in the third experiment, a similar language go/no-go task was employed, but this time participants were instructed to respond only to Dutch words. Like in the previous experiment, longer RTs and more errors were obtained in reaction to interlingual homographs. The results were also influenced by the reading frequency of the interlingual homographs, such that homographs that had low frequency in Dutch and high frequency in English took the longest to respond. According to Dijkstra et al. (2000), this pattern of results is compatible with the activation of both lexicons and failure to completely suppress lexical items from the English lexicon. Overall, the results obtained by Dijkstra et al. (2000) were taken as evidence for language nonselective access, since participants' responses were frequency-dependent in both target and nontarget languages. Presence of items from both of the participants' languages in the stimulus lists prevented them from being able to completely suppress the nontarget language, even if this would optimize their performance in those tasks which required responding only to target language stimuli.

In the domain of lexical-level processing with single items used as experimental stimuli, a number of bilingual studies exploring the activation of phonology have likewise suggested that languages are accessed in a nonselective manner (Dijkstra, Grainger, & van Heuven, 1999). Briefly, those studies have shown that interlingual homographs and *cognates* (words with identical spelling and meaning across languages, for example *HOSPITAL* in Spanish and English) enjoy processing facilitation, in that they are identified faster than matched controls on account of sharing lexical and orthographic representations across languages. Dijkstra et al. (1999) employed a progressive *demasking task* to present Dutch-English bilinguals with Dutch and English stimuli similar in terms of orthography, semantics, and phonology. In a progressive demasking task the participant is shown a target word and a mask which are alternating and is instructed to react as soon as s/he can identify a word. During alternations, the time of the presentation of the mask gradually decreases and the time of the presentation of the target word increases. Participants in Dijkstra et al.'s (1999) study reacted faster to stimuli with orthographic and semantic overlap, whereas they took longer to identify those targets which shared phonology. According to Dijkstra et al. (1999), this *phonological inhibition effect* is caused by the simultaneous activation of two distinct phonological representations which compete at the lexical level, thus incurring delayed identification of the item in the target language. This effect was further replicated in a second experiment, which employed the lexical decision task and asked the Dutch-English bilinguals to decide if the target shown on the screen was an English word or not. Similar to the results from Experiment 1, RTs to items with similar orthography and meaning were facilitated and RTs to stimuli with phonological overlap produced inhibition (i.e., longer RTs). Overall, this brief review of bilingual lexical access studies suggests that both languages are activated simultaneously when bilinguals process individual words. Is the same true for the processing of language at the sentence level, when words are embedded in context? How will bilinguals process code-switched sentences? These questions are addressed in the next section.

Multiple Language Activation in Code-Switched Bilingual Sentence Processing and Connected Speech

In order to study lexical access during the comprehension of code-switched and monolingual sentences, Soares and Grosjean (1984) used a *phoneme-triggered lexical decision task* (Blank, 1980). While hearing sentences presented binaurally, bilinguals listened for a prespecified phoneme (e.g., /g/ for *guerra* in Spanish or /w/ for *war* in English) and decided whether the target containing the phoneme was a word or a nonword. Results showed that bilinguals were faster to make lexical decisions to targets in the monolingual sentences (e.g., [1a] *It is difficult to admit that a WAR sometimes brings more profits than losses*) than in the code-switched sentences (e.g., [1b] *It is difficult to admit that a GUERRA sometimes brings more profits than losses*). Thus, like previous studies (e.g., Macnamara & Kushnir, 1971; see also Kolers, 1966), Soares and Grosjean's findings suggested that word retrieval in mixed-language, as opposed to monolingual sentences, required an extra amount of time. These differences in retrieval have been taken to support the idea of a general *input mechanism* that determines which of the bilingual's two mental lexicons will be *on* or *off* during language processing at a given time (Heredia & Altarriba, 2001; Macnamara & Kushnir, 1971). Accordingly, during the comprehension of a monolingual English sentence, the input switch selects the English linguistic system and the Spanish linguistic system is deselected. Exposure to a language-mixed sentence would require the temporary deactivation of the English linguistic system to properly identify and process the Spanish word.

Other research, however, has focused on identifying some of the factors influencing the comprehension of mixed-language sentences. Li (1996) used a *cue-shadowing task* (Bates & Liu, 1996; Liu, Bates, Powell, & Wulfeck, 1997) and a gating task (Grosjean, 1988) to investigate two important factors of interest. The first factor was a phonological variable concerned with the permissible initial sound sequences in Chinese and English. The English language allows both consonant-consonant (CC) and consonant-vowel (CV) clusters at the beginning of a word. Chinese, on the other hand, allows CVs but lacks CCs. This manipulation examined the extent to which CC clusters, which were marked as belonging to English, would be identified faster than CVs, which were shared by both languages. That is, CC configurations would entail lexical search in only the English lexicon, whereas CV clusters would engage a lexical search in both languages. The second factor was prior context (biased vs. nonbiased preceding contextual information). In the cue-shadowing task, Chinese-English bilinguals listened to Chinese sentences and their task was to shadow or name the embedded English word within the sentence. Participants in this task *were told about the predesignated point [where the target word would appear] before each block of testing* (Li, 1996, p. 770). Overall, unlike the predictions, results revealed that bilinguals were faster to name English code-switched targets with initial CV than CC clusters. These findings suggested to Heredia and Stewart (2002) that the language preceding the code-switched targets determined which phonotactic configuration would be most highly activated (cf. Grosjean, 1988). Moreover, the results also indicate that during the course of sentence processing, information that overlapped across the bilingual's two languages had priority and was activated simultaneously during lexical search. In relation to the second factor of interest, context failed to interact with phonotactics. The general trend was that the critical targets in the biased contextual condition were recognized faster than targets in the nonbiased contextual condition (see also Li & Yip, 1998; cf. Altarriba, Kroll, Sholl, & Rayner, 1996; Chaps. 5 and 8). Unlike the study by Soares and Grosjean (1984), no monolingual sentences were used in this study. Thus, it is difficult to determine differences in lexical access between the code-switched and monolingual sentences.

In another study, Hernández, Bates, and Ávila (1996) set out to explore cross-language priming using a *cross-modal naming task* (CMN; see also Hernández, 2002). In the CMN task, participants name (read out loud) words presented visually on the computer screen while listening to the sentences presented auditorily. Findings from bilingual cross-language priming studies show that naming a word in one language (e.g., *war* in English) is faster when preceded by a related word of a second language (e.g., *paz* Spanish for "peace"), than by an unrelated critical word (e.g., *boca* Spanish for "mouth"). In one language condition, for example, the prime may be in the first language (L1) and the target in the second language (L2) or vice versa (e.g., Fox, 1996; Gollan, Forster, & Frost, 1997; Keatley & De Gelder, 1992; Keatley, Spinks, & De Gelder, 1994). The same logic applies to within-language priming, with the exception that the prime (e.g., *peace*) and the target are both in English (e.g., *war*) or both in Spanish.

Hernández et al. (1996) had bilinguals listen to sentences during which, at a predetermined location, the sentence stopped and a visual related or unrelated target word appeared in the middle of a computer screen. In the within-language condition, sentences were in English with the critical target in English (E-E), such as in sentence (1a) above or all in Spanish (S-S). In the English-Spanish cross-language condition (E-S), an English sentence contained a Spanish target (e.g., sentence 1b), or the Spanish sentence contained an English target (S-E). The beginning of each sentence was always presented auditorily and the target to be named was always presented visually, either immediately or with a delay. All language conditions were either blocked or randomly mixed. In general, cross-language priming was obtained, but only when language conditions were blocked or naming was delayed. When language conditions were mixed, except in the delayed condition, no priming was observed. However, within-language priming was observed for both monolingual conditions, regardless of the experimental condition. This pattern of results led Hernández et al. to conclude that *cross-language priming appears only when participants know what language to expect*, *when they have ample time to generate a response or both* (p. 860). In other words, switching from one language to the other takes time and access to an L2 cannot occur unless the bilingual is in some type of a *bilingual mode*. However, Hernández et al.'s results may have been due to the high predictability of their stimuli, thus forcing participants to develop strategic anticipatory processes.

The Present Study

The present study further investigates bilingual sentence processing at the spoken and connected discourse levels. Specifically, this set of experiments has two important aims. First, a general pattern amongst the studies reviewed here utilized sentences in which the code-switched target is always embedded within the sentence (e.g., sentences 1b–c above). Although it could be argued that such practice truly reflects the manner in which bilinguals communicate, such methodology is problematic because it may encourage bilinguals to simply respond to the language switch of the target word (e.g., Heredia & Stewart, 2002). Because of the distinctiveness of the code-switched target, as the sentence unfolds, participants simply wait for the language cue to respond. Thus, in the cue-shadowing technique (e.g., Li, 1996), for example, it is not clear if the shadowing of the code-switched word is performed with or without the activation of meaning (Bates & Liu, 1996). The present study attempts to overcome this potential drawback by employing the CMN (e.g., Love, Maas, & Swinney, 2003; Heredia & Blumentritt, 2002; Heredia & Stewart, 2002; Stewart & Heredia, 2002; cf. Hernández et al., 1996; Hernández, 2002). An important feature of the CMN is that during sentence presentation, the flow of the sentence is never interrupted (cf. Hernández, 2002; Hernández et al., 1996), thus making it difficult for participants to engage in strategic processing (see for example, Bates & Liu, 1996; Li, 1996). For this reason, bilinguals in the present study are presented with monolingual sentences entirely in English (e.g., 2a) or entirely in Spanish. The participants' objective in this task is to name a target in Spanish or English that is either related or unrelated to the critical prime.

(2a) *It is difficult to admit that a WAR$_{[*1]}$ sometimes brings more profits than losses*
  translation: *Es difícil reconocer que una GUERRA$_{[*1]}$ trae más ganancias que pérdidas.*

(2b) *Soldiers are trained for combat and WAR$_{[*1]}$ and that is why so much is invested in them.*
  translation: *Los soldados se entrenan para el combate y la GUERRA$_{[*1]}$ y por eso se invierte en ellos.*

In the E-S cross-language condition, for instance, as bilinguals listen to sentences (2a–b), the Spanish related *paz* ("peace") or unrelated target *boca* ("mouth") is presented at the offset (depicted by the subscript [*1]) of the critical prime *war*. The general idea here is to obtain a measurement of lexical access by computing a priming effect between the related and unrelated targets. Priming in this case is taken as an index of lexical access. Indeed, this task may be suitable for examining the extent to which the bilingual's L1 lexicon remains active or inactive during L2 sentence processing. Moreover, this technique also overcomes the problem of the *grammaticality of code-switching*, as sentences are presented in one language. Additionally, care should be taken in constructing code-switched sentences because of the possibility of constructing unnatural linguistic groupings. Accordingly, during code-switching, the natural tendency is not to break up linguistic categories such as the noun phrase *the traffic* into *the tráfico* or the infinitival phrase *to drive* into *to manejar* (see also Lederberg & Morales, 1985; Chap. 4, this volume for similar methodological issues). Inspection of the stimuli utilized in some of the studies reviewed here (e.g., Hernández et al., 1996) reveals inconsistencies in relation to the position of the code-switched word and the preceding linguistic category.

The second purpose of the present study was to systematically manipulate context effects to examine specific assumptions of the Revised Hierarchical Model of bilingual memory representation (Kroll & Stewart, 1994). Briefly, this model is based on the assumption that the bilingual's linguistic system is represented at the lexical and conceptual levels. At the lexical level, bilinguals represent their languages in separate, but bi-directionally interconnected lexicons. The link from the L2 to L1 lexicon is stronger than the L1 to L2 link, because it reflects the way the L2 was learned. During L2 acquisition, bilinguals learn to associate every L2 word with its L1 equivalent (e.g., learn *hous*e, associate it with *casa*), thus forming a lexical-level association that remains active and strong (Kroll & Stewart, 1994). At the conceptual level, both languages share one conceptual general store. Meaning or semantic information is represented at this level. Moreover, links from the L1 and L2 lexicons to the conceptual store are bi-directional and differ in strength. The conceptual link from L1 is stronger than the link from L2 to the conceptual store. This difference in strength reflects the fact that L1 is the native language, and bilinguals are more familiar with word meanings in their L1. Although it is theoretically possible that the link from L2 to the conceptual store may develop strong connections (e.g., Altarriba & Mathis, 1997), Kroll and Stewart argue that this link remains relatively weak, even for bilinguals with high L2 proficiency levels (but see Heredia, 1995; Heredia, 1997; Heredia & Altarriba, 2001; Heredia & Brown, 2003).

This model generates two important predictions: (a) Retrieval from L1 to L2 is conceptually mediated and affected by semantic and conceptual factors. Before accessing L2, L1 is more likely to activate the conceptual store, because of its strong connection to it. Thus, activation of the conceptual store should be increased with the manipulation of variables known to evoke semantic/conceptual processing. And, (b) retrieval from L2 to L1 is less likely to be affected by semantic/conceptual factors because it can be performed at the lexical level without recourse to meaning. Therefore, any increase in semantic/conceptual processing should not affect lexical access from L2 to L1.

The model's predictions have been supported empirically in the word translation literature (Kroll & Stewart, 1994; see also, Cheung & Chen, 1998; Sholl, Sankaranarayanan, & Kroll, 1995) and the priming literature. For example, this literature reports asymmetrical cross-language priming effects. Results show that naming an L2 target word is faster, but only if preceded by a related rather than an unrelated L1 prime (e.g., Fox, 1996; Jiang & Forster, 2001, Experiment 1; Keatley et al., 1994; but see Keatley & De Gelder, 1992). In contrast, naming an L1 target is no different than naming a related or unrelated L2 prime (Fox, 1996; Keatley et al., 1994; see also; Gollan et al., 1997). That is, cross-language priming is obtained only if the prime is in L1 and the target is in L2. Indeed, consistent with the Revised Hierarchical Model, the results suggest that accessing the L2 from the L1 lexicon is conceptual because it is achieved via the conceptual store that is the locus of the semantic priming effect (Keatley et al., 1994, p. 77). In contrast, accessing the L1 from the L2 lexicon takes place only at the lexical level, thus producing no semantic priming. This prediction would be more likely to be true for bilinguals whose L2 is not the dominant language (see for example, Heredia, 1997; Heredia & Altarriba, 2001; see also Hernández, 2002).

Evidence for this model is not unequivocal. Some studies have suggested that retrieval from both language directions may be sensitive to meaning-based processing (e.g., De Groot, 1992; De Groot, Dannenburg, & Van Hell, 1994; Heredia, 1995, 1997; La Heij, Hooglander, Kerling, & Van der Velder, 1996; see also Altarriba & Mathis, 1997; Jiang & Forster, 2001). In addition, results at the sentential level suggest that, depending on whether language conditions are presented in blocked or mixed designs, both cross-language conditions exhibit or fail to show priming effects. Hernández et al. (1996) found that when language conditions were randomly mixed as to prevent participants from generating strategies or predicting the language of presentation, both cross-language conditions failed to show priming. When language conditions were blocked, both L1 to L2 and L2 to L1 conditions showed comparable priming effects. This was generally true for blocked and delayed conditions, with the exception of one experiment in which language presentation was mixed but targets were degraded. In this case, L2 to L1 conditions produced significant priming, whereas L1 to L2 conditions did not. Other similar experiments (e.g., Hernández, 2002) using sentences and the priming paradigm show that L2 to L1 language conditions reveal larger priming effects than L1 to L2 conditions. In fact, prior context seemed to increase the priming effect for the L2 to L1 condition and had no effect on the L1 to L2 cross-language conditions (cf. Heredia, 1995, 1997). These findings, as can be seen, are the opposite of what the Revised Hierarchical Model would predict. Clearly, more empirical work is required to determine the usefulness of this model to explain bilingual semantic memory and how the model could be applied to sentence processing.

Previous studies addressing this model have operated under the assumptions that conceptual and semantic information can be obtained by the manipulation of concreteness (De Groot, 1992; De Groot et al., 1994; Heredia, 1995, 1997) or category effects (Kroll & Stewart, 1994) using the isolated word (word pair) or the picture as the experimental unit. The present investigation goes a step further and systematically manipulates the effects of previous sentential context, a variable known to facilitate lexical access in monolinguals (e.g., Herron & Bates, 1997; Marslen-Wilson, 1987; Tabossi, 1988, 1996) and bilinguals (e.g., Altarriba et al., 1996; Li, 1996; Li & Yip, 1998; see also Heredia et al., this volume) during the online comprehension of spoken sentences.

Research Questions

What are the effects of context on bilingual lexical access? Does the preceding context have differential effects on how bilinguals access information from their two lexicons? Is access from the L1 to the L2 bilingual lexicon more likely to be affected by contextual effects than access from the L2 to the L1 bilingual lexicon? In Experiment 1, Spanish-English bilinguals listened to Spanish translations of a sentence (2a), where the preceding context provides no biasing information towards the meaning of the critical prime *guerra*. In Experiment 2, participants listened to Spanish translations of sentence (2b), where the preceding context provides relevant and biasing information about the meaning of the critical prime. As can be seen from sentence (2b), the Spanish *soldados* ("soldiers") and *combate* ("combat") reinforce the meaning of the critical prime *guerra*. In both experiments, sentences were delivered aurally without disruption, and at the offset of the critical prime participants named a related (e.g., *peace*) or unrelated (*road*) English target. Probing was done at prime offset in order to inspect L2 word activation immediately after the processing of the L1 prime. Experiments 1 and 2 represent the S-E cross-language condition or the L1 to L2 condition. In addition to the cross-language condition, a within-language manipulation was included in which participants named Spanish-related (e.g., *paz*) or unrelated (e.g., *boca*) targets. This condition was included to serve as a comparison and a baseline for the bilingual condition, and to examine differences or similarities in lexical access between monolingual and cross-language conditions. Is it possible to retrieve information from L2 as the bilingual speaker processes sentences in L1? Because the critical prime is in L1 and the target is in L2, Experiment 1 should exhibit the cross-language priming effect. In this case, naming related targets should be faster than naming unrelated targets. That is, L2 access should be possible, as the bilingual speaker processes sentences in the L1. Experiment 2 should replicate the results of Experiment 1. However, if L1 to L2 is conceptually mediated and sensitive to semantic/conceptual factors, as predicted by the Revised Hierarchical Model, the presence of prior contextual information should facilitate cross-language lexical access. In this case, cross-language priming should increase from Experiment 1 to Experiment 2.

Experiments 3 and 4 represented the E-S cross-language conditions. English sentences such as in (2a) were used for these experiments. The critical targets for this experiment were in Spanish, to represent the E-S cross-language condition or the L2 to L1, and in English, to represent the E-E or the within-language condition. Are L2 to L1 language directions sensitive to conceptual/semantic factors? A strong version of the Revised Hierarchical Model predicts no cross-language priming from L2 to L1. Therefore, no cross-language priming should be observed from the contextually-unbiased (Experiment 3) to the contextually-biased condition (Experiment 4). If there is cross-language priming, it should remain relatively unaffected by the preceding contextual information of Experiment 4. Alternatively, if L2–L1 language directions are sensitive to the semantic information provided by the contextual information of Experiment 4, an increase in cross-language priming should be observed from Experiment 3 to Experiment 4.

Method

The four experiments reported here used a standard procedure as described in this section. Where an experiment departs from this procedure, exact changes are specified within the description of the experiment.

Materials

The stimuli consisted of 129 Spanish sentences, 69 of which were experimental and 60 sentences served as fillers. Mean word length for the experimental sentences was five letters and word average per sentence was 25. Filler sentences were similar in length to the experimental sentences. Half of the fillers were paired with an unrelated Spanish word and half with an unrelated English word. Four experimental lists were constructed. For the S-S monolingual condition, 17 sentences were paired with a Spanish target word related to a critical prime (e.g., *guerra-paz*), and 17 sentences were paired with a Spanish control word that was unrelated to the critical prime (e.g., *guerra-boca*). Control words were matched in frequency and length to the related targets according to Julliand and Chang-Rodríguez (1964) Spanish word frequency counts. For the S-E cross-language condition, 17 sentences were paired with an English target word related to the critical prime (*guerra-peace*), and 17 sentences were paired with an English control word unrelated to the critical prime (*guerra-road*). Control words were matched in frequency and length to the related words according to Francis and Kučera's (1982) frequency counts. An additional sentence appeared on every list. In one list, this sentence was paired with a Spanish-related target. In another list, this sentence appeared with an English-related target, and so on.

The procedure for creating the experimental sentences was as follows. Sixty-nine nouns and their associates (e.g., *war-peace*) were obtained from Nelson, McEvoy, and Schreiber (1998) English-free association norms. These words were then translated into Spanish (e.g., *guerra-paz*). For every word pair, two Spanish sentences were written. One sentence (for Experiment 1) was written in such a way that the preceding context provided no information towards the critical prime (see sentence 2a, above). A second sentence (for Experiment 2) was written in such a way that the preceding context biased the meaning of the critical prime (see sentence (2b), above).

To assure that the preceding context of the sentences for Experiment 1 was not related or biasing the critical prime, a pretest study was performed. Thirty-two Spanish-English bilinguals were given the stimuli as sentence fragments with the critical word in uppercase (e.g., *it is difficult to admit that a WAR*) and asked to rate, on a 1–7 scale (1 = Not Biased and 7 = Very Biased), the extent to which the preceding context biased the meaning of the word in uppercase. The mean rating for the 69 experimental sentences was 3.2 (SD = .73). The same was done for the contextually-biased sentences for Experiment 2 (sentence 2b above). The mean rating for these sentences was 5.5 (SD = .84). A comparison between contextually-unbiased vs. contextually-biased sentences showed that the mean ratings for the contextually-biased ones were significantly higher, $t(31) = 11.57$, $p < .05$. Finally, care was taken to ensure that information after the critical prime was not related to the target or to the context following the critical prime.

Four lists were required to counterbalance each sentence. Each target word was assigned to one of the four lists using a Latin square design. All 129 sentences were combined in a pseudo-random order, with the only constraint that no more than three items from a given experimental condition occurred consecutively. Additionally, ten sentences (half related and half unrelated) served as practice trials. Fourteen multiple choice comprehension questions were presented throughout each experimental list that asked participants details about a preceding sentence they had just heard. The relatedness proportion (34 related trials out of 129 unrelated trials) was .26 (see Altarriba & Basnight-Brown, 2007; Garcia et al., 2015; Neely, Keefe, & Ross, 1989, for a discussion of the importance of these effects).

Stimuli were recorded by a female native speaker of Spanish. Sentences were directly read into a Sony TCD-8 Digital Audio Tape Corder. The recordings were then entered into a G3 Macintosh using Macromedia SoundEdit 16 Version 2. A sampling rate of 44.1 kHZ with a 16-bit format was used for digitizing. For every wave sound, the offset of a critical prime was located as accurately as possible by using waveforms and auditory feedback. A cue marker was placed at prime offset to indicate to the computer the point at which the visual target was to be presented during sentence presentation. For all experimental sentences, the critical prime appeared in the middle of the sentence. The filler sentences were created the same way, except that cue markers were placed at random points throughout the sentence.

Procedure

Upon arrival, participants read the experimental instructions from the computer screen. They were instructed to listen carefully to the sentences being presented over headphones, understand them, and to pronounce as fast and as accurately as possible a visually presented word, which would appear in the middle of a computer screen. Their responses were recorded and examined for pronunciation errors.

Sentences were delivered uninterrupted at a normal speaking rate. At the offset of the critical prime, a visual target word appeared in front of the computer screen for 300 milliseconds (ms). This short target presentation is standard in the CMN and it controls for any possibility of backward priming (Love & Swinney, 1996; Prather & Swinney, 1988; but see Glucksberg et al., 1986). Response time was measured from the onset of the visually presented target until the participants responded or after a 2300 ms time response window. Sentences were presented over headphones (Optimus Pro-50MX). The experiment was controlled by PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) and participants' responses to the visually presented targets were controlled by the CMU button box (Cohen et al., 1993) connected to a Star Max 3000 Motorola Macintosh compatible computer. The stimuli were played through a set of Apple speakers. After the experiment, each subject completed a Language History Questionnaire.

Experiment 1: Contextually-Unbiased Spanish Sentences

Participants

Forty-five Spanish-English bilinguals participated in the experiment. Participants were students from Texas A&M International University who received course credit for their participation. All bilinguals reported Spanish as their L1 and English as their L2. Participants reported receiving their formal education in English. Analyses were conducted on the participant's language self-ratings on language usage. Self-ratings were based on a 1–7 scale (1 = Not Fluent 7 = Very Fluent). The mean age of the group was 27.6 years. The mean years spent in the United States were 23.4. Mean self-ratings showed that they used Spanish ($M=5.2$) and English ($M=5.7$) equally often. Their speaking ability in Spanish ($M=5.9$) and English ($M=6.4$) was comparable. However, their ability to read English was rated higher ($M=6.4$) than their ability to read Spanish ($M=5.3$), $t(44)=3.8$, $p<.05$. Likewise, their writing ability was rated higher for English ($M=6.5$) than Spanish ($M=5.1$), $t(44)=4.3$ $p<.05$. Understanding English ($M=6.7$) was also rated higher than Spanish ($M=6.1$), $t(44)=2.8$ $p<.05$. It is important to note that bilinguals in the South Texas are known for their ability to mix their two languages simultaneously during their everyday communication.

Results and Discussion

In this and all subsequent experiments, naming responses in milliseconds (ms) above or below 3 SD*s* were treated as outliers. This procedure affected 2.2 % of the total data. Analyses were performed on error rates, and on response latencies (RTs) for correct responses.

Error rates. Pronunciation errors or failure to respond to the visually presented target word were subjected to a 2 (Relatedness: related vs. unrelated) × 2 (Language: Spanish vs. English) within-subjects analysis of variance ANOVA. This analysis was performed for both participants (F_1) and items (F_2). The main effect for relatedness was significant by subjects, $F_1(1,44) = 4.12$, MSE = .0013, $p < .01$, but not by items, $F_2(1,68) = 1.84$, MSE = .0023, $p = .18$. Bilinguals made more mistakes naming unrelated controls ($M = 3.0$ %), than related words ($M = 1.95$ %). The percentage of errors was higher in naming Spanish ($M = 2.9$ %), than naming English target words ($M = 2.0$ %); however, this trend was not statistically reliable by subjects nor by items (all *F*s < 1). The interaction was significant by subjects, $F_1(1,44) = 4.55$, MSE = .0022, $p < .05$, and marginal by items, $F_2(1,68) = 3.78$, MSE = .0019, $p = .06$. The Least Significant Difference (LSD = 1.7 %) multiple comparison (Bruning & Kintz, 1987; Cohen & Cohen, 1983) was calculated to analyze the significant interaction. In all subsequent analyses, the alpha level is set at .05. The LSD revealed that bilinguals made more errors naming Spanish-unrelated ($M = 4.1$ %), than Spanish-related targets ($M = 1.5$ %). Percentage of errors in naming English unrelated ($M = 1.8$ %) vs. related English targets ($M = 2.2$ %) was not statistically significant.

Response latencies. A 2 (Relatedness: related vs. unrelated) × 2 (Language: Spanish vs. English) within-subjects ANOVA was performed. There were main effects for relatedness, $F_1(1, 44) = 42.0$, MSE = 1130.25, $p < .01$; $F_2(1, 68) = 15.35$, MSE = 3718.33, $p < .01$ and language, $F_1(1, 44) = 15.40$, MSE = 5426.91; $F_2(1, 68) = 19.01$, MSE = 4928.50, $p < .01$. Participants were 33 ms faster naming related ($M = 688$ ms, SD = 111) than unrelated controls ($M = 721$ ms, SD = 125). More interesting was the finding that naming an English target ($M = 683$, SD = 103) was 43 ms faster than naming a Spanish target ($M = 726$ ms, SD = 130). This finding, showing that bilinguals were faster in naming English than Spanish targets, is a surprising one considering that sentence presentation was in Spanish. This finding is counter to the *base language effect* (Grosjean, 1997, p. 241) that suggests that the language being spoken has a strong effect on which language will be favored during lexical access. Thus, during the processing of Spanish sentences, naming a Spanish word would be faster than naming an English word. This pattern of results supports the intuition of many bilinguals reporting that when they use their L1 (e.g., Spanish), sometimes they find themselves resorting to their L2 (e.g., English) to communicate.

The interaction between relatedness and language was significant by subjects, $F_1(1, 44) = 13.00$, MSE = 1302.00, $p < .01$, but marginally significant by items, $F_2(1, 68) = 3.503$, MSE = 5408.81, $p = .07$. Simple effects (LSD = 12.80) for the subject means showed that the 52 ms priming effect for the Spanish condition was reliable. Likewise, the 13 ms priming effect for the S-E condition was significant (see Table 6.1).

To summarize, both monolingual and cross-language conditions produced significant priming. Although the priming effect was greater for the monolingual condition (i.e., the within-language), this effect is not surprising given that the spoken sentences were in Spanish. The cross-language priming effect, although reliable only in the analyses by subjects, contrasts with the results reported by Hernández et al. (1996) who reported no cross-language priming under conditions in which participants were unable to predict the language of presentation (their mixed-language

Table 6.1 Mean reaction times and standard deviations (SD) in milliseconds for Spanish-English bilinguals as a function of language and relatedness in Experiment 1

	Relatedness		
Language	Related	Unrelated	Priming
Spanish	700 (119)	752 (137)	52*
English	677 (102)	690 (104)	13*

*$p<.05$

condition). Another important finding here is the main effect for language, which showed that naming English words was actually faster than naming Spanish visual targets. This difference could be due to the fact that bilinguals in the present study reported higher ratings in their reading and writing English ability. Moreover, the results of this experiment are consistent with the predicted priming patterns of the Revised Hierarchical Model. As predicted, we obtained cross-language priming (L1–L2), even when the critical prime was presented aurally and embedded within a sentence. In the following experiment, bilinguals listened to sentences in which the preceding context is biased towards the critical prime. If it is true that L1 to L2 bilingual direction is sensitive to contextual/semantic effects, then manipulation of the preceding context should enhance activation of the L1 concept and strengthen the activation of the conceptual links between the L1 and L2 concepts. In this case, cross-language priming should increase significantly compared to Experiment 1.

Experiment 2: Contextually Biased Spanish Sentences

Participants

Thirty-nine Spanish-English bilinguals participated in this experiment. All participants were students from the University of California, San Diego who received course credit for their participation or were paid $6.00 per hour. The mean age of the group was 21.7, and the mean number of years in the United States was 19. All bilinguals participating in this experiment reported Spanish as their L1 and English as their L2. The majority of the participants reported using Spanish with their family and English with their friends. English was the main language for their education. Analyses were conducted on the participants' responses to a language questionnaire. Mean self-ratings show that participants used English ($M=6.2$) more frequently than Spanish ($M=4.1$), $t(38)=6.7$, $p<.05$). Their speaking ability was greater for English ($M=6.6$) than Spanish ($M=5.8$), $t(38)=4.0$, $p<.05$. Similarly, their ability to read ($M=6.6$) and write English ($M=6.5$) was rated higher than their ability to read ($M=5.7$) and write ($M=5.0$) Spanish, $t(38)=3.60$, $p<.05$, and $t(38)=5.49$, $p<.05$, respectively. Means for understanding English ($M=6.6$) and Spanish ($M=6.3$) were comparable.

Materials and Procedure

Materials and procedures were the same as in Experiment 1, except that the sentences were constructed in such a way that the preceding context was biased towards the meaning of the critical prime, as shown by sentence (2a) above. Mean word length for the experimental sentences was five letters, and word average per sentence was 22.

Results and Discussion

Error rates. The 3 SD cutoff procedure for the exclusion of naming responses constituted 1.5 % of all data. Pronunciation errors or failure to respond to the visually presented target words were subjected to a 2 (Relatedness) × 2 (Language) within-subjects ANOVA. The main effect for relatedness was significant, by subjects, $F_1(1,38) = 11.44$, MSE = .0022, $p < .01$, and by items, $F_2(1,68) = 10.19$, MSE = .0066, $p < .001$). Bilinguals made more errors naming unrelated ($M = 5.2$ %) than related target words ($M = 2.6$ %). The main effect for language was not statistically reliable by subjects nor by items (all Fs < 1). Errors naming English ($M = 3.9$ %) were comparable to naming Spanish target words ($M = 3.9$ %). The interaction was not significant by subjects nor by items (all Fs < 1), suggesting that naming English ($M = 5.1$ %) and Spanish ($M = 5.3$ %) unrelated controls produced comparable naming errors. Likewise, naming English- ($M = 2.8$ %) and Spanish-related targets ($M = 2.5$ %) showed similar error rates.

Response latencies. A 2 (Relatedness) × 2 (Language) within-subjects ANOVA showed a main effect for relatedness, $F_1(1, 38) = 31.41$, MSE = 1054.89, $p < .01$; $F_2(1, 68) = 14.08$, MSE = 4914.56, $p < .01$, and language, $F_1(1, 38) = 10.52$, MSE = 3565.30, $p < .01$; $F_2(1, 68) = 16.86$, MSE = 4806.10, $p < .01$. Participants were 29 ms faster naming related ($M = 628$ ms, SD = 108) than unrelated controls ($M = 657$, 114 ms). Like Experiment 1, naming an English target ($M = 627$ ms, SD = 102) was 31 ms faster than naming a Spanish target ($M = 658$ ms, SD = 119). This finding is important because it replicates the results of Experiment 1 that involved a Spanish-English bilingual population from a geographical area where English is the main language of communication and general interaction.

The interaction between relatedness and language was significant by subjects, $F_1(1, 38) = 9.93$, MSE = 919.03, $p < .01$, but marginally significant by items, F_2 (1, 68) = 3.80, MSE = 4370.47, $p = .06$. Multiple comparisons (LSD = 11.65) for subject means indicate that the difference between the related and unrelated Spanish targets was significant (see Table 6.2). This indicates a significant priming effect for the monolingual condition. Similarly, as can be seen from Table 6.2, the difference between the related and unrelated English target words was significant, thus exhibiting a reliable priming effect. These results follow the same patterns as in Experiment 1. As predicted by the Revised Hierarchical Model, the S-E (L1–L2) conditions exhibited the priming effect.

Table 6.2 Mean reaction times and standard deviations (SD) in milliseconds for Spanish-English bilinguals as a function of language and relatedness in Experiment 2

	Relatedness		
Language	Related	Unrelated	Priming
Spanish	636 (113)	680 (122)	44*
English	620 (105)	634 (101)	14*

*$p < .05$

To further explore the effects of context and bilingual lexical access, an additional 2 (Context: biased vs. unbiased) × 2 (Relatedness: related vs. unrelated) × 2 (Language: Spanish vs. English) was performed. This comparison involved Experiment 1 vs. 2. The main effect for context was significant by subjects, F_1 (1, 82) = 7.124, MSE = 4,5028.75, $p < .01$ and by items, F_2(1, 136) = 132.40, MSE = 4963, $p < .01$. This main effect suggests that bilinguals were 62 ms faster in naming words under contextually-biased (*M* = 643 ms) than under contextually-unbiased conditions (*M* = 705 ms). There was also a language main effect, both by subjects, F_1 (1, 82) = 25.9, MSE = 4554.67, $p < .01$, and by items, F_2(1, 136) = 35.91, MSE = 4867. 30, $p < .01$. English targets (*M* = 657 ms, SD = 106) were 38 ms faster than Spanish targets (*M* = 695 ms, SD = 129). The two-way interaction of relatedness vs. language was also significant, by subjects, F_1 (1, 82) = 22.91, MSE = 1125.56, $p < .01$, and by items, F_2(1, 136) = 7.27, MSE = 4, 889.64, $p < .01$. This two-way interaction qualifies the interactions in Experiments 1 and 2. Simple effects (LSD = 6.12) show a significant 14 ms priming effect between English-related (*M* = 650 ms, SD = 107) and English-unrelated targets (*M* = 664 ms, SD = 106). Likewise, the 49 ms priming effect for Spanish-related (*M* = 670 ms, SD = 118) and unrelated targets (*M* = 719 ms, SD = 135) was reliable. Other two- or three-way interactions were not reliable (all *F*s <1).

Taken together, the results of Experiments 1 and 2 indicate that context had an additive effect on the retrieval of both Spanish and English words. This additive effect may explain why bilinguals were generally faster in Experiment 2, as compared to Experiment 1, when naming Spanish and English targets. Naming a target word in Spanish or English is faster in conditions in which prior context biases the Spanish prime. These findings replicate other studies that have manipulated context and bilingual lexical access (e.g., Li, 1996; Li & Yip, 1998). However, previous context does not increase the priming effect. Especially for the S-E condition, cross-language priming remained constant between Experiment 1 (13 ms), and Experiment 2 (14 ms). Although the results of Experiment 1 and 2 are consistent with the hypothesis that L1 to L2 produces significant priming, the addition of context failed to increase the priming effect. These results suggest that L1 to L2 may not be as sensitive to semantic effects as predicted by the Revised Hierarchical Model.

Moreover, another important finding was that bilinguals in the present study were actually faster in naming English than Spanish targets. This finding is remarkable considering that the sentences were all in Spanish, and from two separate bilingual populations. This issue is further elaborated in the discussion. The next experiment explores E-E and E-S lexical access. In Experiment 3, bilinguals listened to English sentences and named English (the E-E condition) or Spanish (E-S cross-language condition) targets. The purpose of this experiment was to further investigate the extent to which E-S produces the priming effect, and whether or not it is affected by the addition of contextual information (Experiment 4).

Experiment 3: Contextually Unbiased English Sentences

Participants

Fifty Texas A&M International University Spanish-English bilinguals participated in this experiment. Students received course credit for their participation. Bilinguals in this study did not participate in the previous two experiments. All bilinguals reported Spanish as L1 and English as L2, with Spanish used as the family language and English as the language used with their friends. Mean age of the group was 23.2, and the mean number of years in the United States was 20.1. Mean self-ratings showed that participants used Spanish ($M=5.5$) and English ($M=5.5$) equally often, and their speaking ability in both languages was comparable ($M=5.9$ and $M=6.1$, respectively). However, their reading ability was higher for English ($M=6.3$) than for Spanish ($M=5.4$), $t(49)=3.28$, $p<.05$. Likewise, their writing ability was rated higher for English ($M=6.5$) than Spanish ($M=4.6$), $t(49)=5.40$, $p<.05$. Ratings for English understanding ($M=6.6$) were higher than for Spanish ($M=6.1$), $t(49)=3.25$ $p<.05$.

Materials and Procedure

Sentences and stimuli construction followed the same procedure as in Experiment 1. However, sentences for this Experiment were in English (see sentence 2a). Related and unrelated targets were the same as in Experiment 1. Mean word length for the experimental sentences was five letters, and word average per sentence was 21. Sentences were recorded by a native male speaker of English. To assure that the preceding context did not bias the meaning of the critical prime, 48 bilinguals were asked to rate the experimental sentences (1=Not Biased and 7=Very Biased). Participants rated both the contextually-unbiased (Experiment 3) and contextually-biased (Experiment 4) sentences. The mean rating for the contextually-unbiased sentences was 3.4 (SD=.85) and the mean rating for the contextually-biased sentences was 5.4 (SD=.89). Differences between these means were statistically significant, $t(46)=7.8$, $p<.05$).

Results and Discussion

Error rates. The 3 SD cutoff procedure for the exclusion of naming responses constituted 2.8 % of the data. Participants' pronunciation errors or failure to respond to the visually presented target words were subjected to a 2 (Relatedness) × 2 (Language) within-subjects ANOVA. Data from one participant were deleted because of a computer error. The main effect for relatedness was significant by subjects, $F_1(1,48)=21.35$, MSE, .0030, $p<.01$, and by items, $F_2(1,68)=7.32$, MSE=.0120, $p<.01$). Bilinguals made more errors naming unrelated controls ($M=8.2$ %) than related targets ($M=4.6$ %). Likewise, the main effect for language was reliable by subjects, $F_1(1,48)=21.61$, MSE, .0128, $p<.01$, and by items, $F_2(1,68)=27.93$, MSE=.0130, $p<.01$). Bilinguals experienced more errors naming Spanish ($M=10.2$ %) than English targets ($M=2.7$ %). The interaction was significant by subjects, $F_1(1, 48)=21.26$, MSE, .0037, $p<.01$, and by items, $F_2(1,68)=6.80$, MSE=.0139, $p<.01$). This interaction (LSD=2.1) shows that bilinguals had more difficulty naming Spanish ($M=6.4$ %) than English ($M=2.9$ %) related targets. Similarly, Spanish-unrelated words ($M=14.0$ %) produced more naming errors than English-unrelated targets ($M=2.5$ %). Clearly, bilinguals in this experiment experienced interference from English when naming Spanish targets.

Response latencies. A 2 (Relatedness)×2 (Language) ANOVA showed a main effect of relatedness, by subjects, $F_1(1, 48)=73.3$, MSE=863.35, $p<.01$, and by items, $F_2(1, 68)=29.8$, MSE=3182.29, $p<.01$. Participants were 36 ms faster to name related ($M=651$ ms, SD=99.0) than unrelated targets ($M=687$ ms, SD=110.4). The language main effect was also reliable by subjects, $F_1(1, 48)=56.24$, MSE=5219.24, $p<.01$, and by items, $F_2(1, 68)=100.2$, MSE=4260.44, $p<.01$. Participants were about 78 ms faster to name English targets ($M=630$ ms) than Spanish targets ($M=708$).

The interaction between relatedness and language was significant by subjects, $F_1(1, 48)=26.22$, MSE=1034.52, $p<.01$ and by items, F_2 (1, 68)=5.60, MSE=4744.78, $p<.05$. The analysis of simple effects (LSD=11.94) in Table 6.3 shows that bilinguals were faster to name Spanish-related than unrelated targets, thus showing a significant cross-language priming effect of 60 ms. Likewise, naming differences for the English-related and unrelated target were also statistically significant. The surprising result in this experiment was the robust priming effect for the E-S cross-language condition, and the smaller, but significant effect for the within-language condition. Unlike the predictions of the Revised Hierarchical Model, L2 to L1 language directions exhibited the priming effect. However, Experiment 4 is critical in determining the extent to which L2 to L1 language directions are indeed sensitive to context.

Table 6.3 Mean reaction times and standard deviations (SD) in milliseconds for Spanish-English bilinguals as a function of language and relatedness in Experiment 3

	Relatedness		
Language	Related	Unrelated	Priming
Spanish	678 (104)	738 (113)	60*
English	624 (87)	638 (81)	14*

*$p<.05$

Experiment 4: Contextually Biased English Sentences

Participants

Thirty-eight Texas A&M International University Spanish-English bilinguals participated in this experiment in exchange for course credit. All bilinguals reported Spanish as their L1 and English as their L2. Participants reported using Spanish as the family language, and Spanish and English with their friends. Mean age of the group was 23.0, and the mean number of years spent in the United States was 17.5. Mean self-ratings showed that participants used Spanish ($M=5.7$) and English ($M=5.4$) equally often. Their speaking ability for both Spanish ($M=6.1$) and English ($M=6.1$) was comparable. Similarly, their reading ability in Spanish ($M=5.7$) and English ($M=6.0$), and their ability to understand Spanish ($M=6.2$) and English ($M=6.0$) did not differ. However, their writing ability was rated higher for English ($M=6.0$) than for Spanish ($M=5.2$), $t(37)=2.13$, $p<.05$, and their understanding of both languages was rated similarly.

Materials and Procedure

For this experiment, the sentences were in English (see sentence 2b, above) and prepared using the same procedure as Experiments 1, 2, and 3. Mean word length for the experimental sentences was five letters, and word average per sentence was 21. Sentences were recorded by a native female speaker of English.

Results and Discussion

Error rates. The 3 SD cutoff procedure for the exclusion of naming responses constituted 1.5 % of the data. Participants' pronunciation errors or failure to respond to the visually presented targets words were subjected to a 2 (Relatedness)×2 (Language) within-subjects ANOVA. Data from four participants were excluded from the analysis because of computer errors. The main effect for relatedness was significant by subjects, $F_1(1, 33)=6.54$, MSE, .0025, $p<.05$, and marginal by items, $F_2(1, 68)=3.07$, MSE=.0084, $p=.08$. Bilinguals made more errors naming unrelated controls ($M=6.\ 4$ %) than related targets ($M=3.9$ %). The main effect for language was reliable by subjects, $F_1(1, 33)=6.24$, MSE=.0058, $p<.05$, and by items, $F_2(1, 68)=9.99$, MSE=.0106, $p<01$. Participants experienced more errors naming Spanish ($M=6.7$ %) than English targets ($M=3.4$ %). The interaction was significant by subjects, F_1 $(1,33)=5.27$, MSE=.0032, $p<.05$, and marginal by items, $F_2(1, 68)=3.08$, MSE=.0109, $p=.08$. This interaction (LSD=2.3 %) shows that bilinguals made similar mistakes naming Spanish ($M=4.4$ %) and English ($M=3.4$ %) related targets. However, Spanish-unrelated words ($M=8.8$ %) produced more naming errors than English-unrelated targets ($M=3.4$ %).

Response latencies. A 2 (Relatedness) × 2 (Language) within-subjects ANOVA showed a main effect for relatedness for both subjects, $F_1(1, 33) = 25.0$, MSE = 1664. 28, $p < .01$, and items, $F_2(1, 68) = 7.53$, MSE = 13,448.64, $p < .01$. Naming unrelated targets ($M = 708$ ms, SD = 155) was slower than naming related targets (M = *673* ms, SD = 143). There was a reliable language main effect by subjects, $F_1(1, 33) = 12.31$, MSE = 12,545.98, $p < .01$, and by items $F_2(1, 68) = 30.04$, MSE = 10,043.23, $p < .01$. In this case, English targets were named faster ($M = 657$, SD = 119) than Spanish targets ($M = 725$ ms, SD = 169).

The interaction between relatedness and language was marginally significant by subjects, $F_1(1, 33) = 3.66$, MSE = 1709.77, $p = .06$, but not by items, F_2 (1, 68) = 2.23, MSE = 9582.47, $p > .1$. Multiple planned comparisons (LSD = 17. 08) show a significant priming effect for both Spanish and English targets.

To determine the effects of context and bilingual lexical access for the L2 to L1 language direction, a 2 (Context: biased vs. unbiased) × 2 (Relatedness: related vs. unrelated) × 2 (Language: Spanish vs. English) mixed ANOVA was performed. This analysis combines Experiments 3 and 4. The main effect for context was not reliable by subjects, ($F_1 < 1$), however, it was significant by items, $F_2(1, 136) = 8.85$, MSE = 6339.21, $p < .01$. This main effect by items suggests that naming targets under contextually-biased conditions ($M = 690$, SD = 113) was actually 20 ms slower than contextually-unbiased conditions ($M = 670$ ms, SD = 78). These findings suggest that contextually-biased conditions actually inhibited both the naming of Spanish and English targets when listening to English sentences (cf. Altarriba et al., 1996; see also Heredia et al., this volume). The main effect for language was reliable both by subjects, $F_1(1, 181) = 54.36$, MSE = 8204.20, $p < .01$, and items, $F_2(1, 36) = 92.37$, MSE = 7757.31, $p < .01$. English targets ($M = 641$ ms, SD = 100) were named faster than Spanish targets ($M = 715$, SD = 138).

More importantly, relatedness interacted with language. The two-way interaction was reliable by subjects, $F_1(1, 81) = 23.97$, MSE = 1309.62. 28, $p < .01$, and by items, $F_2(1, 136) = 4.73$, MSE = 8045.85, $p < .05$. Multiple comparisons (LSD = 9.38 ms) show that Spanish-related targets ($M = 687$ ms, SD = 130) were responded to faster than unrelated targets ($M = 742$ ms, SD = 142). Thus the 55 ms priming effect is statistically reliable. Likewise, English-related targets ($M = 633$ ms, SD = 101) were faster than unrelated targets ($M = 650$ ms, SD = 100). The 17 ms priming effect is statistically significant. This interaction qualifies the priming effects for Experiments 3 and 4. In short, results from Experiments 3 and 4 showed priming for the E-S conditions. The lack of the 3-way interaction suggests that there was no increase of priming from Experiment 3 to Experiment 4.

Additional analyses were performed to explore differences between S-E and E-S conditions in the four experiments reported here. Data were analyzed on a 2 (Type of sentence: Spanish vs. English) × 2 (Relatedness: related vs. unrelated) × 2 (Language Target: Spanish vs. English) ANOVA. The three-way interaction did not reach significance (all Fs < 1). However, type of sentence (Spanish vs. English) interacted with language target both by subjects, $F_1(1, 92) = 5.19$, MSE = 5318.56, $p < .05$, and by items, $F_2(1, 136) = 13.10$, MSE = 4594.47, $p < .01$. This interaction indicates that when the sentence was in Spanish (Experiment 1), naming an English target took about 683 ms (SD = 103) on average. When the sentence was in English (Experiment 3), naming a Spanish target took about 708 ms (SD = 112). The difference of 25 ms is statistically significant (LSD = 17). Thus, in the language of the Revised Hierarchical Model, the L1 to L2 direction was actually faster than the L2 to L1 direction. When the sentence was in Spanish (Experiment 1) and the target was in Spanish, it took 726 ms (SD = 130) on average to respond. In Experiment 3, when the sentence was in English and the target was in English, it took 631 ms (SD = 84) on average to respond.

A similar analysis was performed to explore differences between Experiments 2 (sentence in Spanish) and Experiment 4 (sentence in English). The three-way interaction was not reliable by subjects nor by items (Fs < 1). The interaction of type of sentence by language was marginal by subjects, $F_1(1, 71) = 3.11$, MSE = 7739.41 $p = .08$, and significant by items, $F_2(1, 136) = 4.18$, MSE = 8030.14, $p < .05$. This interaction by items demonstrates that when the sentence was in Spanish, naming a target in English took 623 ms (SD = 63). When the sentence was in English and the target in Spanish, it took 722 ms (SD = 132). The 99 ms difference is statistically reliable (LSD = 17 ms). Again, unlike the predictions of the Revised Hierarchical Model, L1 to L2 is actually faster than L2 to L1. This pattern replicates across studies when comparing Experiment 1 vs. 3 and Experiment 2 vs. 4.

Summary and Conclusions

The present chapter discussed the issue of multiple activation and cross-language priming in the course of processing language by bilingual speakers. It started with a review of the CMLP, which has been used widely in both monolingual (e.g., Swinney, 1979; Swinney & Osterhout, 1990; Tabossi, 1988, 1996) and bilingual studies (e.g., Li & Yip, 1998) because of its sensitivity to semantic and associative relations, as well as contextual effects. We next provided a brief overview of research concerning multiple language activation at the lexical level, with studies using mostly interlingual homographs and cognates (cf. Libben & Titone, 2009; Schwartz & Kroll, 2006; Whitford et al., this volume), in order to determine whether bilingual lexical access is language selective or nonselective. Then, we looked at studies exploring bilingual processing at the sentence level which identified a number of factors influencing the comprehension of mixed-language sentences, such as context or language-specific phonotactic constraints. Finally, we presented our study consisting of four experiments using the CMN, which is a variant of the CMLP, and explored the effects of context and cross-language priming in Spanish-English bilinguals. We specifically wanted to systematically manipulate the effects of sentential context, a factor known to involve semantic processing, to examine some of the predictions of the Revised Hierarchical Model of bilingual memory representation (Kroll & Stewart, 1994). Previous studies addressing the claims of this model have operated under the assumption that conceptual and semantic information can be obtained by the manipulation of factors such as concreteness (e.g., De Groot, 1992) and category effects (e.g., Kroll & Stewart, 1994), using the isolated word or the picture as the experimental unit (c.f., Hummel, 1993). In the present study, we systematically manipulated the effects of biased vs. unbiased context during the online comprehension of spoken sentences.

In all four experiments, participants listened to sentences containing a critical prime (e.g., *war*); they then named a visually presented target that was either related (e.g., *peace*) or unrelated (e.g., *boca*) to the critical prime. Target words were either in Spanish or English, and target presentation for all experiments occurred immediately at prime offset. In Experiment 1, participants listened to Spanish sentences in which the preceding context was unbiased towards the meaning of the critical prime. The results for this experiment revealed facilitatory priming for both Spanish and English targets, in that naming related targets was faster than naming unrelated targets for both language conditions. The priming effect observed for the Spanish targets replicates the robust within-language effect reported in the bilingual literature (e.g., Hernández, 2002; Hernández et al., 1996; Keatley et al., 1994; Keatley & De Gelder, 1992). More impressive was the cross-language priming effect observed even when the sentences were entirely in Spanish. However, it is important to note that the within-language priming effect was much greater than the cross-language priming effect. This finding, of course, is not a surprising one given that the aurally presented sentences were in Spanish (cf. Grosjean's, 1988, 1997).

Experiment 2 was similar to Experiment 1, except that the preceding context was biased towards the meaning of the critical prime. As in Experiment 1, facilitatory priming was observed for both within- and cross-language conditions. A comparison between these two experiments indicates that, while preceding biased context (Experiment 2) did not increase the priming effect for each language condition, context did have an effect on the overall processing of the target words. That is, participants were faster to name target words under biased- than under unbiased-contextual conditions (for similar results, see Li, 1996; Li & Yip, 1998; see also Becker, 1979).

Our objective for Experiments 1 and 2 was to specifically test the hypothesis generated by the Revised Hierarchical Model that lexical access from L1 to L2 is more likely to be affected by factors known to influence semantic/conceptual processing. In this case, the Revised Hierarchical Model would predict a significant increase in the cross-language priming effect, from Experiment 1 (contextually-unbiased sentences) to Experiment 2 (contextually-biased sentences). The results did not support this hypothesis, as the cross-language priming effect remained relatively constant from Experiment 1 to 2. However, the fact the cross-language priming effect was observed in both experiments supports a weak version of the hypothesis that lexical access from the L1–L2 is somewhat semantically/conceptually oriented.

Experiments 3 and 4 were identical to Experiments 1 and 2 respectively, with the exception that sentences were in English and the visually presented targets were in Spanish or English. Thus, Experiments 3 and 4 represented the L2 to L1 (English-Spanish) cross-language condition, and the within-language (English-English) condition. Both within- and cross-language priming effects were obtained in

Experiments 3 and 4. More surprising was the finding that the cross-language priming effect was actually greater than the within-language priming effect. This result contrasts with both the cross- and within-language priming effects in Experiments 1 and 2, which were the exact opposite. It appears as if more semantic processing had taken place for the cross-language condition than the within-language condition in Experiments 3 and 4. A comparison between Experiments 3 and 4 showed that context actually slowed down the processing of the visually presented targets. That is, bilinguals were actually faster to name Spanish and English targets in the contextually-unbiased condition (Experiment 3), than in the contextually-biased condition (Experiment 4). This finding suggests that the contextual information present in Experiment 4 influenced bilinguals, during the comprehension process, to generate specific predictions and expectations as to the possible targets that were most likely to match the preceding context (cf. Altarriba et al., 1996). Thus, more time was necessary, relative to Experiment 3, to incorporate the visually presented targets into the sentence. Notice that one expectation would be that the Spanish target should be more affected by the contextual information than English one because of the mismatch of language. However, this expectation was not supported by the data, as a 2-way interaction between Language (Spanish vs. English) and Type of Sentence (Spanish vs. English sentence) was not reliable.

Experiments 3 and 4 were designed to specifically examine the extent to which access from L2 to L1 reflected any semantic/conceptual influence. Thus, according to the Revised Hierarchical Model, one would expect no cross-language priming (see for example, Keatley & De Gelder, 1992; Keatley et al., 1994) or relatively small priming effects during bilingual lexical access from the bilinguals' L2 to their L1. A comparison across experiments shows that the priming effect of 60 ms for Experiment 3 (L2–L1) was significantly higher than the 13 ms priming effect for Experiment 1 (L1–L2). Likewise, the priming effect of 49 ms for Experiment 4 (L2–L1) was higher than the 14 ms priming effect for Experiment 2 (L1–L2; see Tables 6.1, 6.2, 6.3, and 6.4). As can be seen, it appears that our results in fact revealed the priming asymmetry predicted by the Revised Hierarchical Model; however, it was in the opposite direction. These results are significant and replicate the same pattern of results observed by Heredia (1995, 1997) using a translation task, and Altarriba (1992) using a translation-priming paradigm.

Other important comparisons across the four experiments revealed interesting findings. Comparisons between Experiment 1 (sentence in Spanish) and Experiment 3

Table 6.4 Mean reaction times and standard deviations (SD) in milliseconds for Spanish-English bilinguals as a function of language and relatedness in Experiment 4

	Relatedness		
Language	Related	Unrelated	Priming
Spanish	700 (161)	749 (176)	49*
English	646 (118)	668 (121)	22*

*$p < .05$

(sentence in English) revealed that speed of response was faster when the sentence was in Spanish (the participants' L1) and the target word was in English (the participants' L2), than when the sentence was in English and the target word was in Spanish. These same patterns were consistent in Experiments 2 and 4, in that bilinguals were faster to retrieve L2 information from their L1, rather than L1 information from their L2. Again, we note the reversal of the language asymmetry.

Overall, with regards to the data presented here, the Revised Hierarchical Model does a good job of predicting some of the current findings. Significant priming was observed in the L1–L2 direction (Experiment 1). That is, the semantic or conceptual information accessed in L1 aided the retrieval and processing of subsequent material appearing in L2. However, adding a biasing context (Experiment 2) did not enhance this effect; that is, there was no further facilitation provided by strengthening the conceptual access of items in L1 on the processing of items in L2. Moreover, priming effects emerged from L2 to L1 (Experiments 3 and 4); effects that are likely based on conceptual or semantic processing more than mere lexical processing, as the model would suggest. In fact, the priming effect for this language direction was significantly higher than the L1 to L2 direction (see Tables 6.1, 6.2, 6.3 and 6.4). Although it could be argued that the present findings were due to the nature of the task, and that the Revised Hierarchical Model was not designed to address units larger than single words, other studies using single words (e.g., Altarriba, 1992; Altarriba & Mathis, 1997; Heredia, 1995, 1997; Heredia & Altarriba, 2001; Hernández, 2002; see also De Groot et al., 1994; Blot, Zárate, & Paulus, 2003) have shown similar patterns.

However, the Revised Hierarchical Model could account for the present results by assuming that, after a certain degree of proficiency in the L2, language retrieval is a function of *language dominance* (e.g., Heredia, 1997; Heredia & Altarriba, 2001). That is, which language is spoken and used more often would determine the ease of accessibility. Thus, according to this view, it would be possible for the bilingual's L2 to become the dominant language and behave as if it were the L1. Indeed, this is what we hypothesize occurred with the bilingual participants in the present study. Although Spanish was clearly our participants' L1, their usage of L2, both at the social and educational levels, was greater than that for L1. In fact, most of our participants reported feeling more comfortable communicating in their L2. They reported more code-switching as they communicated in their L1, that is, more English intrusions as they spoke Spanish, and little or no interference as they communicated in English (see for example, Heredia & Altarriba, 2001). This general experience is not unusual for most bilinguals in the Southwest of the United States. In short, results of the current study support the view of the bilingual memory storage as a dynamic system that can be influenced by such factors as language dominance (cf. Heredia, 1997; see also Cieślicka, Heredia, & Olivares, 2014). Language dominance, context, and other factors affecting bilingual processing certainly await further investigation to help us better understand the multifaceted nature of the bilingual mind. Overall, it seems from the study reported here that, in line with Garcia et al. (2015) and Cieślicka (2006), the CMLP is a sensitive tool that can reveal the nature of bilingual lexical activation. With a sensitive methodology like CMLP or the CMN task, it is possible to show, as we have done in the four experiments reported here, that both languages become activated in the course of bilingual lexical access.

List of Keywords

Backward priming, Base language effect, Biased context, Bilingual lexicon, Code-switching, Cognates, Conceptual store, Cross-language activation, Cross-modal lexical priming (CMLP), Cross-modal naming tasks (CMN), Cue-shadowing task, Demasking task, Go/No-go task, Interlingual homographs, Language mixing, Language dominance, Language selective view, Lexical access, Lexical ambiguity. Lexical decision task, Lexical decision time, lexical priming, Lexical-level processing, Multiple language activation, Naming latency, Naming task, Nonselective view, Phoneme-triggered decision task, Phonological inhibition effect, Phonotactic configuration, Positive priming effect, Prime, Priming, Revised Hierarchical Model, Semantic memory, Within-language, Word frequency

Review Questions

1. What are your thoughts about the bilingual mind and multiple language activation?
2. Think about the different issues addressed in this chapter. What are some of the issues that most impacted your understanding of the bilingual mind? Do you think it is possible for a bilingual/multilingual speaker to walk around with his/her multiple languages activated?
3. Do you code-switch? What is your view on bilinguals mixing their two languages during the communicative process? Why do you think bilinguals code-switch?

Suggested Student Research Projects

1. Design and conduct a simple word association task to see whether bilinguals activate both of their lexicons and whether the language of instruction influences this activation. You will need a group of bilinguals who share the same L1 and L2 (e.g., Spanish-English bilinguals). For your materials, prepare a list of L1 words. The list should contain at least 20 words that are cross-language homographs (e.g., SIN = "wrongdoing" in English and "without" in Spanish). The remaining 20 words should be neither cognates nor cross-language homographs. Present the words one at a time and ask your participants to write down the first word that comes to mind as they see each consecutive word. It is important that they provide immediate associations, without thinking too much. To see whether the language of instruction has an effect on their performance, divide your participants randomly into two groups. Use English only as language of interaction and instruction with one group and Spanish only with the other group. The instructions should be identical for both groups, with the exception

that they will be administered in either L1 or L2. What results did you obtain? Did putting participants in a bilingual mode, by providing instructions in L2, increase the number of cross-language associations? Were cross-language homographs more likely to elicit responses in the other language than nonhomographic targets? What do you conclude about the effects of stimulus materials and language of instruction on the activation of the L1 and L2 lexicons? Do the results support the language selective or nonselective view of bilingual lexical access?

For this research project, you need experimental software that will allow you to record participant's reaction time. You can download the demo version of E-Prime from http://www.pstnet.com/eprime.cfm. Alternatively, you can use OpenSesame, the free open source experiment builder (http://osdoc.cogsci.nl/) You are going to design and run an experiment on cross-language priming. Obtain a group of bilinguals who speak English as a second language and who share their first language background (e.g., Polish-English or Spanish-English bilinguals). Construct 20 word pairs in L1–L2 in which L1 is semantically related to the target L2 (e.g., for a Polish-English bilingual, the related stimuli pair would be KOT-DOG, where KOT = "cat" in English). Next, obtain 20 control word pairs, in which the L1 word is unrelated to the L2 word (e.g., KOC-DOG, where KOC = "blanket"). Assure that your related and unrelated L1 words are matched in terms in frequency, the number of syllables, and length. Comparing RTs to the target following the control vs. the related word will be your measure or dependent variable of the degree of priming. To check for word frequency and other word's characteristics, use the MRC Psycholinguistic Database (http://www.psych.rl.ac.uk/). Moreover, create 40 additional pairs consisting of L1 words and L2 nonwords, such as KOT-PLOG. English Nonwords can be found at http://www.cogsci.mq.edu.au/~nwdb/nonwords.html or by using Wuggy (http://crr.ugent.be/programs-data/wuggy) to create your own nonwords in English or Spanish. It is critical that the nonwords conform to the phonological rules in English, which means they need to be pronounceable. Once all your stimuli are ready, design a lexical decision task, in which participants are presented with the stimulus pairs in such a way that they first see the L1 word for 400 ms, and then are shown the L2 target for the lexical decision (i.e., they have to decide, as fast and as accurately as possible, if the presented L2 string is a legitimate English word or not). You will need to make two different lists for stimulus presentation, so as to avoid showing the same target word twice. For example, if you show the pair KOT-DOG in one presentation list, then the control KOC-DOG needs to be in another list, to avoid priming through repetition of the same item. This is known as counterbalancing. What are your results? Did you find reaction time differences? Did the study show priming from related L1 words, as compared to the control unrelated words? What can you conclude about whether bilingual lexical access is language selective or nonselective?

2. Using materials from the previous research project, you are going to check whether the direction of priming (L1–L2 vs. L2–L1) has an effect on participants' performance. Take all the stimuli from Project 2 and reverse their order (i.e.,

KOT-DOG will now become DOG-KOT). In addition, prepare 20 L2–L1 pairs where L1 are nonwords. Prepare two mixed lists with half of the stimuli in the L1–L2 and half in the L2–L1 direction, making sure the same pair is not repeated in the same list (i.e., if you are presenting KOT-DOG, do not include DOG-KOT in the same list). Compare the priming effects obtained in both directions. Did you find a difference between the two conditions? Can the results be interpreted within the framework of the Revised Hierarchical Model?

Related Internet Sites

CLEARPOND (Cross-linguistic easy-access resource): http://clearpond.northwestern.edu/
Cross-modal priming task: https://en.wikipedia.org/wiki/David_Swinney
English Lexicon Project: http://elexicon.wustl.edu/
Espal (Spanish database): http://www.bcbl.eu/databases/espal/
Multiple lexical access: http://cogweb.ucla.edu/Abstracts/LexicalAccess.html
Semantic priming: https://en.wikipedia.org/wiki/Priming_%28psychology%2
Tatool Web (experiments online): http://www.tatool.ch/
Word frequency (American English): http://expsy.ugent.be/subtlexus/

Suggested Further Reading

Foss, D. (1970). Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *9*, 699–706.
Grosjean, F. (1996). Gating. *Language and Cognitive Processes*, *11*, 597–604.
Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer. *Language and Speech*, *32*, 1–23.
Prather, P. A., & Swinney, D. A. (1988). Lexical processing and ambiguity resolution: An autonomous process in an interactive box. In S. I. Small., G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics*, *neuropsychology & artificial intelligence* (pp. 289–310). San Mateo, CA: Morgan Kaufmann.

Acknowledgements The second author is grateful to the late Elizabeth Bates and David Swinney for their advice, support, and inspiration as postdoctoral advisors, at the Center for Research in Language, and the Center for Information Processing at the University of California, San Diego. We also acknowledge the Center for Research in Language for invaluable assistance with data collection in Experiment 2. Experiment 2 was supported by National Science Foundation, Grant SBR-9520489.

References

Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 157–174). Amsterdam: Elsevier.

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods, 39*, 1–18.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language, 36*, 550–568.

Bates, E., & Liu, H. (1996). Cued shadowing. *Language and Cognitive Processes, 11*, 577–581.

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language, 26*, 658–672.

Becker, C. A. (1979`). Semantic context and word frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 252–259.

Blank, M. A. (1980). Measuring lexical access during sentence processing. *Perception & Psychophysics, 28*, 1–8.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology Learning Memory and Cognition, 19*(2), 295–308.

Blot, K. J., Zárate, M. A., & Paulus, P. B. (2003). Code-switching across brainstorming sessions: Implications for the revised hierarchical model of bilingual language processing. *Experimental Psychology, 50*, 171–183.

Bruning, J. L., & Kintz, B. L. (1987). *Computational book of statistics*. Glenville, IL: Harper Collins.

Burgess, C., Tanenhaus, M. K., & Seidenberg, M. S. (1989). Context and lexical access: Implications of nonword interference for lexical ambiguity resolution. *Journal of Experimental Psychology Learning Memory and Cognition, 15*(4), 620–632.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*, 668–683.

Cheung, H., & Chen, H.-C. (1998). Lexical and conceptual processing in Chinese-English bilinguals: Further evidence for asymmetry. *Memory & Cognition, 26*, 1002–1013.

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by L2 speakers. *Second Language Reserach, 22*, 115–144.

Cieślicka, A. (2007). Language experience and fixed expressions: Differences in the salience status of literal and figurative meanings of L1 and L2 idioms. In M. Nenonen & S. Niemi (Eds.), *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, Finland, May 19-20* (pp. 55–70). Joensuu, Finland: Joensuu University Press.

Cieślicka, A. B., Heredia, R. R., & Olivares, M. (2014). The eyes have it: How language dominance, salience, and context affect eye movements during idiomatic language processing. In L. Aronin & M. Pawlak (Eds.), *Essential topics in applied linguistics and multilingualism. Studies in honor of David Singleton* (pp. 21–42). New York: Springer.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers, 25*, 257–271.

De Groot, A. M. B. (1992). Determinant of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1001–1018.

De Groot, A. M. B., Dannenburg, L., & Van Hell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language, 33*, 600–629.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). Cross-language effects on bilingual homograph recognition. *Journal of Memory and Language, 42*, 45–464.
Dijkstra, T., Van Jaarsveld, H., & Brinke, S. T. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism Language and Cognition, 1*, 51–66.
Fox, E. (1996). Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory and Language, 35*, 353–370.
Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
García, O., Cieślicka, A. B., & Heredia, R. R. (2015). Nonliteral language processing and methodological considerations. In R. R. Heredia & A. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 117–168). New York: Cambridge University Press.
Gerard, L., & Scarborough, D. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology Learning Memory and Cognition, 15*, 305–315.
Glucksberg, S., Kreuz, R. J., & Rho, S. H. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 323–335.
Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and non-cognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122–1139.
Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3*, 233–274.
Grosjean, F. (1997). Processing mixed language: Issues, findings and models. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Mahwah, NJ: Erlbaum.
Heredia, R. R. (1995). *Concreteness effects in high frequency words: A test of the revised hierarchical and the mixed models of bilingual memory representations*. Unpublished doctoral dissertation, University of California, Santa Cruz.
Heredia, R. R. (1997). Bilingual memory and hierarchical models: A case for language dominance. *Current Directions in Psychological Science, 6*, 34–39.
Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*, 164–168.
Heredia, R. R., & Blumentritt, T. L. (2002). On-line processing of social stereotypes during spoken language comprehension. *Experimental Psychology, 49*, 208–221.
Heredia, R. R., & Brown, J. M. (2003). Bilingual memory. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of bilingualism* (pp. 213–228). Malden, MA: Blackwell.
Heredia, R. R., & Muñoz, M. E. (2015). Metaphoric reference: A real-time analysis. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 89–116). New York: Cambridge Press.
Heredia, R. R., & Stewart, M. T. (2002). On-line methods in bilingual spoken language research. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 7–28). North Holland, The Netherlands: Elsevier.
Hernández, A. (2002). Exploring language asymmetries in early Spanish-English bilinguals: The role of lexical and sentential context effects. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 7–28). North Holland, The Netherlands: Elsevier.
Hernández, A. E., Bates, E., & Ávila, L. X. (1996). Processing across the language boundary: A cross-modal priming study of Spanish-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 846–864.
Herron, D., & Bates, E. (1997). Sentential and acoustic factors in the recognition of open- and closed-class words. *Journal of Memory and Language, 37*, 217–239.
Hillert, D., & Swinney, D. (2001). The processing of fixed expressions during sentence comprehension. In A. Cienki, B. J. Luka, & M. B. Smith (Eds.), *Conceptual and discourse factors in linguistic structure* (pp. 107–121). Stanford, CA: CSLI.
Hummel, K. M. (1993). Bilingual memory research: From storage to processing issues. *Applied Psycholinguistics, 14*, 267–284.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*, 32–51.

Julliand, A., & Chang-Rodríguez, E. (1964). *Frequency dictionary of Spanish words*. La Haya, Holland: Mouton.

Keatley, C., & De Gelder, B. (1992). The bilingual primed lexical decision task: Cross-language priming disappears with speeded response. *European Journal of Cognitive Psychology, 4*, 273–292.

Keatley, C. W., Spinks, J. A., & De Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition, 22*, 70–84.

Kolers, P. (1966). Reading and talking bilingually. *American Journal of Psychology, 3*, 357–376.

Kroll, J. F., & Stewart, E. (1994). Category inference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velder, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for conceptual mediation. *Journal of Memory and Language, 35*, 648–665.

Lederberg, A. R., & Morales, C. (1985). Code switching by bilinguals: Evidence against a third grammar. *Journal of Psycholinguistic Research, 14*, 113–136.

Li, P. (1996). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language, 35*, 757–774.

Li, P., & Yip, M. C. (1998). Context effects and the processing of spoken homophones. *Reading and Writing: An Interdisciplinary Journal, 10*, 223–243.

Libben, M., & Titone, D. (2009). The multidetermined nature of idiomatic expressions. *Memory & Cognition, 36*, 1103–1131.

Liu, H., Bates, E., Powell, T., & Wulfeck, B. (1997). Single-word shadowing and the study of lexical access. *Applied Psycholinguistics, 18*, 157–180.

Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology, 50*, 204–216.

Macnamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior, 10*, 480–487.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

Masson, M. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 3–23.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance* (Vol. V). New York: Academic Press.

Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current theories and findings. In D. Besner & G. Humphries (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology Learning Memory and Cognition, 15*(6), 1003–1019.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* http://www.usf.edu/FreeAssociation/

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9*(3), 225–236.

Prather, P. A., & Swinney, D. A. (1988). Lexical processing and ambiguity resolution: An autonomous process in an interactive box. In S. I. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence* (pp. 289–310). San Mateo, CA: Morgan Kaufmann.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behaviour, 23*, 84–99.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489–537.

Sereno, S. C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology Learning Memory and Cognition, 21*(3), 582–595.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science, 6*, 45–49.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120–136.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a speech model: The effect on lexical access. *Memory & Cognition, 12*, 380–386.

Stewart, M. T., & Heredia, R. R. (2002). Comprehending spoken metaphoric reference: A real-time analysis. *Experimental Psychology, 49*, 34–44.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Swinney, D., Love, T., Walenski, M., & Smith, E. E. (2007). Conceptual combination during sentence comprehension: Evidence for compositional processes. *Psychological Science, 18*(5), 397–400.

Swinney, D. A., & Osterhout, L. (1990). Inference generation during auditory language comprehension. *The Psychology of Learning and Motivation, 25*, 17–33.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27*, 324–340.

Tabossi, P. (1996). Cross-modal semantic priming. *Language and Cognitive Processes, 11*, 569–576.

Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 145–162). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tanenhaus, M. K., & Donnenwerth-Nolan, S. (1984). Syntactic context and lexical access. *The Quarterly Journal of Experimental Psychology, 38A*, 649–661.

Titone, D. A., & Connine, C. M. (1994). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology Learning Memory and Cognition, 20*(5), 1126–1138.

Van de Voort, M. E. C., & Vonk, W. (1995). You don't die immediately when you kick an empty bucket: A processing view on semantic and syntactic characteristics of idioms. In M. Everaert, E. J. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 283–300). Hillsdale, NJ: Lawrence Erlbaum Associates.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188–208.

Warren, R. E. (1977). Time and the spread of activation in memory. Journal of Experimental Psychology: Human Learning and Memory, 3, 458–466.

Chapter 7
Reading Integration in Bilingual Speakers

Gary E. Raney and Joanna C. Bovee

Abstract The purpose of this chapter is to describe research methodologies that can be used to examine when and how bilingual readers integrate their representations of texts across languages and how their languages influence these representations. Text representations are described in terms of Kintsch and van Dijk's concepts of *Surface Form, Textbase, and Situation model*. We review research methodologies designed to explore each of these levels of representation, such as script manipulations (surface form manipulation), code switching, and mixed-language reading (textbase manipulations), and transfer benefits and background knowledge (situation model manipulations). For each study reviewed, we describe key independent and dependent variables and key procedural issues that need to be controlled, such as properties of the stimuli. The research described here supports the conclusion that bilinguals can integrate text representations at multiple levels.

Introduction

The purpose of this chapter is to explore research methodologies that can be used to examine when and how bilingual readers integrate their representations of texts across languages. To do this, we examine methodologies designed to explore how bilinguals read, comprehend, and remember text. The focus of the chapter is the research methodologies, not the theories underlying whether representations might be integrated in bilinguals (for reviews of theoretical issues, see Francis, 1999, and Raney, Obeidallah, & Miura, 2002). Of course, research methodologies are designed to test theories; therefore, we will discuss theories to the extent necessary to explain the research methodologies. We also limit our review primarily to behavioral tasks such as those that measure naming time, decision times, and reading time.

G.E. Raney, Ph.D. (✉) • J.C. Bovee
Department of Psychology (MC 285), University of Illinois at Chicago,
1007 W. Harrison Street, Chicago, IL 60607, USA
e-mail: geraney@uic.edu; jbovee2@uic.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_7

For each type of methodology discussed, we provide one or more example studies to illustrate the method, details about the procedure that need to be controlled, and how the methodology helps us understand bilingual language integration. We discuss language integration in terms that roughly correspond with Kintsch and van Dijk's concepts of *Surface Form, Textbase, and Situation model* (Kintsch, 1998; Kintsch & Van Dijk, 1978; Van Dijk & Kintsch, 1983). According to Kintsch and van Dijk, the surface form contains the wording used in the text and reflects lexical and syntactic analysis of a text (Fletcher & Chrysler, 1990; Van Dijk & Kintsch, 1983). The textbase represents the meaning of the words and phrases in a text (i.e., concepts) as an organized network of propositions but is independent of the actual words (Kintsch, 1974; Kintsch & Van Dijk, 1978). For example, the sentences *Jane sat on the couch* and *Jane sat on the sofa* have different surface forms but similar textbases because these sentences express essentially the same concepts. This is relevant to bilingual reading because words from each language can be thought of as synonyms that express the same concept, such as *dog* (English) and *perro* (Spanish). The situation model is an integration of the textbase with a reader's prior knowledge and reflects how the reader interprets a text. The situation model includes information that is not part of the original text, such as general knowledge of the topic, inferences, and other by-products of text comprehension (Fletcher, 1994; Kintsch, 1988; Kintsch, Welsch, Schmalhofer, & Zimny, 1990).

We begin our review by describing methodologies that are designed to explore the representation and integration of languages at the surface level, which is roughly equivalent to what is called the lexical or word level in the literature on bilingualism. Our goal here is to describe research methods that are designed to explore whether and how lexical entries are linked across languages. We then describe methodologies that are designed to examine higher-level representations, such as the textbase and situation model, which is similar to what is called meaning-based or conceptual-based levels in the literature on bilingualism. Our goal here is to describe research methods designed to measure the degree of integration of meaning across languages. As readers will see, many research methodologies are designed to provide information about multiple levels of representation. We conclude by offering methodological suggestions for exploring whether textual memories are integrated.

We discuss research methods in terms of levels of representation for several reasons. First, there is substantial research supporting this model of representation (see Kintsch, 1998; Zwaan & Radvansky, 1998, for reviews). There is also substantial evidence that these levels of representation are somewhat independent, and this supports the contention that bilinguals could form representations that are integrated or independent at each level (see Nassaji, 2002; Raney et al., 2002, for reviews related to bilingual reading).

Although we use Kintsch and Van Dijk's (1978) model as a framework for discussing the research methodologies, we believe our conclusions extend beyond this model. Many models of text and language representation have been proposed, but a common element among several prominent models is the inclusion of multiple levels of representation. There is continuing debate about how many levels of representation are needed, especially at higher levels. For instance, Kintsch (1998), Graesser, Millis, and Zwaan (1997), and Graesser and McNamara (2011) describe models with four or five levels of representation. Our purpose is to discuss research methodologies; we are not endorsing a particular theory of language representation.

A second reason for discussing methods in terms of levels of representation is that several prominent models of bilingual language representation also propose multiple levels of representation. For example, early models proposed by Ervin and Osgood (1954) and Weinreich (1953) proposed a memory structure in which languages were separated, but each language had word-level representations attached to distinct conceptual representations. More recent models, such as the *Word Association Model* by Potter, So, Eckardt, and Feldman (1984) and the *Revised Hierarchical Model* by Kroll and colleagues (Kroll & Stewart, 1994; Kroll, Van Hell, Tokowicz, & Green, 2010) specifically refer to a lexical level of representation and a conceptual level of representation. In these models, there are links between languages at the lexical and conceptual levels, with the direction and strength of the links determining the degree of integration between languages. Even models that are described as being nonhierarchical, such as de Groot's *Distributed Conceptual Feature Model of Bilingual Memory* (De Groot, 1992, 1993; Kroll & De Groot, 1997; Van Hell & De Groot, 1998), include a lexical level and conceptual level of representation. Word- or lexical-level representations are similar to surface representations and conceptual-level representations are similar to textbase and situation model representations.

Before continuing, we want to offer our perspective on research methodologies for studying bilingual language processing. Our viewpoint is that there are no uniquely *bilingual research methodologies*. Instead, we take the perspective that there are research methodologies for exploring cognitive processes that can be more or less easily applied to study bilingualism. If one assumes unique methodologies must be used for studying bilinguals, one is assuming bilingual and non-bilingual cognitive processes are also unique. A concrete illustration might clarify our point. One common area of research on bilingualism is language selection. The general goal of this research is to determine how bilinguals select which of their languages should be used in a particular situation. A common research methodology used to explore this issue is the language switching task. One way this task is performed is to present pictures to bilinguals and ask them to name the pictures in one language or the other. The language to be spoken on each trial is cued in some manner, such as by the presence of a colored square around the picture to be named (Meuter & Allport, 1999). For instance, English-Spanish bilinguals might be asked to name pictures in English if the square is in green and to name pictures in Spanish if the square is in blue. Naming time is then measured for two different types of trials, repeated trials, and switch trials. In repeated trials, the current trial and the previous trial are named in the same language. In switch trials, the current trial and the previous trial are named in different languages. A robust finding is that the time needed to name pictures on switch trials is longer than the time needed to name pictures on repeated trials (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Philipp & Koch, 2009). This is known as a switch cost. The more integrated the languages are in memory, the smaller the switch cost should be.

The language switching task might seem unique to bilinguals because two languages are required. However, the underlying methodology reflects a standard task switching paradigm. Task switching paradigms have been used to examine many cognitive functions (see Kiesel et al., 2010, for a review). For instance, Prior and MacWhiney (2010) examined attentional control by showing participants geometric shapes that were presented in one of two colors, such as a triangle in green ink or a square in red ink. Participants were cued to identify either the shape of the figure or the color of the figure. As in the bilingual language switching task, participants were presented with repeated trials, such as identifying the shape on two consecutive trials, or switch trials, such as identifying the shape on the current trial and the color on the previous trial. Both of these experiments use task switching methodologies, one of which is simply adapted to study issues related to bilingual lexical access. In each method, the primary independent variable is trial type (repeated trials and switch trials) and the primary dependent measure is response time.

In the remainder of this chapter, we present research methodologies that have been adapted to study bilingual language processing. We begin by describing methodologies that have manipulated lower-level linguistic features.

Script Manipulations

Script manipulations, both within a single language and across languages, can be used to examine how readers' L1 (first language) orthographical knowledge can interfere or facilitate L2 (second language) word recognition and comprehension. Such manipulations can use a single language and vary properties within a language, such as case alternation (e.g., *cAsE aLtErNaTiOn*) or font change; use two or more languages, such as script substitution (e.g., including a Sanskrit symbol in an English passage or alternating between the scripts of two languages), or vary the text based on the different properties of the two languages. The dependent measure is most often reading time, though reading comprehension data are often collected as well. Languages compared often differ in type of orthographical system, such as logography (one grapheme represents an entire word or morpheme), syllabary (one grapheme represents a syllable), or alphabetical (one grapheme represents a phoneme).

One type of script manipulation is case alternation, which can be used to look at the underlying mechanisms involved in word processing. Akamatsu (2003) used this methodology to examine how orthographic features from an L1 affect word recognition in L2. Participants had different L1 backgrounds, but all had English as their L2, and they were matched on English proficiency. All participants read English texts with letters alternating between upper- and lower-case letters. Participants whose L1s were Chinese and Japanese, which are non-alphabetical languages, were predicted to have more difficulty reading a text in their L2 than those whose L1 was Persian, an alphabetic language. Akamatsu's rationale was that word recognition in Chinese and Japanese depends more on word shape information; thus, recognizing words with the added variation from alternating case will require a skill they would not have highly developed. In contrast, participants whose L1 was Persian will have already developed the skill to recognize individual letters and will, therefore, be less affected by the case manipulation.

The basic design was as follows. Twelve texts of approximately 100 words were presented to the participants via computer, half of which were presented normally (typical case usage) and half were presented with alternating case. Each text was followed by 4–6 multiple choice comprehension questions. Participants were instructed to read each passage as quickly as possible and that reading time for each passage would be recorded. Additionally, they were told that once they were finished reading a passage, they would answer comprehension questions based on the passage. Looking back at the passage during the comprehension portion was not permitted. Participants pressed the spacebar to begin reading (the passage appeared on the computer screen) and pressed the spacebar when they finished reading the passage. Reading time was recorded from stimulus onset (initial spacebar press) to participant response (second spacebar press).

The primary independent variables in this experiment were case use (typical, alternating) and L1 background (Chinese, Japanese, Persian). The dependent measures were overall passage reading time and comprehension question accuracy. Akamatsu found a significant interaction between case and L1. All participants read the normal case passages faster than the alternating case passages, but Chinese and Japanese L1 participants were slowed down more by the case manipulation than Persian participants. All participants also comprehended less when reading alternating script passages, although there was no interaction with L1. These results suggest that the orthography of the L1 may influence how L1 orthographical knowledge is integrated with L2 orthographical knowledge, ultimately affecting word recognition efficiency in text reading, but likely does not influence overall comprehension.

This study has several methodological strengths. To start, the researchers used 12 texts, which provides multiple (6) texts for each case manipulation (normal vs. alternating). Using multiple passages provides more data points, which in turn should provide more reliable data. Additionally, by presenting the passages via the computer rather than on paper, a more precise measure of reading time is collected. Lastly, rather than using only one language group for the non-alphabetic group, the researchers used native Chinese and Japanese bilinguals. This increases generalizability and allows us to be more confident that the differences between conditions are not solely language based, but are orthography based.

Another type of script manipulation is script substitution. Koda (1990) used this methodology to examine how L1 orthographic knowledge is integrated with L2 phonological encoding strategy. Koda included participants from four different L1 backgrounds: Arabic, Spanish, Japanese, and English (control group). All participants were skilled readers in English. Participants read English passages that contained some Sanskrit symbols or some pronounceable English nonsense words. Koda predicted that participants who were morphographic readers (Japanese L1 speakers) would be less impaired when phonological information was not available in the graphemic representation, as in the case of Sanskrit symbols, than would be the phonographic readers (English, Spanish, and Arabic L1 speakers).

The basic methodology was as follows. Each participant read two passages containing approximately 350 words per passage. One passage was about fictitious types of fish and the other was about five fictitious cocktails. In the experimental condition, fish and cocktail names were replaced with Sanskrit symbols and in the control condition names were presented as pronounceable English nonsense words. Passage presentation order (fish first vs. cocktail first) and experimental condition (Sanskrit vs. nonsense word) were counterbalanced across L1 groups in a repeated measures design. Participants were told their reading times would be measured, that each passage would be followed by a recall test, and that they would rate their confidence in their accuracy on the recall test.

The primary independent variables in this experiment were L1 background (English, Spanish, and Arabic) and critical word type (Sanskrit, nonsense word). The dependent measures were overall passage reading time, comprehension accuracy, and comprehension test confidence ratings. When analyzing reading times, Koda found that all participants spent more time on passages with Sanskrit symbols than English nonsense words; however, the language groups differed in the degree to which they were affected by phonological inaccessibility. Japanese readers were the least affected and, in half the cases, actually spent less time on the passages containing Sanskrit symbols than on passages containing English nonwords. English readers were the most adversely affected by the inclusion of Sanskrit symbols. Analysis of passage recall indicated that all groups performed best on the passages that included nonsense words, although there was no difference as a function of language background. Lastly, there were no differences in confidence ratings for the comprehension tests. Together these results suggest that L1 orthographic strategies do transfer to L2 reading processes and orthographic level representations can be integrated between languages to facilitate L2 processing.

Koda's (1990) methodology has several strengths. Although the experiment only included two passages, the passages were longer in length than those used by Akamatsu (2003). This allows for many instances of the symbol/nonwords manipulation to occur. Additionally, the script manipulation was within-subjects so that all participants were exposed to both levels, which allowed better control of individual differences between participants.

An additional way of examining how L1 script affects L2 reading is to manipulate the script at the word level instead of the passage level. This is commonly done using a lexical decision task. The lexical decision task entails presenting the participant a letter string, such as *word* or *wrod*, and the participant must decide whether the letter string is a word or a nonword. Participants can be instructed to decide whether the string is a word in a single language (e.g., English) or multiple languages (e.g., English or Spanish), or whether the string was previously seen or not.

Muljani, Koda, and Moates (1998) used the lexical decision task to examine the relative influence of L1–L2 orthographic distance and L2 word frequency on L2 word recognition. They predicted that higher L2 word frequency would lead to faster word recognition times regardless of orthographic distance between L1 and L2. They also predicted that the orthographic distance between L1 and L2 would influence L2 word recognition. Specifically, participants with a close L1–L2 orthographic distance would be more efficient in L2 word recognition than participants with a large L1–L2 orthographic distance. Furthermore, participants with a close L1–L2 orthographic distance would be more sensitive to incongruencies between consonant and vowel patterns in their L2 relative to their L1.

The basic methodology was as with. Researchers compared English lexical decision task performance between three groups of participants: Chinese (logographic) L1 speakers, Indonesian (alphabetical) L1 speakers, and English (alphabetical) L1 speakers. Chinese and Indonesian speakers were matched on their L2 English proficiency. All participants were presented with four sets of 24 critical English words, 24 critical English nonwords, 8 filler English words, and 8 filler nonwords. Real words varied based on word frequency. Real words and nonwords varied based on the similarity of their syllabic pattern to Indonesian. For example, Consonant-Vowel-Consonant (CVC) is a common letter pattern in both English and Indonesian, whereas CCCVCC only occurs in English. Frequency and congruency was counterbalanced across word sets. Within a set, word orders were randomized. Stimuli were presented via computer. For each trial, participants were presented with a "get ready" signal for 1000 ms in the center of the screen. Next, a letter string was presented in uppercase letters until the participant indicated it was a real word or a nonword by pressing one of two keys on a keyboard. Once the response was made, the letter string disappeared and the ready screen for the next trial appeared. Participants had been informed that they would be measured on speed and accuracy.

The primary independent variables in this experiment were L1 background (English, Chinese, and Indonesian), critical word type (critical word, filler word), letter string type (word, nonword), word frequency (high, low), and spelling pattern congruency (English and Indonesian congruent, English only congruent). The dependent measures were reaction time and lexical decision task accuracy.

Muljani et al. (1998) found that L1 orthographic background had a significant effect on overall reaction time, with English readers being faster than Indonesian readers who were faster than Chinese readers. Additionally, reaction time was faster for high than for low frequency words. Finally, spelling pattern congruency also had a significant effect on reaction time such that responses were faster for congruent words than for incongruent words. Spelling pattern also interacted with language background. English speakers were not affected by word frequency or congruency manipulations, whereas Indonesian speakers were faster to respond to words with congruent spelling patterns as compared to words with incongruent spelling patterns, regardless of word frequency. Chinese speakers showed the same congruency effect as the Indonesians, but only for low frequency words. Error rates were low for all groups. Results for the nonword analysis essentially mirrored those of the real word data, though some effects were slightly attenuated.

These results indicate that high frequency words in L2 are less affected by L1 knowledge than are low frequency words, and that word recognition efficiency is affected by degree of orthographic similarity between two languages. Taken together, these results support the conclusion that knowledge of L1 and L2 letter patterns can be integrated as orthographic similarity can either facilitate or interfere with word processing.

Cognates

Cognates are words that look or sound similar and share a common meaning across languages (e.g., English and Spanish word pairs *funeral* and *funeral*, and *problem* and *problema*). Cognates share perceptual and conceptual representations across languages whereas non-cognates only share conceptual features (Altarriba, 1992; Chen & Ng, 1989; De Groot, 1993; De Groot & Nas, 1991). In discourse terms, cognates share surface features and contribute the same meaning to the textbase, whereas non-cognates contribute the same meaning to the textbase but do not share surface features. Presenting cognates in a text is a useful method to examine how concepts are stored in the mental lexicon and how reading a word can activate its concept and translation equivalent in long-term memory.

According to the *Bilingual Interactive Activation Plus Model* (BIA+), when a concept is activated in memory, so are words from both of a bilingual's lexicons (Dijkstra & Van Heuven, 2002). A consequence of this phenomenon, known as the *cognate facilitation effect*, is that cognates are often processed and recognized faster than non-cognates. Access to cognates during reading is facilitated because there are two sources of activation in the lexicon that match across languages, conceptual activation and orthographic activation. Non-cognates, such as *dog* (English) and *perro* (Spanish), share one source of lexical activation, namely, conceptual activation. Cognates and non-cognates are often compared with interlingual homographs, which are words that look and sound similar, but do not share meaning, such as the English word *sensible* (appropriate, proper) and the Spanish word *sensible* (sensitive). *Interlingual homographs* can interfere with reading by activating an incorrect meaning in the lexicon that competes with the correct meaning, thus slowing down reading. These findings are robust and occur in both single-word contexts (Costa, Caramazza, & Sebastian-Galles, 2000) and sentence contexts (Schwartz & Kroll, 2006; Van Hell & de Groot, 2008).

Van Assche, Duyck, Hartsuiker, and Diependaele (2009) measured eye movements of bilingual participants as they read sentences in their native languages that either contained cognates or control (non-cognate) words. The purpose of the experiment was to examine how knowledge of a second language influences reading in a first language. They were specifically exploring whether bilinguals could restrict lexical retrieval when reading in their native language or whether their second language would be activated as well.

The basic design was as follows. Participants were native-Belgian speaking university students who began learning English around the age of 15. Their eye movements were recorded while they read 40 sentences with low-constraint contexts in their native language and 40 filler sentences. A sample sentence is presented below. Experimental sentences had been designed so that either the cognate or the non-cognate control word could be used in a natural manner. Cognate and non-cognate control words were counterbalanced across participants so that each participant only saw each sentence and critical word once. Sentences were presented in random order, one at a time, on a computer screen. All sentences occupied one line on the monitor. Participants started and stopped a trial by pressing a button on the keyboard. To ensure participants read for comprehension, a comprehension question was asked after a sentence 25 % of the time.

1. *Ben heft een oude OVEN/LADE gevonden tussen de rommel op zolder.*
   Ben found an old OVEN/DRAWER among the rubbish in the attic.

This design has several strengths. First, by using the same sentence contexts to present both cognate and control words, any differences seen in eye movement measures can be attributed to the manipulation rather than another aspect of the sentences. Additionally, by using low-constraint sentences, the researchers ensured that participants would not be able to predict the upcoming critical word when reading the sentence. The use of filler sentences serves to mask the purpose of the study, reducing the likelihood that participants will infer the cognate/non-cognate comparison. Lastly, by measuring eye movements, the researchers are able to measure online reading times for individual words without requiring secondary tasks. The primary independent variables in this experiment were sentence type (low-constraint context, filler) and critical word type (cognate, control). The dependent measures were various eye tracking measure of reading time.

Van Assche et al. (2009) found that first fixation duration (the duration of the initial fixation made on a word) and gaze durations (the sum of all fixations made on a word before moving to another word) were shorter for cognates than for non-cognate control words. They took this as evidence for a cognate facilitation effect. That is, representations of the less proficient second language were strong enough to aid in the word recognition process in the more proficient, native language.

Libben and Titone (2009) employed a methodology similar to that of Van Assche et al. (2009), but used cognates, control words, and interlingual homographs (e.g., the word *coin*, which means *money* in English and *corner* in French). Additionally, they varied the degree of semantic constraint. The purpose of the study was to investigate the effect of semantic constraint on nonselective lexical access.

The basic design was as follows: participants were French-English bilinguals (dominant in French). Eye movements were recorded while participants read sentences in English that contained a target word that was a cognate, an interlingual homograph, or a control word (32 of each type). Target words were matched on word length, frequency, and neighborhood density (the number of words that are highly similar to it). Sentences were composed of two clauses. The first clause varied on semantic constraint; high-constraint clauses were semantically biased toward the target word and low-constraint clauses were unbiased. The second clause contained the target word. Target words and constraint conditions were counterbalanced across sentences such that each target and its control appeared once in a high-constraint clause and once in a low-constraint clause. To ensure participants were reading for comprehension, comprehension questions were asked following 25 % of the sentences. The primary independent variables in this experiment were sentence type (low-constraint, high-constraint) and target word type (cognate, interlingual homograph, control word). The dependent measures included several measures of reading time, such as first fixation duration and total reading time.

Libben and Titone (2009) found that interlingual homographs were read more slowly, as measured by first fixation duration, than their matched control words in low- and high-constraint sentences. In comparison, cognates were read faster than interlingual homographs in both types of semantic constraint conditions. Both cognates and interlingual cognates were read faster in high-constraint than in low-constraint sentences. For total reading time, interlingual homographs and cognates only differed from their control items in the low-constraint sentences. The authors concluded that cross-language inhibition or facilitation only occurs when little semantic context is provided. With respect to integration, lexical representations from both languages were used in processing only when there was little supportive semantic context.

This study has several strengths. First, it uses 32 sentences for each condition (word type and constraint type). The large number of sentences is likely to increase the reliability of the data. The within-subject design controls for individual differences among participants, since every participant is exposed to all possible conditions. Additionally, the use of both low- and high-constraint sentences allows for more complete understanding of the effect of semantic context than did Van Assche et al.'s (2009) study, which only used low-constraint sentences.

Text Repetition and Transfer Methodologies

When texts are read twice, reading times typically decrease during the second reading. This is known as a *text repetition effect*. Reading one text can also facilitate processing of a different text, which is known as a *transfer benefit*. Because repetition effects and transfer benefits reflect the processes used to comprehend and represent a text in memory, the size of repetition effects and transfer benefits can be used as a measure of comprehension and memory (Levy, 2001; Raney, 2003). Raney (2003) suggests that repetition effects and transfer benefits can reflect processing of the surface form, textbase, and situation model, which makes it a useful method for examining integration of text representations across languages. To illustrate the basic logic of this task, we briefly summarize a non-bilingual study conducted by Raney, Therriault, and Minkoff (2000).

The basic design was a follows. Participants read a set of short passages, each twice in succession, and reading times were measured. The second reading was either the same text (repeated) or a paraphrased version of the original text. Paraphrases were formed by replacing approximately one third of the original words with synonyms. To illustrate, the sentences below (2–3) were taken from an original and paraphrased text about the creation of soap. Words that were changed between versions are underlined.

Original Passage

2. *The invention of soap was probably accidental. One Roman <u>legend suggests</u> that <u>fat</u> from candles that were <u>burned</u> in a <u>sacred ritual</u> was <u>mixed</u> with wood ashes.*

Paraphrased Passage

3. *The discovery of soap was most likely unplanned. One Roman story proposes that lard from candles that were lighted in a holy ceremony was blended with wood ashes.*

The independent variables in this experiment were reading (first, second) and passage type during the second reading (identical, paraphrased). The dependent variable was passage reading time.

Raney et al. (2000) found that both repeated and paraphrased texts were read significantly faster during the second reading, but the repetition effect was slightly smaller for paraphrased texts. Because large repetition effects were found for paraphrased texts, they concluded that text repetition effects primarily reflect repetition of concepts (textbase), not repetition of individual words (surface form). The importance of conceptual repetition has been shown in several other studies (Levy, Barnes, & Martin, 1993; Levy et al., 1995; O'Brien, Raney, Albrecht, & Rayner, 1997; Raney & Rayner, 1995).

Raney, Atilano, and Gomez (1996) extended the text repetition methodology to study bilingual text comprehension. Their research design was as follows: skilled and novice (not proficient) bilinguals read a text in one language (Spanish or English) and then read the same text again (same language) or read a translated version of the text (different language) while their eye movements were monitored. This created four conditions (English-English, Spanish-Spanish, English-Spanish, Spanish-English), and two texts were read in each condition. From a methodological standpoint, changing the language manipulates the surface form while keeping the textbase relatively constant. This is analogous to changing words to synonyms during the second reading of a text. Presenting texts in two languages provides a direct method to examine the integration of text representations at the word and conceptual levels. One way this can be done is by examining reading times for words that are cognates or non-cognates. Embedded in each passage are two cognate (C) and two non-cognate (N) target words. Below are two sample sentences from a pair of passages.

English Passage

4. *Many politicians ignore issues (N) related to gangs and the problem (C) just seems to be getting worse.*

Spanish Passage

5. *Muchos políticos ignoran los asuntos (N) relacionados con las pandillas y el problema (C) parece estar empeorando.*

The independent variables in this experiment were reading (first, second), passage language (English, Spanish), and target word type during the second reading (cognate, non-cognate). The dependent measures were overall passage reading time and fixation time on the target words.

During the second readings, overall reading times decreased for same-language texts and for translations, but novice bilinguals showed very little repetition benefit for English texts preceded by Spanish texts. That is, reading Spanish texts did not help novice bilinguals process the English translations. This implies that the text representations were not integrated at the meaning level because reading Spanish did not facilitate reading English. This conclusion was supported by the fixation time data for the cognate and non-cognate target words. For skilled bilinguals, fixation times on identical words (same language on each reading), cognates, and non-cognates were reduced during the second readings and there were no reliable differences in fixation times between cognates and non-cognates. This implies that repetition effects were based primarily on conceptual features, which indicates that the text representations were integrated at this level. The pattern was less clear for novice bilinguals. Fixation times on identical words, cognates, and non-cognates were slightly reduced during the second readings, but cognates were read faster than non-cognates when the target words were in Spanish on the second reading. This implies that novices benefitted from the repetition of surface level, lexical properties (i.e., spelling and sound), as well as conceptual properties when reading translations that were presented in their second, less fluent language. Other researchers have also found that non-proficient bilinguals emphasize surface features when reading in their second language (Koda, 1996; Nagy, García, Durgunoğlu, & Hancin-Bhatt, 1993).

Friesen and Jared (2007) conducted a similar study. Their basic research design was as follows. Participants read a text in one language (English or French) and then read the text again in the same language or the different language. Embedded in the passages were 20 cognate or non-cognate target words. Friesen and Jared also manipulated the degree of conceptual overlap between the first and second readings. Specifically, participants read a passage in one language and then read a second passage that was identical (same story, same language), a translation that included the target cognates (same story, cognates overlap across readings), a translation in which the target cognates were replaced by non-cognate synonyms (same story, no cognate overlap across readings), a different story that included the target cognates (different story, cognates overlap across readings), or a different story that did not include the target cognates (different story, no cognate overlap across readings).

The primary independent variables in this experiment were passage language during the second reading (same language, different language), story overlap (same story, different story), and target word overlap (cognates, non-cognate synonyms). The dependent measures were overall passage reading time and fixation time on the target words.

A preliminary analysis indicated similar rereading benefits for English and French (i.e., faster reading during the second readings); therefore, subsequent analyses focused on comparing reading times and fixation times between conditions during the second readings. Friesen and Jared (2007) found faster reading times during the second readings (i.e., larger transfer effects) when the stories were the same (meaning overlap) than when they were different. Interestingly, when the stories did not overlap, only less skilled readers benefitted from word-level repetition (i.e., shorter fixation times for repeated cognates during the second reading than for the synonyms during the second readings).

Taken together, Raney et al.'s (1996) and Friesen and Jared's (2007) studies support the conclusion that, for skilled bilinguals, text repetition effects are mediated primarily by conceptual properties, whereas for novice bilinguals, both conceptual and lexical properties, such as word type (cognate or non-cognate), mediate repetition effects. In essence, these studies support a model in which bilinguals' representations of texts are integrated or shared across languages, but the source of integration reflects different levels of representation as a function of proficiency.

These two studies can be used to illustrate several methodological issues that must be considered when performing cross-language repetition studies. In Raney et al.'s (1996) study, the English and Spanish passages were written using a similar number of content words, which were matched on average word frequency. Normative testing was performed to make sure that the original and paraphrased versions of each text were read at similar rates. To avoid bias associated with the development of translations, half the passages were initially written in English and then translated into Spanish, and half the passages were initially written in Spanish and then translated into English. Cognate and non-cognate target word pairs (e.g., *issues* and *asuntos*) were individually matched on word frequency and length when possible. Friesen and Jared (2007) used similar controls at the word level, but used a different method to ensure accurate translations. They initially translated a set of English passages into French. The French versions were then translated back into English by a new translator who had not seen the English passages. This ensured that translations matched the intended meanings. Without accurate translations, conceptual transfer could be greatly reduced due to lack of meaning overlap.

Code Switching and Language Mixing

Another method for exploring integration at the textbase or conceptual level is code switching and language mixing. Code switching can be described as an intentional (voluntary) alternation between languages during speech (Poplack & Meechan, 1998; see Gollan & Ferreira, 2009 for a discussion of voluntary and cued code switching). Language mixing generally refers to situations in which individuals are presented stimuli in two languages (as opposed to producing code switched language), such as being given a Spanish sentence to read that includes some English words.

A common finding is that there is a cost to switching languages within a sentence, both in production and comprehension (Gollan & Ferreira, 2009; Heredia & Altarriba, 2001). *Cost* is typically defined as slower response times and/or reduced comprehension. Given that there is often a cost to switching languages, it is reasonable to ask why bilinguals code switch. Resolving this issue is beyond the scope of this chapter, but one reason for switching languages is that a word needed to describe a concept is not readily available in the currently used language, so the speaker switches languages (Heredia & Altarriba, 2001). The important point here is that when language switches occur within sentences, the associated costs with switches can be used to examine language integration. Although there is strong evidence for a cost associated with switching languages within a sentence, recent research by Gullifer, Kroll, and Dussias (2013) indicates there is little or no cost to switching (alternating) languages between sentences (e.g., sentence 1 is Spanish, sentence 2 is English). They take this finding as further evidence that lexical access is nonselective in bilinguals.

One piece of evidence that within-sentence code switching reflects integration of languages at the textbase is that code switches tend to occur at phrase boundaries for the target language. Likewise, mixed-language sentences are generally easier to comprehend when language switches occur at normal phrase boundaries (Dussias, 2001). There are exceptions to this general trend. For example, in English and Spanish there is frequent mixing between the determiner and the noun, such as when switching from Spanish to English produces *la flag* instead of *the flag*. Mixing of determiners and nouns in this case might reflect the fact that the determiner is not contributing to the meaning of the phrase (Dussias, 2001). Another basic finding is that as proficiency in both languages increases, the degree of code switching also increases and the ease of processing mixed-language sentences increases (Miccio, Scheffner-Hammer, & Rodriguez, 2009). Interestingly, attention does not need to be focused on both languages to obtain effects of language mixing (Amrhein, 1999). These findings support the conclusion that integration of languages at the textbase or meaning level increases as proficiency increases.

Dussias (2001) reviews several studies that show how within-sentence code switching can be used to examine the difficulty of comprehending specific grammatical aspects of language and, therefore, the degree of integration across languages at the textbase level. Dussias (1997) provides a clear example of the methodology. The basic methodology was as follows: Spanish-English bilinguals read sentences in which the first part of the sentence was in one language (e.g., Spanish) and the second part was in the other language (English). Sample sentences are provided below with the point at which the language switches underlined.

6. *La maestra compró <u>los books</u> for the children.* [mismatch]
   The teacher bought the books for the children.
7. *La maestra compró <u>the books</u> for the children.* [match]
8. *The teacher bought <u>the libros</u> para los niños.* [mismatch]
9. *The teacher bought <u>los libros</u> para los niños.* [match]

Sentences 6, 7, 8, and 9 were presented as pairs on a computer monitor and participants indicated whether the first sentence was identical to the second sentence (i.e., a sentence-matching task). For example, Sentence 6 would be presented twice (*La maestra compró <u>los books</u> for the children/La maestra compró <u>los books</u> for the children*) and the participant should respond *same*. Reading time for each sentence was measured. An additional set of sentences pairs were presented that did not match, such as, *El señor se olvidó los books for his friends/El señor se olvidó los boots for his friends* (*books* becomes *boots*). The correct response here is *different*. The critical task appears to be a simple matching task for mixed-language sentences. Of importance is the location of the language switch and the effect on reading time for a sentence. In Sentences 6 and 8, the language of the determiner (*the*) and the noun (*books*) are mismatched. In Sentences 7 and 9, the language of the determiner and the noun are matched. Sentences were counterbalanced such that participants read only one version of each sentence (e.g., a participant who read Sentence 6 did not read Sentences 7, 8, or 9). Participants also read a set of filler passages that contained language switches in other locations than between a determiner and a noun to mask the common location of the code switch in the experimental sentences.

The key independent variables were initial language of a sentence (Spanish, English) and whether the language of the determiner and noun matched (match, mismatch). Reading times for *different* sentence pairs were not analyzed; therefore, *same/different* was not included as an independent variable. The dependent variable was reading times for the sentences.

Dussias found that reading times for Sentence 7 (match) were actually longer than for Sentence 6 (mismatch). Finding longer reading times when the language of the determiner and noun match (Sentence 7) relative to the mismatch (Sentence 6) at first seems odd. Dussias suggests that the proficient Spanish readers use the determiner to gain information about the upcoming noun; therefore, *los* in Sentence 6 provides information that is not provided by *the* in Sentence 7 (e.g., gender). This facilitates reading time. Dussias found no difference in reading times for Sentences 8 (mismatch) and 9 (match). Thus, when switching to the dominant language, whether or not the language of the determiner and noun matched did not influence processing time. Asymmetries in switch cost have been reported in several other studies involving language-based code switching tasks (e.g., Grosjean, 1988; Myers-Scotton, 1995, 1997). The asymmetry in switch costs is a hallmark of switching from the dominant language to the less-dominant language and is thought to also reflect the need for greater inhibition of the dominant language when switching to the less-dominant language (Costa & Santesteban, 2004).

The sensitivity of the code-switching/language-mixing task to grammatical boundaries indicates that the readers formed representations at the textbase level that were based on grammatical properties of both languages. That is, the representations were integrated at the textbase level. The asymmetry in switch costs demonstrates that the lexical features of the dominant language are more active than the subordinate language.

Another clear example of the mixed-language methodology is Altarriba, Kroll, Sholl, and Rayner (1996). Here, we describe their first experiment. The basic design was as follows. Proficient bilingual readers were presented with sentences containing English words only, or English sentences that contained one Spanish word. Participants read the sentences while their eye movements were measured. Two sample sentences are presented below.

High Constraint

10. He wanted to deposit all of his *money/dinero* at the credit union.

Low Constraint

11. He always placed all of his *money/dinero* on a silver dish on his dresser.

Participants read each sentence with the English (*money*) or Spanish (*dinero*) target word. Target words were either high or low frequency (defined as frequency of use in English). The sentences also varied in the degree of constraint for the given target word. For example, the context in Sentence 10 constrains the range of words that could fill the target word slot to a greater degree than does Sentence 11. In other words, the target word is more predictable in the high-constraint sentences. In this experiment, the independent variables were sentence constraint (high, low), target word frequency (low, high), and target word language (English, Spanish). The dependent variable was fixation times for the target words. Target words did not appear in italics.

Altarriba et al. (1996) found that average fixation times were shorter for English target words than for Spanish target words. When the English target words were examined, they found clear effects for word frequency (longer fixation time for low frequency words) and sentence constraint (longer fixation time for low constraint sentences). When the Spanish target words were examined, they again found word frequency effects (longer fixation time for low frequency words), but word frequency interacted with sentence constraint for first fixation duration (the initial fixation on a word). For low frequency words, fixation time was less for low-constraint sentences, whereas for high frequency words, fixation time was less for high-constraint sentences. This interaction was not found for gaze duration (sum of all fixations made on a word before moving to a new word).

Altarriba et al. (1996) interpreted the interaction found for first fixation duration as demonstrating that sentence contexts restricted the range of expected semantic features for upcoming words (i.e., what words will fit the context) *as well as* lexical features (i.e., what language will be accessed). These bilingual participants slowed down when reading Spanish target words in high-constraint English contexts because the Spanish words did not match the expected lexical features despite being conceptually appropriate. The findings that sentence constraint influences activation of conceptual and lexical features of upcoming words (i.e., word meaning and orthography) implies that the languages are integrated at these levels.

Altarriba et al.'s (1996) study demonstrates several important methodological controls. First, Spanish words were presented in Spanish orthography with required accent marks. Second, target words were counterbalanced across sentences such that participants read only one sentence containing a critical English/Spanish word pair (i.e., participants who read Sentence 10 did not read Sentence 11). Third, sentence order was counterbalanced across participants so that for half of the participants an English sentence was followed by a Spanish sentence, and the reverse was true for the other half of the participants. Fourth, participants were presented with comprehension questions to ensure they attended to the reading task. Fifth, the target words were positioned near the middle of each sentence. In this type of study, target words should not be the last words of sentences because reading times would be inflated due to natural sentence wrap-up effects (readers typically read the last word of a sentence slower than preceding words). Fifth, target words were always congruous with the sentence contexts, regardless of language. This is necessary to ensure that interpretation of target words does not change as a function of sentence endings.

Background Knowledge and Schemas

The research methodologies we have described so far have been aimed at exploring representations that reflect the content of texts, that is, the surface form and textbase. Some of the studies reviewed do have implications for higher-level representations, such as the situation model or pragmatic level. For example, the transfer studies by Raney et al. (1996) and Friesen and Jared (2007) support the conclusion that for non-proficient bilinguals (e.g., novice language learners), knowledge obtained from reading in L2 might not be available to guide processing when reading L1. Another way of stating this is that a reader does not always develop a situation model after reading an L2 text that can be used to guide processing when reading the L1 text.

The influence of higher-level representations on bilingual reading is an area of research that deserves more attention. One topic that has been explored is the role of background knowledge and schemas on second language reading (Carrell, 1983; Lee, 1986; Nassaji, 2002). We combine these together as a single topic because, from a methodological perspective, research on background knowledge and schemas both explore *what the reader brings to the text* and how this influences processing. Furthermore, background knowledge is sometimes described as a component of schemas (Rumelhart, 1980).

After an initial spurt of studies (e.g., Bernhardt, 1986; Carrell, 1983; Horiba, van den Broek, & Fletcher, 1993; Lee, 1986; McLeod & McLaughlin, 1986; Roller & Matambo, 1992), research in this area has waned, possibly due to the theoretical difficulty of describing the components of situation models, pragmatic representations, communicative contexts, or other high-level representations. We believe this is a loss from a theoretical perspective because understanding how bilingual and non-bilingual readers use their prior knowledge to comprehend texts in their L1 and L2 is important. Because our purpose is to illustrate research methodologies, not resolve theoretical difficulties, we describe one line of research to illustrate how research has attempted to explore high-level text representations.

Carrell (1983), Lee (1986), and Roller and Matambo (1992) each explored the role of background knowledge on text processing in L1 and L2. The basic research design was as follows: bilingual participants read a set of texts that varied in familiarity, the presence/absence of an interpretive context, and transparency (i.e., clarity of content). They each did this by presenting participants with texts that were easy to comprehend with a title and difficult to comprehend without a title (adapted from Bransford & Johnson, 1972) in their L1 or L2. Daniel and Raney (2007) have shown that removing the title from these types of texts has little or no impact on comprehension of the surface form, slightly decreases comprehension of the textbase, and substantially decreases situation model level comprehension. This demonstrates that the inclusion/exclusion of the title is an effective technique for manipulating the formation of a coherent situation model when reading vague texts such as those designed by Bransford and Johnson.

In the initial study, Carrell (1983) defined familiarity based on knowledge and experience with the content of a passage. Specifically, there was a passage about washing clothes (familiar) and a passage about serenading someone in a tall building by playing a guitar with a speaker attached to balloons that floated up to the individual in the building (something that was not likely to have been experienced). The presence/absence of a context was manipulated by including a title with a picture illustrating the scene described by the text (context present) or not including the title and picture (context absent). A transparent and non-transparent version of each passage was formed by including concrete or vague words when referring to objects or events (e.g., *music* vs. *sound*, respectively). Carrell had intermediate-level bilinguals read the clothes washing and serenading passages in one of the eight conditions in either their L1 or L2 and then recall the passages in their L2. Thus, the independent variables in this experiment were passage language (L1, L2) familiarity (familiar, unfamiliar), context (present, absent), and transparency (transparent, not transparent). The dependent variable was the number of idea units (meaning units) recalled. Carrell found that familiarity, context, and transparency influenced the number of idea units recalled for passages read in L1 but not in L2. That is, readers were not using their prior knowledge or the context to help them comprehend the texts when reading in their L2.

Lee (1986) repeated Carrell's (1983) study with an important methodological difference: Lee had participants recall the texts in their L1, whereas Carrell had participants recall the texts in their L2. Lee found that all three variables influenced recall. Thus, when using one's L1 to recall a text read in L2, familiarity, context, and transparency influenced recall. Lee suggested that recalling the text in their L2 was too difficult in Carrell's study and that it eliminated the effects of familiarity, context, and transparency. Roller and Matambo (1992) repeated Lee's (1986) procedure but used advanced bilinguals. They reasoned that highly proficient bilinguals should be influenced by familiarity, context, and transparency in each of their languages. They found a complex pattern of results, including no effect of context on recall for both L1 and L2, but for our purpose the key point is they did not replicate Carrell's or Lee's findings.

The studies by Carrell (1983), Lee (1986), and Roller and Matambo (1992) produce a seemingly incoherent set of results. However, one finding was relatively consistent across all three studies: When reading in one's L2, recall of the unfamiliar passages actually exceeded recall of the familiar passages. Why might this be? Roller and Matambo (1992) and Nassaji (2002) suggested this could have reflected how the passages were written, as well as how the passages were comprehended. Two points are relevant. First, the unfamiliar passages seemed to contain more concrete words than the familiar passages. This would make the textbase easier to construct in the unfamiliar passage. Second, when reading in one's L2, less emphasis might have been placed on situation-level comprehension and more emphasis might have been placed on textbase-level comprehension. In essence, the construction of the texts might have facilitated textbase comprehension, and reading in L2 might have both decreased situation model comprehension and emphasized textbase comprehension. This led to enhanced recall of the actual texts when they were unfamiliar.

These studies provide several methodological lessons. First, comprehension will differ as a function of language proficiency, but task demands can lead to a nonlinear relation between proficiency and comprehension (i.e., better proficiency does not guarantee better comprehension). Second, experimental texts must be precisely matched. Equating passages on the number of words alone is not sufficient; passages must be equated on the type of words (e.g., number of concrete words, number of high frequency words) as well as content. Third, seemingly small changes in procedures can have large changes on outcomes. For example, recalling a passage in L1 might increase the amount recalled as well as the type of information recalled. Fourth, reading in L1 or L2 does not merely influence the difficulty of reading the text; this manipulation can also alter reading strategies, such as focusing less on the situation model and more on the textbase when reading in L2 relative to L1. These three studies clearly demonstrate the complexity of conducting research on bilingual text comprehension.

Summary and Conclusions

Overall, our purpose in this chapter was to describe research methodologies that can be used to examine how and when bilingual readers integrate their representations of texts across languages. We approached the review by describing methods based on what level of text representation was examined. For the surface level, we described studies in which the script itself was manipulated (Akamatsu, 2003; Koda, 1990; Muljani et al., 1998) and studies in which cognates and non-cognates were examined (Libben & Titone, 2009; Raney et al., 1996; Van Assche et al., 2009). At the level of the textbase, we described studies in which texts were read twice and repetition effects were examined (Friesen & Jared, 2007; Raney et al., 1996) and studies in which the language was mixed within a text (Altarriba et al., 1996; Dussias, 1997). At the level of the situation model, we described studies in which schemas and background knowledge were examined (Carrell, 1983; Lee, 1986; Nassaji, 2002). Although the levels of representation distinction was convenient for reviewing research methodologies, it is important to keep in mind that text representations cannot always be separated into discrete levels. By necessity, representations must overlap. For instance, a reader must comprehend the wording (surface form) to build the textbase, and the textbase must be understood to build a complete and coherent situation model.

We provided several example studies to illustrate common methods used and provided details about what parts of a procedure need to be carefully controlled (e.g., word length and frequency, passage order, readers' language backgrounds, how a text is presented on a display). We also attempted to show the wide range of variables manipulated, such as proficiency level, orthography, text language, task demands, and semantic context. Common dependent measures presented included sentence reading times, eye movement measures (e.g., first fixation duration, total gaze duration) comprehension accuracy, and reaction times. The materials and apparatus used in the research were as simple as a printed booklet and as complicated as an eye tracking equipment.

We mentioned that many of these methods are not uniquely “bilingual” research methods, but rather methods that have been used frequently in other studies of language processing. For instance, lexical decision tasks are commonly used in “monolingual” studies. Case alternation and variation of letter pattern familiarity are manipulations that can be performed within a single language and used to answer questions about differences in reading processing between good and poor readers. Although cognate words can only be recognized by bilinguals as cognates, they can be included in single language experiments (i.e., bilinguals reading one language) and dual-language experiments (i.e., bilinguals reading in two languages). One can think of cognates as analogous to synonyms in paraphrased texts in monolingual studies. Text repetition has been used to study higher levels of representation and to examine what is transferred from one reading to the next in the same language.

The final point we want to make is that studying bilingual text representation provides insight not just about how bilinguals represent text, but how texts are processed and comprehended in general. Conducting research using bilingual populations also provides methodological opportunities that do not exist for non-bilingual populations, such as comparing cognates and non-cognates. These opportunities require researchers to be especially attentive to the methodologies employed.

List of Keywords

Bilingual Interactive Activation Plus Model (BIA+), Case alternation, Code switching, Cognate facilitation effect, Cognates, Communicative context, Communicative model, Concepts of surface form, Conceptual-based level, Eye movements, First fixation duration, Gaze duration, Higher-level representations, Integrated representations, Interlingual homographs, Interpretive context, Language integration, Language mixing, Language switching task, Lexical decision task, Lexical level, Meaning-based level, Mixed-language, Multiple levels of representation, Neighborhood density, Orthographic activation, Orthographical knowledge, Phonological encoding, Pragmatic-Communicative Model, Repetition effects, Revised Hierarchical Model (RHM), Schemas, Scripts, Situation Model, Surface level, Switch cost, Syntactic analysis, Task switching paradigm, Text repetition effect, Textbase, Transfer benefit, Transparency, Word Association Model, Word frequency, Word recognition, Word-level repetition, Wrap-up effects.

Review Questions

1. What factors (both for the text and the bilingual) might influence whether a cognate has a facilitative effect during reading?
2. Assume a group of native English speakers who are learning Spanish read two texts, one in each language. They are then given a comprehension quiz that tests

their knowledge of the surface form, textbase, and situation model for each text. How might their comprehension vary across text as a function of language and level of comprehension? Which level of representation is most likely to be integrated across languages for these speakers?

3. We suggested there are no uniquely *bilingual research methodologies* and that research methodologies for exploring cognitive processes can be more or less easily applied to study bilingualism. What types of research methodologies are easily adapted to study bilingual cognitive processes? Do you think there are research methodologies that cannot be adapted to study bilingual cognitive processes?

Suggested Student Research Projects

1. Design a study in which you can test whether cognates in a bilingual's lexicon are integrated cross-linguistically. Possible methodological considerations include type of phonological similarity and word frequency.
2. Design a study to determine if the time needed to access a word from memory (i.e., the lexicon) is different for proficient bilinguals, non-proficient bilinguals (people who know a second language but are not highly proficient), and non-bilinguals (people with very little experience with a second language).
3. Design a study to determine whether code switching slows reading time more when switching from one's first language to one's second language, or vice versa.
4. Design a study to determine whether reading a text in one language (e.g., Spanish) and then reading the same text in a second language (e.g., English) is more beneficial for texts in which the reader has strong background knowledge or weak background knowledge.

Related Internet Sites

Bilingual Memory Models: http://en.wikipedia.org/wiki/Bilingual_memory

Cross-Linguistic Easy-Access Resources: http://clearpond.northwestern.edu/

Corpus of Contemporary American English (COCA): http://www.wordfrequency.info/

Suggested Further Reading

Bovee, J.C., & Raney, G.E. (2015). Evaluating missing letter effects and comprehension in proficient and non-proficient languages. *Applied Psycholinguistics*: doi: http://dx.doi.org/10.1017/S0142716414000563.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 797–810.

Lim, J. H., & Christianson, K. (2013). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition, 16,* 518–537.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visual Experimentation, 83,* e50780, doi:10.3791/50780.

References

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning, 53*, 207–231.

Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. *Advances in Psychology, 83*, 157–174.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*, 477–492.

Amrhein, P. C. (1999). On the functional equivalence of monolinguals and bilinguals in "monolingual mode": The bilingual anticipation effect in picture-word processing. *Psychological Science, 10*, 230–236.

Bernhardt, E. B. (1986). Cognitive processes in L2: An examination of reading behaviors. In J. Lantolf & L. Labarca (Eds.), *Research in second language learning: Focus on the classroom* (pp. 35–51). Norwood, NJ: Ablex.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717–726.

Carrell, P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning, 33*, 183–203.

Chen, H. C., & Ng, M. L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition, 17*, 454–462.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1283–1296.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491–511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1057–1074.

Daniel, F., & Raney, G. E. (2007). Capturing the effect of a title on multiple levels of comprehension. *Behavior Research Methods, 39*, 892–900.

De Groot, A. M. B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Advances in psychology: Orthography, phonology, morphology, and meaning* (pp. 389–412). Amsterdam: Elsevier.

De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 27–51). Amsterdam: John Benjamins.

De Groot, A., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90–123.

Dijkstra, A. F. J., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition, 5*(3), 175–197.

Dussias, P. E. (1997). Sentence matching and the functional head constraint in Spanish/English codeswitching. *Spanish Applied Linguistics, 1*, 114–150.

Dussias, P. E. (2001). Psycholinguistic complexity in codeswitching. *International Journal of Bilingualism, 5*, 87–100.

Ervin, S., & Osgood, C. (1954). Psycholinguistics: A survey of theory and research problems. In C. Osgood & T. Sebeok (Eds.), *Psycholinguistics* (pp. 139–146). Baltimore, MA: Waverly Press.

Fletcher, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 589–607). San Diego: Academic Press.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. *Discourse Processes, 13*, 175–190.

Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin, 125*, 193–232.

Friesen, D. C., & Jared, D. (2007). Cross-language message-and word-level transfer effects in bilingual text processing. *Memory & Cognition, 35*, 1542–1556.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 640–665.

Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of discourse comprehension. *Topics in Cognitive Science, 3*, 371–398.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual review of psychology* (pp. 163–189). Palo Alto, CA: Annual Reviews.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3*, 233–274.

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology, 4*, 1–13.

Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*, 164–168.

Horiba, Y., van den Broek, P. W., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure-preserving top-down processing. *Language Learning, 43*, 345–372.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., et al. (2010). Control and interference in task switching: A review. *Psychological Bulletin, 136*, 849–874.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363–394.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29*, 133–159.

Koda, K. (1990). The use of L1 reading strategies in L2 reading. *Studies in Second Language Acquisition, 12*, 393–410.

Koda, K. (1996). L2 word recognition research: A critical review. *Modern Language Journal, 80*, 450–460.

Kroll, J. F., & De Groot, A. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Mahwah, NJ: Lawrence Erlbaum.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism, 13*, 373–381.

Lee, J. F. (1986). Background knowledge and L2 reading. *The Modern Language Journal, 70*, 350–354.

Levy, B. A. (2001). Text processing: Memory representations mediate fluent reading. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger III (Eds.), *Perspectives on human memory and cognitive aging: Essay in honour of Fergus Craik* (pp. 83–98). Philadelphia: Psychology Press.

Levy, B. A., Barnes, L., & Martin, L. (1993). Transfer of fluency across repetitions and across texts. *Canadian Journal of Experimental Psychology, 47*, 401–427.

Levy, B. A., Campsall, J., Browne, J., Cooper, D., Waterhouse, C., & Wilson, C. (1995). Reading fluency: Episodic integration across texts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1169–1185.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 381–390.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automatization? Reading in a second language. *Language Learning, 36*, 109–127.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25–40.

Miccio, A. W., Scheffner-Hammer, C., & Rodriguez, B. (2009). Code-switching and language disorders in bilingual children. In B. E. Bullock & J. A. Toribio (Eds.), *Cambridge handbook of linguistic code-switching* (pp. 259–270). Cambridge, England: Cambridge University Press.

Muljani, D., Koda, K., & Moates, D. R. (1998). The development of word recognition in a second language. *Applied Psycholinguistics, 19*, 99–113.

Myers-Scotton, C. (1995). A lexically based model of codeswitching. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages: Cross-disciplinary perspectives on codeswitching* (pp. 233–256). Cambridge, England: Cambridge University Press.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching* (2nd ed.). Oxford: Oxford University Press.

Nagy, W. E., García, G. E., Durgunoğlu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Literacy Research, 25*(3), 241–259.

Nassaji, H. (2002). Schema theory and knowledge-based processes in second language reading comprehension: A need for alternative perspectives. *Language Learning, 52*, 439–481.

O'Brien, E. J., Raney, G. E., Albrecht, J. E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes, 23*, 1–24.

Philipp, A. M., & Koch, I. (2009). Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1187–1195.

Poplack, S., & Meechan, M. (1998). Introduction: How languages fit together in codemixing. In S. Poplack & M. Meechan (Eds.), Instant loans, easy conditions: The productivity of bilingual borrowing. *International Journal of Bilingualism, 2 [special issue]*, (pp. 127–138). London: Kingston Press.

Potter, M. C., So, K. F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23–38.

Prior, A., & MacWhiney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition, 13*, 253–262.

Raney, G. E. (2003). A context-dependent representation model for explaining text repetition effects. *Psychonomic Bulletin & Review, 10*, 15–28.

Raney, G. E., Atilano, R., & Gomez, L. (1996, November). Language representation and bilingual reading. Presented at the 37th annual meeting of the Psychonomic Society, Chicago, IL.

Raney, G. E., Obeidallah, S., & Miura, T. (2002). Text comprehension in bilinguals: Integrating perspectives on language representation and text processing. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 165–186). Amsterdam: Elsevier.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49*, 151–172.

Raney, G. E., Therriault, D., & Minkoff, S. (2000). Repetition effects from paraphrased text: Evidence for an integrative model of text representation. *Discourse Processes, 28*, 61–81.

Roller, C. A., & Matambo, A. R. (1992). Bilingual readers' use of background knowledge in learning from text. *TESOL Quarterly, 26*, 129–141.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. E. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Lawrence Erlbaum.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197–212.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science, 20*, 923–927.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Van Hell, J. G., & De Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*, 193–211.

Van Hell, J. G., & de Groot, A. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*(3), 431–451.

Weinreich, U. (1953). *Languages in contact*. New York: The Linguistic Circle of New York.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.

Chapter 8
Eye Movement Methods to Investigate Bilingual Reading

Veronica Whitford, Irina Pivneva, and Debra Titone

Abstract This chapter provides a general overview of the use of eye movement recordings to investigate the cognitive processes that underlie natural reading, including first- (L1) and second-language (L2) reading in bilinguals. We focus on two important issues arising from bilinguals' divided L1/L2 knowledge and use: cross-language activation (i.e., nonselective activation of both L1 and L2 lexical representations) and reduced lexical entrenchment (i.e., delayed lexical access resulting from lower baseline activation levels of L1 and L2 words and/or weakened L1 and L2 lexical memory representations). Prior work has used eye movement recordings to independently examine these two issues; however, in a reanalysis of recent work from our laboratory (Whitford and Titone, Psychon Bull Rev 19:73–80, 2012), we examine their joint impact on eye movement measures of bilingual reading. We find that cross-language activation and reduced lexical entrenchment mutually constrain L1 and L2 reading in bilinguals, which suggests that they may be two sides of the same coin.

Introduction

Reading is an essential life skill developed over many years of formal instruction and practice, and is arguably one of humankind's greatest achievements. Indeed, unlike spoken language processing, which evolved over several thousands of years, written language processing, which has only existed for a few thousand years, is a relatively recent development in the history of humankind (e.g., Immordino-Yang & Deacon, 2007). The scientific study of reading through the use of eye movement recordings first emerged in the late 1800s (e.g., Huey, 1908), and has since culminated in an extensive body of research (reviewed in Rayner, 1998, 2009; Rayner,

V. Whitford (✉) • I. Pivneva, Ph.D. • D. Titone, Ph.D.
Department of Psychology, Centre for Research on Brain, Language and Music, McGill University, Stewart Biology Building, 1205 Dr. Penfield Avenue, Montreal, QC, Canada, H3A 1B1
e-mail: veronica.whitford@mail.mcgill.ca; irina.pivneva@mail.mcgill.ca; dtitone@psych.mcgill.ca

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_8

Pollatsek, Ashby, & Clifton, 2012). However, from that point onward, the primary focus of this research has been on the cognitive processes underlying reading performance in English monolinguals, who represent a relatively small percentage of the world's population. More recently, studies have begun to use eye movements to examine the cognitive processes underlying reading performance in bilinguals, who represent a relatively larger percentage of the world's population. In fact, current estimates suggest that more than half of the world's population is bilingual, if not multilingual (Grosjean, 2010).

Bilinguals differ from monolinguals in a number of important ways. Perhaps the most crucial difference is that relative to monolinguals, bilinguals, by virtue of knowing and using both first (L1) and second language (L2), necessarily have divided L1/L2 knowledge and use. Consequences of this divided L1/L2 knowledge and use include *cross-language activation*, that is, simultaneous and nonselective activation of both L1 and L2 lexical representations and reduced *lexical entrenchment,* that is, delayed lexical access resulting from reduced L1 and L2 baseline activation levels and/or weakened L1 and L2 lexical memory representations. Although few in number, eye movement studies examining bilingual reading have *independently* tapped into these two key issues. In this chapter, we review this literature and present a reanalysis of recent work from our laboratory, demonstrating that both cross-language activation and lexical entrenchment *mutually constrain* bilingual reading—ultimately demonstrating that they are two sides of the same coin with respect to natural bilingual reading.

Eye Movement Reading Research

Reading involves the integration of many complex cognitive and oculomotor processes, which ultimately result in word recognition and comprehension. At its most basic level, reading requires a series of eye movements called *saccades*, separated by brief pauses called *fixations*, which direct printed information onto the fovea (i.e., the central region of the retina with highest visual acuity) for detailed linguistic processing (see, for example, Liversedge, Gilchrist, & Everling, 2011; Radach & Kennedy, 2013; Rayner, 1997, 1998, 2009; Rayner et al., 2012). This linguistic processing is hierarchical in nature: sublexical processing (i.e., lower-level analysis of a word's orthographic, phonological, and morphological properties) precedes supralexical processing (i.e., higher-level analysis of a word's semantic and syntactic properties within a sentence), which, in turn, precedes postlexical processing (e.g., discourse-level semantic integration). In addition, some low-level processing (e.g., word length, word shape, letter features) occurs beyond the fovea, that is, in the parafoveal region (i.e., 2–5° of visual angle from fixation). This area of effective vision from which useful information can be extracted during reading is called the *perceptual span* (for a recent review, see Schotter, Angele, & Rayner, 2012). The perceptual span extends 3–4 characters to the left and 14–15 characters to the right of fixation in skilled readers of left-to-right orthographies, such as English (e.g., McConkie & Rayner, 1975, 1976; Rayner & Bertera, 1979; Rayner, Well, & Pollatsek, 1980), and is functionally reversed in skilled readers of right-to-left orthographies, such as Arabic, Hebrew, and Urdu (Jordan et al., 2014; Paterson et al., 2014; Pollatsek, Bolozky, Well, & Rayner, 1981). The perceptual span mediates eye movement control during reading, especially in the selection of upcoming saccadic targets, and trades-off with the ease of textual processing: more parafoveal information is extracted from the right of fixation when the text is easier to process.

Eye tracking technology, also referred to as eye movement recordings, has played a crucial role in elucidating the cognitive and oculomotor processes implicated in reading. In this method, a camera monitors the saccades and fixations made by one or both eyes as participants read text presented on a computer screen. Most contemporary eye trackers do this by generating corneal reflections through infrared pupil illumination; the vector between the corneal reflection and the pupil's center is used to assess the eye's position on the screen (Hansen & Ji, 2010). Calibration and validation procedures are routinely performed before and during the experiment to ensure tracking accuracy. A sampling rate of 1000 Hz (or one time every millisecond) is typically used to capture the eye movements characteristic of reading.

There are a number of advantages and disadvantages associated with the use of eye movement recordings of reading. The advantages of an eye movement approach include increased ecological validity—participants read text in a relatively naturalistic context. This contrasts with response-based tasks, such as lexical decision (i.e., classifying stimuli as words or non-words as quickly and accurately as possible) and progressive demasking (i.e., identifying visually degraded stimuli as words as quickly as possible), which use decontextualized stimuli (e.g., single words presented in isolation), and require overt decisions that are not normally part of natural language comprehension (potentially resulting in dual-task situations). Moreover, eye movement recordings allow for a temporally precise measure of the cognitive processes implicated in reading. For example, eye movement recordings can capture differences in early- vs. late-lexical processing stages (e.g., Rayner, 1998, 2009), which contrasts with many response-based tasks. The disadvantages of an eye movement approach include somewhat cumbersome equipment that is difficult to transport. This contrasts with response-based tasks which can be run using a laptop and button-box. Moreover, the data generated by eye movement recordings, although rich in nature (see below for further details), require an extensive amount of pre-processing and analysis. Finally, the eye movement approach largely assumes that what happens at a fixation reflects attention that is directed at that fixation (i.e., the *Eye-Mind Hypothesis*; Just & Carpenter, 1980). However, as previously mentioned, attention may not necessarily be directly linked to where the eyes are fixated, which may have implications for how we interpret different data patterns. These and other possible disadvantages notwithstanding, we strongly believe that the advantages associated with eye movement recordings far outweigh the disadvantages. Without this method, a deeper understanding of the moment-to-moment processes implicated in reading would be impossible.

A wealth of eye movement reading measures can be extracted from the eye movement record. These measures can largely be divided into global measures of processing difficulty across an entire sentence/passage of text and local measures of processing difficulty for specific words within a sentence/passage of text. Global measures of reading performance include average reading rate (normally 250–300 words/min), average saccade length (normally 7–9 characters in length), average fixation duration (normally 200–250 ms), percent regressive saccades (backward saccades reflecting re-reading, normally 10–15 % of all saccades), and total reading time (in ms), again, across an entire sentence/passage of text (Rayner, 1998, 2009). Such measures are modulated by text difficulty, participants' reading ability, and participants' reading goal (e.g., thorough reading for comprehension vs. scanning for key words). Specifically, increased text difficulty, reduced reading ability, and more thorough reading generally result in slower reading rates, shorter saccade lengths, longer fixation durations, more regressions, and longer total reading times (Rayner, 1998, 2009).

Local measures of reading performance can be divided into those reflecting early-stage lexical processing (i.e., the earliest stages of lexical access) vs. those reflecting late-stage, post-lexical processing (i.e., semantic integration, revision). Early-stage measures include first fixation duration (duration of the very first fixation on a word), single fixation duration (duration of the only fixation made on a word), gaze duration (sum of all fixations made on a word on the first pass), and skipping (probability of fixating a word on the first pass). Late-stage measures include go-past time (the sum of all fixation durations on a word starting from the first fixation duration until a rightward saccade is made past the word), percent regressions (backward saccades), and total reading time (sum of all fixation durations) of a word (Rayner, 1998, 2009). Like global measures of reading performance, local measures of reading performance are also modulated by text difficulty, participants' reading ability, and participants' reading goal; however, they are also modulated by the linguistic properties of a word (e.g., length, frequency, contextual predictability). Specifically, longer words (e.g., *CONSCIENTIOUSNESS* vs. *BLISS*), less frequent words (e.g., *FERN* vs. *HOME*), and less predictable words (e.g., *Every morning, Mary drinks a cup of <u>wine</u>* vs. *Every morning, Mary drinks a cup of <u>tea</u>*) generally have longer fixation durations, reduced skipping rates, more regressions, and longer total reading times. Of note, the estimates provided above are based on monolingual English readers; however, reading behavior in bilinguals should, in principle, also be modulated by text difficulty, reading ability, and reading goal. Because bilinguals generally have reduced L2 vs. L1 exposure, and consequently, reduced L2 vs. L1 reading ability, bilinguals generally exhibit both reduced global and local measures of L2 vs. L1 reading performance (see Whitford & Titone, 2012).

Experiments using both sentence-level and paragraph-level materials can be programmed using a number of software packages; however, because the eye tracking system used in our laboratory is the EyeLink 1000 (see "Related Internet Sites" section), we have used Experiment Builder software (i.e., SR-Research's default software package), due to its relatively straightforward drag-and-drop procedures.

A sample experimental script for paragraph-level materials is provided in the Appendix. We also use open-source EyeTrack software developed at the University of Massachusetts Amherst (see "Related Internet Sites" section) to program sentence-level experiments, which, of note, can also be used to run paragraph-level experiments. An advantage of using EyeTrack software is that stimuli created within text files and Microsoft Excel or OpenOffice spreadsheets can be easily transformed into items that EyeTrack can read.

Perhaps the most important component in experimental design is generating well-controlled language materials. When designing sentence-level materials, it is crucial to ensure that experimental target words are matched to control words on key linguistic variables known to affect word processing (e.g., word length, frequency, contextual predictability, to name just a few), and also that experimental sentences are matched to control sentences on overall sentence structure and length (both of the target item and also potentially the words preceding and following the target item, given the potential for parafoveal preview). WordGen (see "Related Internet Sites" section) is a convenient tool for generating bilingual stimuli (target and control words) in English, French, Dutch, and German (Duyck, Desmet, Verbeke, & Brysbaert, 2004). When designing paragraph-level materials, which are inherently more naturalistic and complex in nature (i.e., less amenable to experimental control), it is crucial to code the constituent words on a word-by-word basis for key linguistic variables (e.g., word length, frequency, contextual predictability).

Cross-Language Activation and Lexical Entrenchment in Bilingual Reading

Now that we have covered some background information on eye movement reading research, let us turn to two key issues within bilingual reading: cross-language activation (i.e., simultaneous and nonselective activation of both L1 and L2 lexical representations) and reduced lexical entrenchment (i.e., delayed lexical access resulting from reduced L1 and L2 baseline activation levels and/or weakened L1 and L2 lexical memory representations). As previously mentioned, eye movement studies examining bilingual reading have *independently* examined these issues. We first review the literature on cross-language activation, followed by the literature on lexical entrenchment, and lastly, present a reanalysis of recent work from our laboratory demonstrating that both cross-language activation and lexical entrenchment *mutually constrain* bilingual reading. The motivation behind this reanalysis is a recent response-based study (i.e., progressive demasking) by Diependaele, Lemhöfer, and Brysbaert (2013), which casts cross-language activation and lexical entrenchment as distinct mechanisms that independently modulate word recognition. It is our hypothesis that cross-language activation and lexical entrenchment may not be as theoretically and empirically distinct as Diependaele et al. (2013) maintain.

Cross-Language Activation

Cross-language activation involves the simultaneous and automatic activation of both L1 and L2 lexical representations (including meaning) of lexically ambiguous words (reviewed in Van Assche, Duyck, and Hartsuiker, 2012). For example, when an English–French bilingual reads the isolated word *PIANO* (i.e., a musical instrument in English and French) in their L2 (French), they will automatically activate both the L2 (French) and L1 (English) meaning. This finding, referred to as *nonselective access,* lends support for an integrated bilingual lexicon, where both languages share a common memory store. Although the above example involves nonselective access of a word presented in isolation, it also occurs for words embedded in sentences. In fact, nonselective access can be heightened or attenuated as a function of contextual constraint, among other factors, such as reading time-course (see below).

Nonselective lexical access is measured by examining how fast bilinguals read words that share lexical characteristics across languages (e.g., cognates, interlingual homographs, and cross-language orthographic neighbors) vs. language-specific words (e.g., *HOUSE* in English). Specifically, cognates are words that share orthography (written form) and semantics (meaning) across languages (e.g., *PIANO* is a musical instrument in both English and French); interlingual homographs are words that share orthography, but not semantics across languages (e.g., *CHAT* is a *conversation* in English vs. a *cat* in French); and cross-language orthographic neighbors are the number of nontarget-language words that differ from a target word by a single letter only (e.g., the French word *feu* "fire" has the following English neighbors: *fee, fen, few, fez, flu*; Coltheart, Davelaar, Jonasson, & Besner, 1977).

Although cognates, interlingual homographs, and cross-language orthographic neighbors are all ways of examining cross-language activation, most eye movement studies have employed cognates and interlingual homographs. These studies have shown that bilinguals process cognates and interlingual homographs differently from language-unique words, partially because of the nature of the cross-linguistic overlap (Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Libben & Titone, 2009; Titone, Libben, Mercier, Whitford, & Pivneva, 2011; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Assche et al., 2012; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). In particular, bilinguals read cognates embedded in sentences faster than language-specific control words. The difference in processing time between cognates and language-unique words is known as *cognate facilitation.* Cognate facilitation can be observed during the earliest stages of reading, and is modulated by the amount of orthographic overlap between the two languages (e.g., Van Assche et al., 2011). Conversely, bilinguals generally read interlingual homographs embedded in sentences slower than language-unique control words. The difference in processing time between interlingual homographs and language-specific words is known as *interlingual homograph interference*. Interlingual homograph interference can also be observed during the earliest stages of reading, and is also modulated by the amount of orthographic overlap between the two languages.

Researchers use the magnitude of the cognate facilitation and interlingual homograph interference effect to index the amount of nonselective access. Cognate facilitation and interlingual homograph interference often persist throughout early- and late-stage reading, but vary as a function of sentential constraint (e.g., Libben & Titone, 2009).

By embedding cognates and interlingual homographs in sentences, researchers can use contextual constraint to bias the L1 vs. L2 meaning of these lexically ambiguous words. For example, the sentence *Since they liked to gossip, they had an extended CHAT that lasted all night* more strongly biases the English meaning of the interlingual homograph *CHAT* (i.e., a conversation in English) than the sentence *Since they liked each other, they had an extended CHAT that lasted all night*. By comparing how bilinguals read cognates or interlingual homographs vs. language-specific control words embedded in high- vs. low-constraint sentences, researchers can assess whether sentential constraint can attenuate cross-language activation, and if so, determine the time-course of such effects. Below, we review the effects of sentential constraint on nonselective access during L2 reading, followed by the effects during L1 reading.

L2 Sentence Reading Studies

Duyck et al. (2007) were the first to use eye movements to investigate the time-course of cross-language activation during L2 reading. Specifically, Dutch–English bilinguals read low-constraint sentences that contained cognates (e.g., *Hilda bought a new RING and showed it to everyone; RING* is a piece of jewelry in Dutch and English) or matched language-specific control words (e.g., *Hilda bought a new COAT and showed it to everyone*; *COAT* is an English-specific word). Cognate facilitation was observed during the earliest stages of reading (i.e., first fixation duration), and persisted throughout early- and late-stage reading. However, the effect was strongest when the written form was identical (e.g., *RING*) vs. nonidentical (e.g., *SCHIP* is Dutch for *SHIP*) across the two languages. Thus, Duyck et al. (2007) ultimately demonstrated that cross-language activation can occur during the earliest stages of L2 reading, and importantly, that cross-language activation varies as a function of the amount of cross-linguistic overlap.

A follow-up study by Van Assche et al. (2011) investigated how the amount of sentential constraint and cross-linguistic overlap relate to cross-language activation during L2 reading. Specifically, Dutch–English bilinguals read cognates and matched control words embedded in high- vs. low-constraint sentences. Similar to Duyck et al. (2007), the authors observed cognate facilitation during early- and late-stage reading. However, the amount of cognate facilitation was not significantly reduced by high-constraint sentences. Thus, Van Assche et al. (2011) ultimately demonstrated that bilinguals automatically activate both languages even in the presence of language-biasing contextual constraint during L2 reading. However, one point to consider is that cognates, by virtue of sharing orthography and semantics

Table 8.1 Sample stimuli from Libben and Titone (2009)

Word type	High-constraint sentence	Low-constraint sentence
Cognates		
Target word	When they were on the safari, they saw an enormous *jungle* that was dark and scary	When they were on their trip, they saw an enormous *jungle* that was dark and scary
Matched control	When she was chewing her gum, she blew an enormous *bubble* that was pink and shiny	When she waited for her friend, she blew an enormous *bubble* that was pink and shiny
Interlingual homographs		
Target word	Because she knew the change was counterfeit, the brown colored *coin* was thrown out	Because it was completely worthless, the brown colored *coin* was thrown out
Matched control	Because it didn't clean and lather well, the brown colored *soap* was thrown out	Because it smelled really bad, the brown colored *soap* was thrown out

across languages, necessarily have a higher form-frequency, which might render them less sensitive to the impact of sentential constraint—either towards an L1 or L2 meaning. Thus, interlingual homographs might serve as a better index of cross-language activation: word meaning *differs* across languages, which can be consistent or inconsistent with the sentential context.

Libben and Titone (2009) investigated just that by examining interlingual homographs in addition to cognates. In their study, French–English bilinguals read form-identical cognates, interlingual homographs, and matched control words embedded in high- vs. low-constraint sentences. High-constraint sentences biased the L2 (English) meaning of the cognates and interlingual homographs. Sample experimental stimuli from Libben and Titone (2009) are presented in Table 8.1. Participants responded to comprehension questions on 25 % of the trials to ensure adequate comprehension. Participants also completed a Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) to assess their language background. For low-constraint sentences, the authors observed cross-language activation, indexed by cognate facilitation and interlingual homograph interference, during both early- (i.e., first fixation duration, gaze duration) and late-stage reading (i.e., go past time, total reading time). However, for high-constraint sentences, the authors observed cross-language activation *only* during early-stage reading (e.g., gaze duration). In other words, cross-language activation was attenuated for high-constraint sentences during late-stage reading (e.g., total reading time). Moreover, post hoc analyses revealed that the degree of cognate facilitation was modulated by L2 proficiency: high vs. low L2 proficiency bilinguals demonstrated increased cognate facilitation. Thus, Libben and Titone (2009) ultimately demonstrated that cross-language activation occurs during early-stage L2 reading (even in the presence of L2-biasing contextual constraint), but that high contextual constraint can attenuate cross-language activation during late-stage L2 reading. Moreover, their results highlight the notion that individual differences in L2 proficiency among bilinguals also modulate the magnitude of nonselective lexical access during L2 reading (e.g., Degani & Tokowicz, 2010).

Recently, Pivneva, Mercier, and Titone (2014) conducted an eye movement study that examined how individual differences among bilinguals in nonlinguistic executive control relate to nonselective lexical access during L2 sentence reading. Executive control has been shown to significantly influence both bilingual language comprehension (e.g., Blumenfeld & Marian, 2013; Mercier, Pivneva, & Titone, 2013; Shook & Marian, 2013) and production (reviewed in Kroll & Gollan, 2014). The authors administered the same materials as Libben and Titone (2009), but also assessed individual differences in executive control among bilinguals using a nonlinguistic executive control battery (e.g., arrow versions of Simon and Stroop tasks; Blumenfeld & Marian, 2011); Antisaccade task (Hallett, 1978); and Number Stroop task (Mercier et al., 2013; Pivneva, Palmer, & Titone, 2012). A composite executive control score (based on performance on the executive control battery) was computed.

While sentence constraint did not significantly reduce cross-language activation in this study, individual differences among bilinguals on executive control and L2 proficiency did. Specifically, greater executive control was associated with reduced cross-language activation (indexed by reduced interlingual homograph interference), while greater L2 proficiency was associated with reduced cross-language activation (indexed by reduced cognate facilitation). These effects were observed during early-stage reading, suggesting that individual differences among bilinguals can attenuate cross-language activation earlier than sentence constraint. Thus, Pivneva et al. (2014) ultimately demonstrated that nonselective access might vary as a function of the nature of the cross-linguistic overlap, and also, individual differences in executive control and L2 proficiency among bilinguals during L2 reading.

To summarize, eye movement studies of bilingual L2 reading generally show that nonselective access, indexed by cognate facilitation and interlingual homograph interference, can be attenuated by individual differences in L2 proficiency and executive control during early-stage reading, but by sentential constraint during late-stage reading.

L1 Sentence Reading Studies

Eye movement studies investigating nonselective access have predominantly examined such effects during L2 reading. However, two published studies have examined these effects during L1 reading (Titone et al., 2011; Van Assche et al., 2009).

Van Assche et al. (2009) were the first to use eye movement recordings to investigate the time-course of cross-language activation during L1 reading. Specifically, Dutch–English bilinguals read cognates and matched control words embedded in low-constraint sentences (e.g., *Ben heft een oude OVEN/LADE gevonden tussen de*

Rommel op zolder: "Ben found an old OVEN/DRAWER among the rubbish in the attic"). Similar to the findings from the L2 reading studies, the authors observed cognate facilitation during early-stage reading, which was modulated by the amount of cross-linguistic overlap: more overlap resulted in greater cognate facilitation. Thus, Van Assche et al. (2009) ultimately demonstrated that nonselective lexical access occurs even when bilinguals read in their more dominant L1.

Subsequently, Titone et al. (2011) examined whether sentential constraint can attenuate nonselective lexical access during L1 reading to the same extent as it does during L2 reading. Across two experiments, English–French bilinguals read form-identical cognates, interlingual homographs, and matched control words embedded in high- vs. low-constraint sentences. Their materials were also taken from Libben and Titone (2009). In Experiment 1, the authors observed cross-language activation during early-stage reading (indexed by cognate facilitation), irrespective of contextual constraint. Interestingly, cross-language activation during early-stage reading was heightened for bilinguals with an early L2 age of acquisition (AoA), again, irrespective of contextual constraint. However, during late-stage reading, cross-language activation was attenuated for high-constraint sentences across all bilinguals. In Experiment 2, the authors inter-mixed L2 filler sentences with the L1 sentences from Experiment 1 to examine whether increased L2 saliency would increase cross-language activation during L1 reading. Their results were consistent with this aim: cognate facilitation was observed during both early- and late-stage reading, irrespective of contextual constraint. Of note, interlingual homograph interference was only observed during late-stage reading across both experiments. The authors maintain that processing differences in cognates vs. interlingual homographs are driven by representational differences in the bilingual mental lexicon.

To summarize, eye movement studies of bilingual L1 reading show that nonselective lexical access occurs even when bilinguals read in their more dominant L1, especially with early L2 AoA. However, consistent with the findings from the L2 sentence reading studies, contextual constraint can attenuate nonselective lexical access during late-stage L1 reading. Moreover, nonselective lexical access can be heightened by increasing L2 saliency (e.g., L2 filler sentences), which in turn, can counteract the effects of contextual constraint during L1 reading.

Lexical Entrenchment

An important consequence of bilinguals' divided L1/L2 exposure is that, relative to monolinguals, bilinguals have delayed lexical access resulting from reduced baseline activation levels of words (e.g., Bilingual Interactive Activation Plus Model, BIA+; Dijkstra & Van Heuven, 2002) and/or weakened lexical memory representations due to reduced integration of word-related information (e.g., orthography, phonology, semantics) in memory (e.g., *Weaker-Links Hypothesis*; Gollan, Montoya, Cera, & Sandoval, 2008; *Frequency-Lag Hypothesis*; Gollan et al., 2011). Moreover, because bilinguals generally have reduced L2 vs. L1 exposure, bilinguals have delayed L2 vs. L1 lexical access. Word frequency effects can serve as a proxy for lexical activation, particularly because they are observed during early-stage reading (e.g., first fixation duration)—reflecting the earliest stages of lexical access (Rayner, 1998, 2009). Word frequency effects are the finding that high-frequency (HF) words (e.g., *HOME*) are recognized more accurately and rapidly than low-frequency (LF) words (e.g., *FERN*; e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986). Although the specific locus of word frequency effects is disputed among cognitive models of language processing (e.g., Monsell, 1991; Murray & Forster, 2004), word frequency indisputably impacts the structure of the mental lexicon (Rayner, 1998, 2009). However, the extant research has disproportionately examined word frequency effects in monolinguals, and thus, may not accurately reflect word frequency effects in bilinguals. Specifically, knowledge and use of an L2 may differentially impact lexical access across the L1 and L2, resulting in differential L1 and L2 word frequency effects. Of note, word frequency effects are asymptotic in nature, such that increased exposure to LF words decreases lexical access times (i.e., LF words are shifted towards asymptote), whereas increased exposure to HF words marginally affects lexical access times (i.e., HF words are already near or at asymptote). Accordingly, increased L2 exposure should strengthen L2 lexical entrenchment, thereby reducing L2 word frequency effects (Dijkstra & Van Heuven, 2002; Gollan et al., 2008, 2011). Conversely, decreased L1 exposure should weaken L1 lexical entrenchment, thereby increasing L1 word frequency effects (Dijkstra & Van Heuven, 2002; Gollan et al., 2008, 2011).

The few studies that have examined word frequency effects during bilingual reading can be divided into those using response-based tasks vs. those using eye movement recordings. The findings from studies using response-based tasks suggest that bilingual L2 word frequency effects are larger than monolingual L1 word frequency effects, and that bilingual L2 word frequency effects are larger than bilingual L1 word frequency effects. For example, Lemhöfer et al. (2008) found larger bilingual L2 vs. monolingual word frequency effects using a progressive demasking task; however, bilingual L1 word frequency effects were not examined (see also Diependaele et al., 2013). Moreover, Duyck, Vanderelst, Desmet, and Hartsuiker (2008) found larger bilingual L2 vs. L1 word frequency effects using a lexical decision task; however, no differences were found between bilingual L1 vs. monolingual L1 word frequency effects. Finally, consistent with Gollan et al. (2011), Lemhöfer et al. (2008), found larger bilingual L2 vs. monolingual L1 word frequency effects using a lexical decision task; however, bilingual L1 word frequency effects were, again, not examined.

The findings from studies using eye movement recordings, which we believe reflect more naturalistic language processing (Rayner, 1998, 2009), are mixed. Gollan et al. (2011) found no differences between bilingual L2 vs. monolingual L1 word frequency effects during L2 sentence reading, and no differences in bilingual L2 word frequency effects across different bilingual groups (e.g., L1- vs. L2-dominant bilinguals). In contrast, Whitford and Titone (2012) found larger L2 vs. L1 word frequency effects across L1-dominant bilinguals, which were modulated by individual differences in current L2 exposure. Specifically, bilinguals with high vs. low L2 exposure exhibited smaller L2 word frequency effects during L2 paragraph reading, but larger L1 word frequency effects during L1 paragraph reading. Thus, Whitford and Titone (2012) ultimately demonstrate that different patterns of bilingual L1 and L2 word frequency effects (in the *same* group of bilinguals) are driven by individual differences in L2 exposure.

However, the Gollan et al.'s (2011) and Whitford and Titone's (2012) studies differ in a number of potentially important ways. For example, Gollan et al. (2011) did not examine bilingual L1 word frequency effects within and across bilinguals, while Whitford and Titone (2012) did not examine bilingual L1/L2 vs. monolingual L1 word frequency effects. Moreover, the studies evaluated different types of bilinguals. Specifically, Gollan et al. (2011) examined L2-dominant Spanish–English bilinguals in the USA and highly L1-dominant Dutch–English bilinguals in the Netherlands, while Whitford and Titone (2012) examined a wide range of L1-dominant French–English bilinguals in Canada, from balanced to highly L1-dominant (see Baum & Titone, 2014; Titone & Baum, 2014 for discussion of how the language environment might impact language processes).

Regardless of the type of bilingual being evaluated, it is important to accurately assess L1/L2 exposure and proficiency. Consistent with prior bilingual work, Whitford and Titone (2012) assessed L1/L2 exposure and proficiency using a questionnaire modeled after the LEAP-Q. This questionnaire, which strongly correlates with objective measures of L1/L2 ability, measures current percent L1/L2 exposure and L1/L2 proficiency on a scale from 1 (*beginner*) to 7 (*native-like*). Although the questionnaire can serve as a reliable estimate of language ability, a better option is to administer the questionnaire in conjunction with objective measures of language ability, such as L1/L2 speeded lexical animacy judgment tasks, which measure how rapidly and accurately participants classify words as living (e.g., GIRL) or nonliving (e.g., BOOK; e.g., Segalowitz & Frenkiel-Fishman, 2005). Of note, Whitford and Titone (2012) did not administer objective measures of language ability, as the paragraphs were originally created as objective measures of L1/L2 ability for Titone et al.'s (2011) L1 sentence reading study.

Whitford and Titone's (2012) paragraphs were representative, naturalistic texts that readers may encounter in everyday life (see also Whitford & Titone, 2014). Specifically, the paragraphs were brief, informational articles about Canadian events (e.g., effects of a hurricane, benefits of new transportation infrastructure). There were two paragraphs of approximately 150 words each, with officially translated English and French versions, coded and matched for word frequency, length, and contextual predictability.

Whitford and Titone's (2012) experimental materials were created using Experiment Builder software. As previously mentioned, an advantage of creating paragraph materials with this experimental application is that it uses relatively straightforward drag-and-drop procedures. Specifically, experimental components, such as actions (e.g., displaying static text on the screen) and triggers (e.g., launching or terminating the display of static text on the screen upon a key press) can be dragged and dropped into a workspace. Sequential experimental components can be connected using directional arrows, resulting in a structure that resembles a flow chart (see Fig. 8.4 in the Appendix). Moreover, the properties of the experimental components can be modified. For example, by double-clicking on any DISPLAY_SCREEN, different fonts, font sizes, and background colors, can be selected. Whitford and Titone's (2012) paragraphs were presented in yellow, 14-point Courier New font (due to equidistant character spacing) on a black background (to minimize eye fatigue). Each paragraph was double-spaced on a single screen, with a maximum of 14 lines of text, and 70 characters per line. Roughly three characters subtended 1° of visual angle.

Whitford and Titone's (2012) data were collected using the Eye Link 1000 tower-mounted system, with a 1000 Hz sampling rate. Data visualization and pre-processing procedures (e.g., filtering) were performed using the EyeLink Data Viewer tool (see "Related Internet Sites" section). For example, blinks and short fixations were removed by unselecting *Display Blink Saccades* and setting the *Fixation Duration Threshold* to 100 ms within the *Data Filters* tab under *Preferences* (for more detailed instructions, see the Data Viewer user manual). Although an upper cutoff of 1000 ms is also customarily applied in eye movement reading research, Whitford and Titone (2012) did not apply an upper cutoff for more data inclusion (see also Whitford & Titone, 2014). Moreover, any drift in gaze position caused by the nervous systems' inexact control over the oculomotor system was corrected (see Rayner, 1998). Thus, interest areas were created around the boundaries of each word for drift correction and report generation purposes. Rectangular interest areas can be automatically generated or manually created in DataViewer. For example, the sentence, *Every morning, Mary drinks a cup of tea*. would have the following interest areas around each of its constituent words:

Every	morning,	Mary	drinks	a	cup	of	tea.

These interest areas circumscribe the fixations and saccades that are linked to processing a particular word. Interest area reports can be generated by selecting *Analysis* in Data Viewer's upper toolbar, then *Reports,* and then *Interest Area Report*. Although not exhaustive, some variables of interest that can be extracted include first fixation duration (IA_FIRST_FIXATION_DURATION), gaze duration (IA_FIRST_RUN_DWELL_TIME), and total reading time (IA_DWELL_TIME). A description of each variable is provided upon its selection. The reports are outputted in Microsoft Excel format, at which point, all 0 values in fixation measures can be removed using Excel's search and replace function. Finally, the eye movement data spreadsheet can be combined with separate spreadsheets containing participant linguistic/demographic information and word-level information (e.g., length, frequency, contextual predictability). A relatively easy and straightforward way is by using the merging function within R (R Development Core Team, 2010; see also "Related Internet Sites" section). Of course, all of this pre-processing would be easier to do if one were in a position to write software that pre-processes the data semiautomatically (we say "semiautomatically" rather than "automatically" because we strongly believe in the need to continually visually inspect one's data to be certain that the software's algorithms are operating as expected).

Consistent with prior work (Miellet, Sparrow, & Sereno, 2007; Pollatsek, Reichle, & Rayner, 2006), Whitford and Titone (2012) excluded all words situated at the beginning and end of each line of text from analyses. Proper nouns, repeated words, function words, punctuated words, cognates, and interlingual homographs were also excluded.

The data were analyzed using linear mixed-effects (LME) models, which are models that contain both fixed effects (i.e., independent variables or predictors) and random effects (i.e., sources of random variation) via the *lme4* package within R (version 2.13.1; Baayen, Davidson, & Bates, 2008; Bates, 2007; R Development Core Team, 2010). There are several analytical advantages of using LME models over standard analyses of variance. For example, trial-level data serves as input, thus, there is no loss of information by averaging over participants and items; heterogeneity of variance from both participants and items (i.e., random effects variables) can be simultaneously accounted for; statistical outliers have a reduced impact; and statistical power is increased (e.g., Baayen et al., 2008).

The same model was applied to each eye movement measure or dependent variable (gaze duration, total reading time), which was broken up into L1 and L2 reading. Fixed factors (i.e., the independent variables or predictors) included word frequency (continuous) and current L2 exposure (continuous). Random factors (i.e., sources of random variation) included participants and items (random intercepts only). Of note, random slope adjustments were not included as Barr, Levy, Scheepers, and Tily's (2013) paper on maximal random effects structure had not been published at that point in time. In particular, Barr et al. (2013) maintain that the inclusion of the maximal random effects structure justified by the experimental design improves the generalizability of the experimental treatment effects. Control predictors (i.e., covariates) were also included to account for variance in word length (continuous), contextual predictability (continuous), L2 age of acquisition (AoA; continuous), and participant native language (treatment coded: English vs. French, English = baseline). All predictors were centered, that is, the mean of all data points was subtracted from each data point, to reduce collinearity and to facilitate interpretation of the impact of the continuous independent variables (i.e., word frequency, current L2 exposure) on the dependent variables (i.e., gaze duration, total reading time). *P-values* for all fixed factors were obtained using Markov chain Monte Carlo (MCMC) sampling tests ($n = 10{,}000$), which are data augmentation algorithms used to obtain likelihood inference when data points are missing or when dependence structures are complex in nature (see Geyer, 2011).

As previously mentioned, Whitford and Titone (2012) found that graded differences in current L2 exposure among bilinguals modulate word frequency effects, and thus, lexical access across the L1 and L2. Specifically, all bilinguals demonstrated larger L2 vs. L1 word frequency effects; however, bilinguals with high vs. low L2 exposure exhibited smaller L2 word frequency effects, but larger L1 word frequency effects due to more divided L1/L2 use. Of note, Whitford and Titone (2012) examined word frequency effects using a continuous frequency measure, which improves upon prior bilingual work using binary frequency measures, such as comparisons between HF vs. LF words (e.g., Duyck et al., 2008; Gollan et al., 2011).

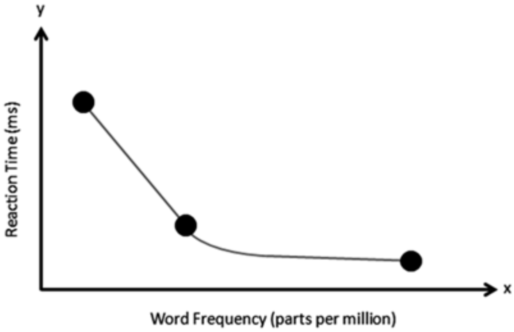


Fig. 8.1 Graphical illustration of a restricted cubic spline using artificial word frequency values and reaction times (knots are represented by the *black circles*)

However, prior work has reliably demonstrated a nonlinear relationship between word frequency effects and reaction times, although mostly during monolingual/L1 comprehension (e.g., Baayen, Feldman, & Schreuder, 2006; Keuleers, Diependaele, & Brysbaert, 2010). Thus, nonlinearity may also characterize the relationship between word frequency and eye movement measures of bilingual reading, although, of note, very few studies have examined this issue, including that of Whitford and Titone (2012). Indeed, most studies have modeled the effects of word frequency using simple linear curves for ease of data interpretation. However, a notable exception is recent work by Diependaele et al. (2013), who modeled the nonlinear effects of word frequency during performance on a progressive demasking task.

Specifically, in a recent reanalysis of Lemhöfer et al.'s (2008) data, Diependaele et al. (2013) examined whether larger bilingual L2 vs. monolingual L1 frequency effects were driven by cross-language activation (indexed by cross-language orthographic neighborhood density) or lexical entrenchment (indexed by performance on a vocabulary measure) using nonlinear models. In particular, Diependaele et al. (2013) used restricted cubic splines, which are special mathematical functions (i.e., mathematical curves) defined in a piecewise fashion by polynomials (see Harrell, 2001). In particular, cubic splines are defined piecewise by third-order polynomials. Piecewise polynomial functions are obtained by dividing the independent variable of interest (in this case, word frequency) into adjoining intervals or segments, which are represented by separate polynomials. The endpoints of each interval or segment are joined together via *knots*, which allow for a smooth, continuous function (i.e., curve). The number of knots can be chosen a priori or based on the pattern of data. Usually, a small number of knots (i.e., 3–5) are needed to model the nonlinear relationship between a dependent variable (e.g., reaction time, eye movement measure) and an independent variable (e.g., word frequency). Consistent with this range, Diependaele et al. (2013) used three knots to model nonlinearity in their data. A graphical illustration of a restricted cubic spline using artificial data can be found in Fig. 8.1.

Diependaele et al. (2013) found that larger bilingual L2 vs. monolingual L1 frequency effects, which of note, indeed had nonlinear components, were fully accounted for by individual differences in performance on the vocabulary measure. The authors maintain that their results exclusively lend support for a lexical entrenchment account, as differential frequency effects were not modulated by cross-language activation (i.e., cross-language orthographic neighborhood density). Interestingly, Diependaele et al.'s (2013) results are consistent with those of Whitford and Titone (2012), who found differential bilingual L1 and L2 frequency effects among bilinguals given individual differences in L2 exposure using exclusively linear models. Although Whitford and Titone's (2012) results are also consistent with a lexical entrenchment account: increased L2 exposure strengthened L2 lexical integration, but weakened L1 lexical integration, their results are also consistent with a cross-language activation account: larger bilingual L2 vs. L1 frequency effects could be driven by heightened activation from orthographically similar L1 word forms, which arguably have higher subjective frequencies than L2 word forms (Dijkstra & Van Heuven, 2002). Moreover, individual differences in L2 exposure and proficiency could modulate the degree of cross-language activation: highly proficient bilinguals, whose L2 word forms arguably have higher subjective frequencies than less proficient bilinguals, could encounter less activation from orthographically similar L1 word forms.

Reanalysis of Whitford and Titone's (2012) Data

To examine the source of differential L1 and L2 frequency effects among bilinguals in the Whitford and Titone's (2012) study, we reanalyzed their data using nonlinear mixed-effects models, with a particular focus on whether within- and cross-language orthographic neighborhood density modulate these effects. L1 and L2 orthographic neighborhood density was computed using WordGen (Duyck et al., 2004).

Prior work has generally reported inhibitory effects of increased cross-language neighborhood density (i.e., longer response times) on word recognition using response-based (Beauvillain, 1992; Bijeljac-Babic, Biardeau, & Grainger, 1997; Grainger & Dijkstra, 1992; Van Heuven, Dijkstra, & Grainger, 1998) and electrophysiological measures (Grossi, Savill, Thomas, & Thierry, 2012; Midgley, Holcomb, Van Heuven, & Grainger, 2008). For example, Van Heuven et al. (1998) found inhibitory effects in recognizing L1 and L2 words with many vs. few cross-language neighbors using both language-specific and language-general progressive demasking and lexical decision tasks. More specifically, bilinguals identified target words and decided whether a target was a word/non-word in the L1 or L2, respectively (see also Grainger & Dijkstra, 1992). Within-language neighborhood density effects varied as a function of language: inhibitory effects were found in recognizing L1 words with many within-language (L1) neighbors; however, facilitatory effects (i.e., faster response times) were found in recognizing L2 words with many within-language (L2) neighbors. This reversal in the impact of within-language neighborhood density, which contrasts with their monolingual findings (facilitatory effects were found), may be driven by lexical organization differences across the L1 and L2.

However, not all studies have replicated these findings (de Groot, Borgwaldt, Bos, & Van den Eijnden, 2002; Lemhöfer et al., 2008). For example, using an L2 progressive demasking task, Lemhöfer et al. (2008) found both negligible cross-language (L1) neighborhood density and neighborhood frequency effects (orthographic neighbors' summed frequencies) when bilinguals processed L2 words. This suggests that L1 knowledge did not affect L2 processing; however, the reverse was not tested, in that the effects of L2 knowledge on L1 word recognition were not examined. In contrast, both inhibitory within-language (L2) neighborhood density and neighborhood frequency effects were found in recognizing L2 words with many HF neighbors.

Thus, the potential effects of cross-language neighborhood density on bilingual reading are unclear, and two possible patterns are logically possible. Given Van Heuven et al.'s (1998) results (see also Grainger & Dijkstra, 1992), it is possible that cross-language activation of orthographically similar L1 word forms (i.e., L1 neighbors) during L2 word recognition results in competition—L1 word forms arguably have higher subjective frequencies than L2 word forms, thereby interfering with L2 lexical access, and ultimately, increasing L2 frequency effects. However, the above-cited studies that support this view involve overt responses to isolated words, which may not reflect natural reading. Thus, another possibility is that cross-language activation of L1 neighbors during L2 word recognition results in facilitation—identification of target L2 word forms vs. other competing, visually similar word forms is reinforced, thereby speeding L2 lexical access, and ultimately, reducing L2 frequency effects. Of note, this view has obtained substantial support from the monolingual response-based literature (e.g., Andrews, 1989, 1992; Carreiras, Perea, & Grainger, 1997; Pollatsek, Perea, & Binder, 1999, Experiment 1; Sears, Hino, & Lupker, 1995, but see Coltheart et al., 1977; Snodgrass & Mintzer, 1993), where increased neighborhood density facilitates lexical access, particularly for LF words (for reviews, see Andrews, 1997; Grainger, 1992; Perea & Rosa, 2000).

However, this view has received little support from the monolingual eye movement reading literature. For example, Pollatsek et al. (1999) found inhibitory effects of increased neighborhood density (i.e., longer fixation durations) on lexical access during sentence reading (Experiment 2). However, when the number of HF neighbors was controlled for, the authors found facilitatory (i.e., shorter fixation durations) effects during early-stage reading, but inhibitory effects during late-stage reading, suggesting word misidentification or misreading on the first pass (Experiment 3; for similar neighborhood frequency effects, see Perea & Pollatsek, 1998; Slattery, 2009, but see Sears, Campbell, & Lupker, 2006). Thus, it remains unclear whether facilitatory or inhibitory within- and cross-language neighborhood density effects would extend to bilingual reading, and how these effects relate to word frequency.

Following Whitford and Titone (2012), we report the results for gaze duration (reflecting early-stage reading) and total reading time (reflecting late-stage reading). We applied the same model to each eye movement measure, which was broken up into L1 and L2 reading. Random factors included random intercepts for participants and items, and by-participant random slope adjustments to the linear and cubic frequency components (consistent with Diependaele et al., 2013; see also Barr et al., 2013). Fixed factors included linear and cubic frequency components (continuous), current L2 exposure (continuous), and log-transformed L1 or L2 neighborhood density (continuous). Control predictors included the three-way interaction between the linear and cubic frequency components, current L2 exposure, and L1 or L2 neighborhood density (i.e., L1 neighborhood density was a control predictor when L2 neighborhood density was a fixed factor, and vice versa), word length (continuous), contextual predictability (continuous), AoA (continuous), and participant native language (deviation coded: English vs. French, −0.5, +0.5). Thus, our models allowed us to examine the effects of L1 and L2 neighborhood density simultaneously. All predictors were centered (i.e., the mean of all data points was subtracted from each data point) to reduce collinearity. Of note, although all fixed effects were analyzed continuously in all models, they were dichotomized using a median split to facilitate interpretation of higher-order interactions. Significant effects were reported following the convention $t > 1.96$, which represents a .05 alpha level. Frequency was divided into HF (log subtitle word frequency >2.45) and LF words (log subtitle word frequency <2.45); current L2 exposure was divided into high (>30 %, but <50 %) and low L2 exposure (≤30 %); and L1 and L2 neighborhood density were divided into low (value of 0) and high (value >0).

Of note, modeling nonlinear relationships within linear regression models requires data transformation procedures. Thus, the data were analyzed using nonlinear mixed-effects models, implemented within the *nlme* package of R (version 2.13.1; Baayen et al., 2008; Bates, 2007; R Development Core Team, 2010). Following Diependaele et al. (2013), we used restricted cubic splines with three knots, implemented within the *rcspline.eval Hmisc* function of R (Harrell, 2011). An advantage of using restricted cubic splines is that small data changes within a particular interval or segment of a polynomial expansion have a reduced impact on the curvature of the piecewise polynomial fit (see Harrell, 2001). This ultimately makes the results more generalizable.

Gaze duration—L1 model. Only main effects of linear ($b = -19.71$, $SE = 5.35$, $t = -3.69$) and cubic ($b = 16.44$, $SE = 5.91$, $t = 2.79$) word frequency occurred, where gaze durations were shorter for HF (254 ms) vs. LF words (299 ms). No interactions reached significance. Thus, this result represents the standard word frequency effect, that is, the finding that HF words are processed more rapidly than LF words (Rayner, 1998, 2009). In addition, this result suggests that there is indeed a nonlinear relationship between word frequency and gaze durations during L1 reading. As well, this result suggests that L1 lexical access is insensitive to the effects of within-language (L1) and cross-language (L2) neighborhood density, as interactions between word frequency and neighborhood density did not reach significance.

Gaze duration—L2 model. Main effects of linear ($b = -43.25$, $SE = 6.88$, $t = -6.28$) and cubic ($b = 36.24$, $SE = 8.13$, $t = 4.46$) word frequency occurred, where gaze durations were shorter for HF (258 ms) vs. LF words (361 ms). Again, this result represents the standard word frequency effect, and suggests that there is indeed a

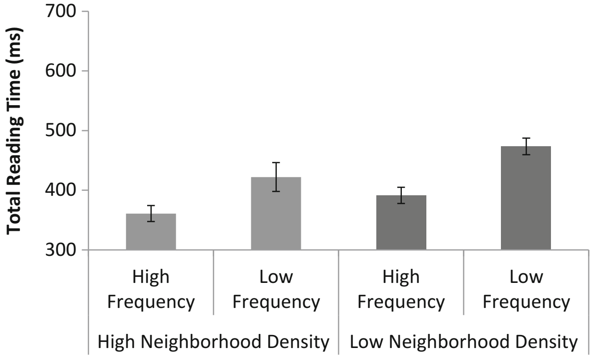


Fig. 8.2 Graphical representation of the interaction between L1 word frequency and L2 (cross-language) neighborhood density for total reading time (means and standard errors are plotted)

nonlinear relationship between word frequency and gaze durations during L2 reading (consistent with Diependaele et al., 2013).

Moreover, the three-way interaction between linear word frequency, current L2 exposure, and within-language (L2) neighborhood density was significant ($b=1.59$, $SE=0.76$, $t=2.08$), where word frequency effects were larger for bilinguals with low vs. high L2 exposure (consistent with Whitford & Titone, 2012), particularly for words with low within-language (L2) neighborhood densities. Of note, word frequency effects for words with high within-language (L2) neighborhood densities were particularly small in bilinguals with high L2 exposure. This result suggests that L2 lexical access is facilitated by within-language (L2) neighborhood density, but is, however, insensitive to cross-language (L1) neighborhood density.

Total reading time—L1 model. Main effects of linear ($b=-38.45.25$, $SE=9.93$, $t=-3.87$) and cubic ($b=37.09$, $SE=10.68$, $t=3.47$) word frequency occurred, where total reading times were shorter for HF (323 ms) vs. LF words (450 ms). Again, this result represents the standard word frequency effect, and suggests that there is indeed a nonlinear relationship between word frequency and total reading times during L1 reading. A main effect of cross-language (L2) neighborhood density also occurred ($b=146.86$, $SE=46.89$, $t=3.13$), where total reading times were shorter for words with high (386 ms) vs. low cross-language (L2) neighborhood densities (455 ms). This result suggests that late-stage L1 word recognition is facilitated by cross-language (L2) neighborhood density.

Moreover, a two-way interaction between linear word frequency and cross-language (L2) neighborhood density ($b=44.84$, $SE=14.76$, $t=3.04$), as well as cubic word frequency and cross-language (L2) neighborhood density occurred ($b=-40.23$, $SE=13.98$, $t=-2.88$), where word frequency effects were smaller for L1 words with high vs. low cross-language (L2) neighborhood densities (see Fig. 8.2). More specifically, LF-L1 words benefitted most from increased cross-language (L2) neighborhood density. Again, this result suggests that late-stage L1 word recognition is facilitated by cross-language (L2) neighborhood density.

Total reading time—L2 model. Main effects of linear ($b=-77.29$, SE$=11.64$, $t=-6.64$) and cubic ($b=72.12$, SE$=13.70$, $t=5.27$) word frequency occurred, where

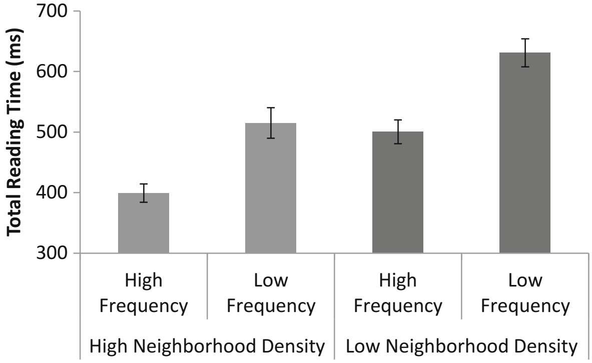


Fig. 8.3 Graphical representation of the interaction between L2 word frequency and L1 (cross-language) neighborhood density for total reading time (means and standard errors are plotted)

total reading times were shorter for HF (357 ms) vs. LF words (567 ms). Again, this result represents the standard word frequency effect and suggests that there is indeed a nonlinear relationship between word frequency and total reading times during L2 reading. Main effects within-language (L2) neighborhood density ($b=204.45$, $SE=53.97$, $t=3.79$), as well as cross-language (L1) neighborhood density also occurred ($b=-123.61$, $SE=54.77$, $t=-2.26$), where total reading times were shorter for words with high vs. low within-language (L2) and cross-language (L1) neighborhood densities (374 vs. 579 ms; 415 vs. 605 ms), respectively. These results suggest that late-stage L2 word recognition is facilitated by within-language (L2) and cross-language (L1) neighborhood density.

Moreover, a two-way interaction between linear word frequency and within-language (L2) neighborhood density ($b=61.95$, $SE=16.02$, $t=3.87$), as well as cubic word frequency and within-language (L2) neighborhood density ($b=-64.99$, $SE=17.31$, $t=-3.76$) occurred, where word frequency effects were smaller for L2 words with high vs. low within-language (L2) neighborhood densities. More specifically, LF L2 words benefitted most from increased within-language (L2) neighborhood density. Similarly, a two-way interaction between linear word frequency and cross-language (L1) neighborhood density ($b=-39.39$, $SE=15.29$, $t=-2.58$), as well as cubic word frequency and cross-language (L1) neighborhood density ($b=39.32$, $SE=17.63$, $t=2.23$) also occurred, where again, word frequency effects were smaller for L2 words with high vs. low cross-language (L1) neighborhood densities (see Fig. 8.3). More specifically, LF-L2 words benefitted most from increased cross-language (L1) neighborhood density. Again, these results suggest that late-stage L2 word recognition is facilitated by within-language (L2) and cross-language (L1) neighborhood density.

Summary of results. Taken together, the reanalysis of Whitford and Titone's (2012) data suggests that both cross-language activation (indexed by the effects of cross-language neighborhood density) and lexical entrenchment (indexed by the effects of word frequency) synergistically mediate bilingual L1 and L2 reading. More specifically, we found that during late-stage L1 and L2 reading (total reading times), both L1 and L2 word recognition was facilitated by increased cross-language neighborhood density (i.e., more L2 and L1 orthographic neighbors, respectively), indexed by smaller L1 and L2 word frequency effects. Conversely, we found no impact of cross-language neighborhood density on early-stage L1 and L2 reading (gaze durations); however, we did find that increased within-language (L2) neighborhood density facilitated early-stage L2 reading, indexed by smaller L2 word frequency effects. Accordingly, cross-language activation from orthographically similar word forms has a late temporal course, impacting later stages of L1 and L2 post-lexical processing only. Interestingly, although bilinguals generally have more L1 than L2 experience, which potentially results in L1 word forms being more lexically entrenched (i.e., higher baseline activation levels of L1 words and/or increased strength of L1 word-related information in memory), these results suggest that L1 lexical representations are indeed sensitive to the impact of cross-language activation from relatively less entrenched L2 word forms.

Interestingly, facilitatory effects of increased cross-language neighborhood density on L1 and L2 word recognition are inconsistent with prior work from the bilingual response-based literature, where largely inhibitory influences have been reported (e.g., Grainger & Dijkstra, 1992; Van Heuven et al., 1998). However, the nature of the tasks used (e.g., overt responses to isolated words) may not reflect the cognitive processes implicated in natural reading. Thus, one possibility is that activation of visually similar word forms (even from the nontarget language) facilitates target word recognition among other visually similar word forms during reading. Indeed, similar findings have been reported in the monolingual response-based literature (for reviews, see Andrews, 1997; Grainger, 1992; Perea & Rosa, 2000). Consistent with this literature, we also found facilitatory within-language neighborhood density effects during both early- and late-stage L2 reading. Specifically, increased within-language (L2) neighborhood density resulted in smaller L2 word frequency effects, where LF words benefitted most from having more within-language (L2) neighbors (see also Grainger & Dijkstra, 1992; Van Heuven et al., 1998, for similar findings during L2 within-language word recognition). Of note, however, these findings are inconsistent with the monolingual eye movement literature, where inhibitory effects of increased neighborhood density have been reported during both early- and late-stage reading (Pollatsek et al., 1999, Experiment 2). However, when these studies controlled for *neighborhood frequency*, increased neighborhood density had facilitatory effects on early-stage reading, but inhibitory effects on late-stage reading. This finding suggests that participants might have been misreading target words as their higher-frequency neighbors during the first pass (Pollatsek et al., 1999, Experiment 3; for similar neighborhood frequency effects, see Perea & Pollatsek, 1998; Slattery, 2009). However, our lack of inhibitory effects during late-stage reading, in conjunction with the fact that participants in Whitford and Titone's (2012) study had relatively high reading comprehension performance (93 % average accuracy), suggest that participants were unlikely misreading words during the first pass.

On the whole, our findings suggest that the theoretical and empirical contrast pointed out by Diependaele et al. (2013), that is, that cross-language activation and lexical entrenchment are distinct mechanisms that independently modulate bilingual language processing, may not accurately reflect natural bilingual L1 and L2 reading.

Our findings also have implications for current models of bilingual visual word recognition, such as the Weaker Links (Gollan et al., 2008) or Frequency Lag Hypothesis (Gollan et al., 2011) and the Bilingual Interactive Activation Plus Model

(BIA+, Dijkstra & Van Heuven, 2002). In particular, the Weaker Links or Frequency Lag Hypothesis is primarily concerned with how divided L1/L2 knowledge and use among bilinguals (and in contrast to monolinguals) impact L1 and L2 word recognition. In particular, it posits that L1 and L2 lexical representations are weakened by reduced L1 and L2 experience, respectively, resulting in larger word frequency effects. Thus, it emphasizes the role of lexical entrenchment in bilingual L1 and L2 word recognition; however, the present work's findings suggest that cross-language activation is an equally important mediator of bilingual L1 and L2 recognition. In contrast, BIA+ is concerned with how divided L1/L2 knowledge and use among bilinguals (and in contrast to monolinguals) impact L1 and L2 word recognition and, also, the processing of lexically ambiguous words. In particular, it posits that baseline activation levels of L1 and L2 words decrease with reduced L1 and L2 experience, respectively, resulting in larger word frequency effects. As well, it posits that cross-language neighborhood density has inhibitory effects on single-word recognition: orthographically similar word-forms compete for activation, resulting in larger word frequency effects. Thus, in contrast to the Weaker Links or Frequency Lag Hypothesis, BIA+ has the added advantage of emphasizing both the roles of lexical entrenchment and cross-language activation in bilingual L1 and L2 word recognition. However, contrary to its predictions regarding the impact of cross-language neighborhood density, we find that cross-language neighborhood density has facilitatory effects on target-word recognition: orthographically similar word-forms boost target-word activation, resulting in smaller word frequency effects. Thus, our findings suggest that the direction of cross-language neighborhood density effects may be task-specific (i.e., inhibitory when using response-based measures to isolated words vs. facilitatory when using eye movement recordings during natural reading). As such, BIA+'s authors may want to implement changes to accommodate cross-language neighborhood density effects during more contextualized reading.

Conclusions

To conclude, in this chapter we reviewed some background information on the use of eye movement recordings to examine the cognitive processes underlying reading (e.g., how eye movement data are acquired, what measures can be extracted from the eye movement record), and two key issues in the bilingual eye movement reading literature: cross-language activation and reduced lexical entrenchment, which have generally been independently examined. We then presented a reanalysis of Whitford and Titone's (2012) published data suggesting that these two key issues are indeed two sides of the same coin with respect to natural bilingual reading. In so doing, we presented some of the technical considerations involved in implementing eye movement studies of reading, as well as the different kinds of analytic strategies one could take in understanding the data. Taken together, these findings speak to the importance of using eye movement recordings to elucidate the cognitive processes implicated in bilingual reading. As should not be surprising, given the tenor of this review, we strongly believe that there is simply no better behavioral measure available for understanding bilingual language processing (indeed, language processing generally) at the sentence and discourse level than through the eye movement approach. Although eye movement measures of reading are purely behavioral and do not necessarily tell us about the neural mechanisms underlying reading, recent successes in co-registering eye movement measures with neural measures (e.g., Henderson, Luke, Schmidt, & Richards, 2013; Kliegl, Dambacher, Dimigen, Jacobs, & Sommer, 2012) will undoubtedly make eye movement measures indispensable for understanding the cognitive neuroscience of reading in short time.

List of Keywords

Age of acquisition (AoA), Automatic activation, Backward saccades, Baseline activation, Cognate, Cognate facilitation, Contextual predictability, Cross-language activation, Demasking task, Discourse-level, Early-stage, Ecological validity, Entrenchment, Executive control, Eye movements, Eye-Mind Hypothesis, First fixation, Fixation duration, Frequency-Lag Hypothesis, Gaze duration, Global measure, Go-past time, High-frequency words, Interest area report, Interlingual homographic effect, Late-stage, Lexical decision, Lexical representation, Lexically ambiguous words, Local measures, Low-frequency words, Mental lexicon, Neighborhood density, Nonselective activation, Oculomotor processes, Orthographic neighbors, Parafoveal region, Paragraph-level, Perceptual span, Phonological, Post-lexical processing, Random effect variables, Random factors, Reading goal, Reading rate, Reading time, Regressive saccades, Response-based tasks, Saccadic eye movements, Semantic integration, Sentence-level, Simon task, Simultaneous activation, Single fixation, Sublexical processing, Total reading time, Weaker-Links Hypothesis, Word length.

Review Questions

1. The present work examines the effects of cross-language activation and reduced lexical entrenchment using behavioral measures (i.e., eye movement recordings). What do we know about the neural correlates of these bilingual reading processes?
2. The present chapter reviews work suggesting that cross-language activation and reduced lexical entrenchment are behaviorally related. Based on what we know about the neural correlates of these processes, do you think they can be neurally disentangled?
3. How might the issues reviewed in this chapter extend to other bilingual populations, such as bilinguals who know languages that have different L1 and L2 writing scripts (e.g., English and Arabic), or deep vs. shallow mapping between orthography and phonology within a similar script?

Suggested Student Research Projects

1. Most of the studies reviewed in this chapter did not explicitly manipulate language mode, that is, whether people were put into a first or second language frame of mind during the experiment by some deliberate manipulation ahead of the experiment. Generate hypotheses about this question and design a study that tests the impact of language mode on bilingual reading.
2. Design an experiment that tests whether the simultaneous effects of cross-language activation and reduced lexical entrenchment change as a function of different kinds of context (e.g., sentence-level vs. paragraph-level contexts; semantically vs. syntactically biased contexts).
3. Design an experiment that tests how lexical entrenchment and cross-language activation interact for trilinguals or polyglots, compared to bilinguals who know only two languages. What predictions would you make about people who speak more than two languages with respect to how these processes manifest themselves during bilingual reading?

Related Internet Sites

Data Viewer: http://www.sr-research.com/dv.html
Experimenter Builder: http://www.sr-research.com/eb.html
EyeLink 1000: http://www.sr-research.com
EyeTrack Application: http://www.psych.umass.edu/eyelab/software
GazeAlyze: Eye Movement Analysis: http://gazealyze.sourceforge.net/
LibreOffice Word Processing Suite: http://www.libreoffice.org
OpenOffice Word Processing Suite: http://www.openoffice.org
R Software for Statistical Computing: http://www.r-project.org/
WordGen: http://www.wouterduyck.be/?page_id=29
Wuggy: http://crr.ugent.be/programs-data/wuggy

Suggested Further Reading

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first and second language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology, 66*, 843–863.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

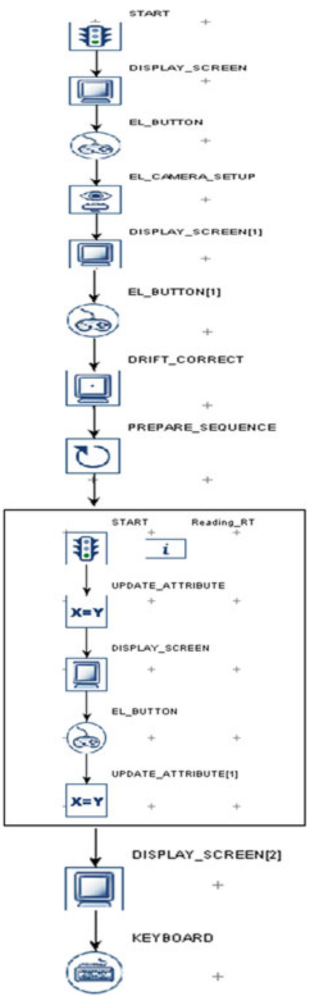
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology, 3*, 1–8

Whitford, V., & Titone, D. (2012). Second language experience modulates first- and second- language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review, 19,* 73–80.

Appendix

Appendix



START

DISPLAY_SCREEN (calibration instructions for participant)

EL_BOTTON (participant key press upon completion of calibration instructions)

EL_CAMERA_SETUP (calibration procedure: 9-point grid for paragraphs)

DISPLAY_SCREEN[1] (task instructions for participant)

EL_BOTTON[1] (participant key press upon completion of task instructions)

DRIFT_CORRECT (drift correction fixation point before paragraph presentation)

PREPARE_SEQUENCE

START

UPDATE_ATTRIBUTE

DISPLAY_SCREEN (paragraph)

EL_BOTTON (participant key press upon completion of paragraph)

UPDATE_ATTRIBUTE[1]

DISPAY_SCREEN[2] (participant notification of experiment completion)

KEYBOARD (experimenter key press to terminate experiment)

Fig. 8.4 Whitford and Titone's (2012) sample Experiment Builder script

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15*, 802–814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18*, 234–254.

Andrews, S. (1997). The effects of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review, 4*, 439–461.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baayen, R. H., Feldman, L. F., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 53*, 496–512.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bates, D. M. (2007). *Linear mixed model implementation in lme4*. Unpublished manuscript. University of Wisconsin, Madison, WI.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics, 5*, 857–894.

Beauvillain, C. (1992). Orthographic and lexical constraints in bilingual word recognition. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 221–235). Amsterdam: Elsevier.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition, 25*, 447–457.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118*, 245–257.

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25*, 547–567.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*, 857–871.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Lawrence Erlbaum.

de Groot, A. M. B., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47*, 91–124.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first and second language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology, 66*, 843–863.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. *Quarterly Journal of Experimental Psychology, 63*, 1266–1303.

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments, & Computers, 36*, 488–499.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*, 663–679.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review, 15*, 850–855.

Geyer, C. J. (2011). Introduction to Markov Chain Monte Carlo. In S. Brooks, A. Gelman, G. K. Jones, & X. L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo* (pp. 3–48). London, England: Chapman & Hall/CRC.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*, 787–814.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology. General, 140*, 186–209.

Grainger, J. (1992). Orthographic neighborhoods and visual word recognition. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 131–146). Amsterdam: Elsevier.

Grainger, J., & Dijkstra, A. (1992). On the representation and use of language information in bilinguals. In R. J. Haris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: Elsevier.

Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press.

Grossi, G., Savill, N., Thomas, E., & Thierry, G. (2012). Electrophysiological cross-language neighborhood density effects in late and early English-Welsh bilinguals. *Frontiers in Psychology, 3*, 408.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*, 1279–1296.

Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 478–500.

Harrell, F. E. (2001). *Regression modeling strategies with applications to linear models, logistic regression and survival analysis*. New York: Springer.

Harrell, F. E. (2011). Hmisc: Harrell miscellaneous (R package version 3.9-0). Retrieved from http://CRAN.R-project.org/package=Hmisc.

Henderson, J. M., Luke, S. G., Schmidt, J., & Richards, J. E. (2013). Co-registration of eye movements and event-related potentials in connected-text paragraph reading. *Frontiers in Systems Neuroscience, 7*, 1–13.

Huey, E. B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.

Immordino-Yang, M., & Deacon, T. (2007). An evolutionary perspective on reading and reading disorders? In K. W. Fischer, J. H. Bernstein, & M. H. Immordino-Yang (Eds.), *Mind, brain, and education in reading disorders* (pp. 16–29). Cambridge, England: Cambridge University Press.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431–439.

Jordan, T. R., Almabruk, A. A. A., Gadalla, E. M., McGowan, V. A., White, S. J., Abedipour, L., et al. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review, 21*, 505–511.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology, 1*, 1–15.

Kliegl, R., Dambacher, M., Dimigen, O., Jacobs, A. M., & Sommer, W. (2012). Eye movements and brain electric potentials during reading. *Psychological Research, 76*, 145–158.

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 165–181). Oxford: Oxford University Press.

Lemhöfer, K., Dijkstra, T., Sciefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in second language: A megastudy. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*, 12–31.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 35*, 381–390.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *Oxford handbook of eye movements*. Oxford, England: Oxford University Press.

Marian, V., Blumenfeld, K. H., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing, 50*, 940–967.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–586.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*, 365–368.

Midgley, K. J., Holcomb, P. J., Van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research, 1246*, 123–135.

Miellet, S., Sparrow, L., & Sereno, S. C. (2007). Word frequency and predictability effects in reading French: An evaluation of the E-Z Reader model. *Psychonomic Bulletin & Review, 14*, 762–769.

Mercier, J., Pivneva, I., & Titone, D. (2013). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition, 17*, 89–117.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Lawrence Erlbaum.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111*, 721–756.

Paterson, K. B., McGowan, V. A., White, S. J., Malik, S., Abedipour, L., & Jordan, T. R. (2014). Reading direction and the central perceptual span in Urdu and English. *PloS One, 9*, 1–8.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology. Human Perception and Performance, 24*, 767–779.

Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicológica, 21*, 327–340.

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*, 787–796.

Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology, 3*, 1–18.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*, 174–180.

Pollatsek, A., Perea, M., & Binder, K. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology. Human Perception and Performance, 25*, 1142–1158.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye movement control. *Cognitive Psychology, 52*, 1–56.

R Development Core Team. (2010). *R: A language and environment for statistical computing.* (Version 2.13.1). Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology, 66*, 429–452.

Rayner, K. (1997). Understanding eye movements in reading. *Scientific Studies of Reading, 1*, 317–339.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. E. (2012). *The psychology of reading*. New York: Psychology Press.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27*, 537–544.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*, 468–469.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74*, 5–35.

Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighbor frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology. Human Perception and Performance, 32*, 1040–1062.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood frequency and neighborhood size effects in visual word recognition. *Journal of Experimental Psychology. Human Perception and Performance, 21*, 876–900.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory & Cognition, 33*, 644–653.

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*, 304–324.

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology. Human Perception and Performance, 35*, 1969–1975.

Snodgrass, J. G., & Mintzer, M. (1993). Neighborhood effects in visual word recognition: Facilitatory or inhibitory? *Memory & Cognition, 21*, 247–266.

Titone, D., & Baum, S. (2014). The future of bilingualism research: Insufferably optimistic and replete with new questions. *Applied Psycholinguistics, 5*, 933–942.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*, 1412–1431.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*, 88–107.

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology, 3*, 1–8.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? *Psychological Science, 20*, 923–927.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483.

Whitford, V., & Titone, D. (2012). Second language experience modulates first- and second- language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review, 19*, 73–80.

Whitford, V., & Titone, D. (2014). The effects of reading comprehension and launch site on frequency-predictability interactions during paragraph reading. *Quarterly Journal of Experimental Psychology, 67*, 1151–1165.

Chapter 9
Connectionist Models of Bilingual Word Reading

Daniel Holman and Michael J. Spivey

Abstract A number of connectionist models (inspired by biological neural networks) have been designed to simulate human data in bilingual word reading tasks. These models have in common a reliance on neuron-like nodes that are connected by a distributed pattern of synapse-like connections. When some nodes become active due to linguistic input, this activation pattern spreads throughout the network and eventually activates other nodes that correspond to word recognition states and/or motor responses. Various models differ with one another on certain architectural details. Some models use localist representations, where a single node represents each word, while others use distributed representations, where each word is represented by a pattern of activation across many nodes. Some models focus on how the connection strengths are developed via a learning algorithm, while others focus more on real-time processing dynamics. By fitting existing human data, and then making explicit predictions for future experiments, these different models steadily advance our understanding of how bilinguals comprehend written words.

Introduction

Bilingualism offers an interesting challenge to the models of word and language comprehension already in place. It is generally agreed upon in the research that making sense of a written or spoken word is a process that requires first reconstructing the word itself by recognizing features, letters, and ultimately words, and then matching the word to a separate, although connected, meaning. This is to say that when reading the word *CAT*, for example, the visual system must first recognize the components of the word, such as its contours and shapes in order

D. Holman • M.J. Spivey, Ph.D. (✉)
Cognitive and Information Sciences, University of California, Merced,
5200 N. Lake Road, Merced, CA 95343, USA
e-mail: dholman@ucmerced.edu; spivey@ucmerced.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_9

to determine the letters they form, then relate the letters to each other to form the word, then lastly link the word *CAT* to the concept of a four-legged furry house pet (along with alternative meanings). There are several proposed models of this process, including the *Dual-Route Model* (Coltheart, Curtis, Atkins, & Haller, 1993), *Localist Connectionist Models* (McClelland & Rumelhart, 1981), and *Distributed Connectionist Models* (Aisa, Mingus, & O'Reilly, 2008; Seidenberg & McClelland, 1989). However, the addition of a second language (L2) adds significant complexity to this problem.

Language itself is a rather tricky thing, and it is truly remarkable how well people can make efficient use of it, especially considering the preponderance of ambiguity that it entails. Setting aside the feat of parsing a sentence to deduce which words act upon which others, individual words can be rather problematic themselves. Many words have multiple potential meanings, in the sense of abstract concepts as well as contextual differences (for instance, "pool," in the sense of a body of water, as opposed to a billiard game). There are also inconsistencies in the rules for spelling and pronunciation that every language exhibits, such as the "ough" sound in "plough" versus "tough," or the "i before e" rule that doesn't apply to words like "weird." In addition, there are words that are spelled exactly the same as one another, but pronounced differently and carrying different meanings, such as the word "wound" in *The bandage was wound around the wound*. Rounding out the mix are words that are spelled differently, but end up sounding essentially the same, such as "sinking" and "syncing." Coltheart et al.'s (1993) Dual-Route Model is aimed at identifying these rules and their exceptions, with a special focus on past-tense verbs. The Dual-Route Model places emphasis on a functional description of the rule-based processes carried out by a brain that is reading words, and it steers away from providing an explanation of how the neural implementation of those processes is achieved. That is, it focuses on the *software of the mind* rather than on the *hardware of the mind*. This model posits one processing route containing a set of rules for turning certain verbs into their regular past-tense forms, and a second route devoted entirely to storing the specific exceptions to those rules. Hence, words like "frost," "walk," "love," and "open" all follow the generalized rule of adding the suffix "-ed" to become past-tense, and would fit within the primary route. In contrast, words such as "go," "eat," "find," and "tell" do not follow this rule, and so the second route accounts for each of these exceptions by rote memory. In both cases, the word representations are distinct from their meaning representations, which is to say that the concept described by both "go" and "went" is the same, and connected to each of these words, rather than being duplicated and existing separately for each.

When applied to bilingualism, the Dual-Route Model is slightly different, but fundamentally theorizes that each language is kept separate, both linking to one semantic system. In this way, "cat" and "gato" both map onto the same concept in the same way as "eat" and "ate." If this were true, one would expect that individuals who are fluent in multiple languages would be equally fast at recognizing words from either language, as both have direct mappings to the meaning. However, this does not appear to be the case, even in individuals who learned both languages at birth. One of the languages always comes out at least slightly dominant (Sebastián-Gallés, Echeverría,

& Bosch, 2005). And the later the second language (L2) is learned, the more dominant the first language (L1) is. These age-of-acquisition effects can pose some challenges for a Dual-Route Model, which does not place much emphasis on the gradual temporal dynamics of the learning process. Neural network models of the connectionist variety can place substantial emphasis on those developmental dynamics, and thus are able to provide natural accounts of these age-of-acquisition effects (Hernandez & Li, 2007) and similar critical period phenomena (Munro, 1986).

In addition to providing a better fit to existing data, there is another reason to explore connectionist models of bilingual word reading. In science, there is a meta-theoretical reason to explore the *underlying* mechanisms of one's phenomena of interest. Every science has gradually (sometimes painfully) learned that, in order to develop a proper understanding of the mechanisms and processes that make their phenomena work, one needs to look at the *underlying* spatial and temporal scales of their phenomena (i.e., *reductionism*) and also at the *overlying* spatial and temporal scales of their phenomena (i.e., *emergentism*). Unfortunately, treating one's favorite scientific phenomena (be they chemical, biological, or cognitive) as though they belong to a *special science* (Fodor, 1974), which is not reducible or expandable, too often amounts to little more than an excuse to not have to read about partially overlapping scientific disciplines. A responsible scientific investigation of how any aspect of the mind works, such as bilingual word reading, will clearly benefit greatly from a synthesis of the larger social and cultural context in which that cognitive process takes place, a functional level of description of the cognitive operations involved, as well as an analysis of how the physical material of a brain (which we know relies on parallel processing and distributed representations) can implement those cognitive operations. The latter is our focus in this chapter.

There is also a practical reason to develop computational simulations of the mechanisms that implement one's phenomena of interest. When one builds a model simulation of the underlying mechanisms of a cognitive process, one should expect to *learn* something from it. Simply fitting a model to preexisting data, and stopping there, merely serves as an existence proof. It shows that, among the many ways that a computational process could implement the phenomena of interest, this is one of them. This is progress, to be sure, because there are always some computational processes that are simply not capable of implementing the phenomena of interest. In the process of developing a model that does work, some of those models that do not work will indeed be discovered. However, a more important form of progress can be obtained by allowing the model to generate new unexpected aspects of the simulated phenomena that can serve as explicit predictions from the theory, which can then be tested in the laboratory. Computational modeling is most beneficial to science when it engages in the recurrent loop of a model simulation making predictions, laboratory experimentation testing those predictions, and then the model being refined to better approximate the new data it helped produce, and finally, making further new predictions, for example. Only when that recurrent loop is realized, do the model and its theory become equally important as the experimental data.

It can be argued that the connectionist modeling tradition—and really the broader neural network approach in general—combines these meta-theoretical and practical motivations more evenly than other computational modeling traditions. Connectionist models respect the *underlying* physical constraints of biological networks of neurons better than most other computational modeling traditions, and they have a proven history of producing unexpected nonlinear behavior that allows their practitioners to *learn* something new from the simulation process. With this as our reason and motivation, we review here some models that attempt to provide insight into the underlying neural processes that carry out bilingual word reading.

Interactive Activation

The *Interactive Activation Model* is at the heart of the history of connectionism (Rumelhart & McClelland, 1982, 1986). There are similar models that contributed to the development of the field, such as Anderson, Siverstein, Ritz, and Jones's (1977) *Brain-State-in-a-Box Model* and Dell's (1986) *Spreading Activation Model*, but the specific architecture of the Interactive Activation Model has persisted in a number of contemporary bilingual language processing simulations (e.g., Dijkstra & Van Heuven, 1998; Grainger & Dijkstra, 1992; Scheutz & Eberhard, 2004). Interactive activation is based on the premise that the sensory system recognizes certain distinct features of the world, such as horizontal or vertical edges or curves in the case of letter recognition, and combines these individual features in order to identify larger representations (e.g., letters and words). For written word recognition in English, the original Interactive Activation Model comprises three levels: the feature level, the letter level, and the word level (see Fig. 9.1). The feature level is the lowest level,

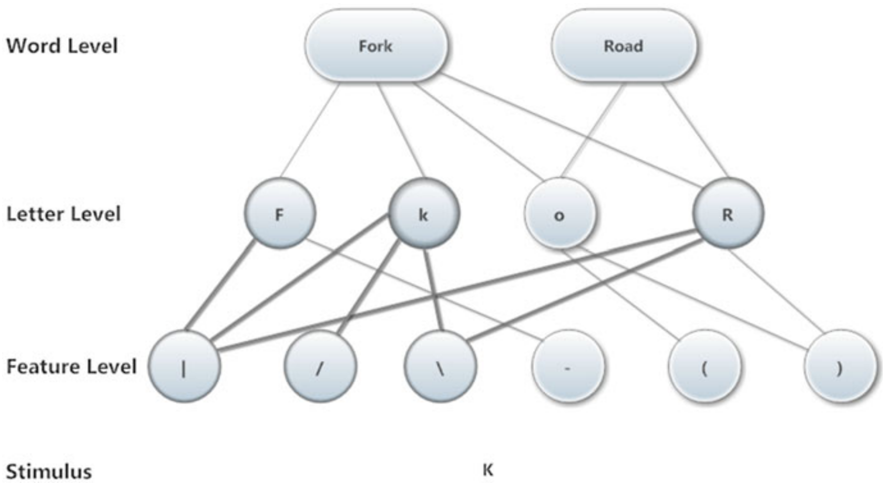


Fig. 9.1 Basic architecture of the interactive activation connectionist network, with localist representations

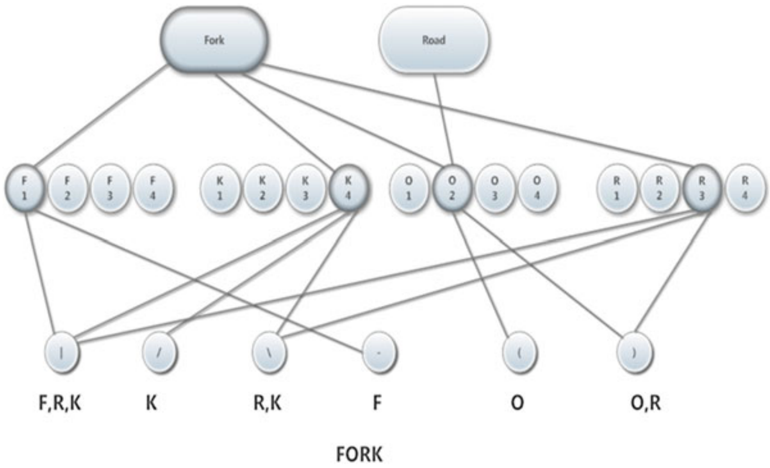


Fig. 9.2 Interactive activation with letter position specified

responding directly to visual input, and is triggered by various straight or curved lines (the model designates twelve specific *feature units*). The features chosen then activate the letter level, which consists of every known letter in a language, prompting the letters with features found in the stimulus to become active. Then, the word level, consisting of all known words, is activated by the letter level in order to discern which word has been seen. The following example demonstrates how the Interactive Activation model might work to recognize the letter "k" in "Fork."

In Fig. 9.1, the vertical and two diagonal lines that form the "K" trigger those particular feature units, which in turn activate letters with those features. The darker lines in this diagram indicate activation, which demonstrates that not only does the "k" at the letter level receive stimulation, but "F" and "R" as well, since there are features in both of those letters that are also found in "k." However, only the "k" in this instance is fully activated, having all its inputs satisfied, although "R" is a close second. This demonstrates how the model is able to determine which specific letter is being viewed, but with a little adjustment, the system also responds to specific words. In Fig. 9.2, there are units at the letter level that respond not only to the letter, but to its position in a word.

In this way, the model can accommodate words with similar letters in a different order, so even though there is an "R" in "Roof," it is in the wrong location for the stimulus, and therefore "Roof" receives less activation than "Fork." Moreover, due to position coding, words like "Dog" and "God" will not be confused by the network. Finally, in order to explain the word superiority effect (Reicher, 1969)—the fact that people respond to and recognize letters more quickly and accurately when they are part of a word than when either presented alone or as part of a non-word—the Interactive Activation Model suggests that when a word is recognized, it sends feedback activation to the letter units. In this way, the letters receive more activation when recognized as part of a word than they would otherwise.

Bilingual Interactive Activation

Following Grainger and Dijkstra (1992), Van Heuven, Dijkstra, and Grainger (1998) expanded the Interactive Activation Model to include two languages, English and Dutch. There are four levels in the Bilingual Interactive Activation (BIA) Model, consisting of 14 features for each letter position, 26 letters for each position in a word, and 1324 word nodes from English along with 978 word nodes from Dutch (see Fig. 9.3). The feature and letter levels function exactly the same way as in the original Interactive Activation Model, responding to the shapes and position of letters within a word. The word level in this model sends positive feedback activating the corresponding letters in the level below, and every word node is connected to every other word node, allowing for inhibition for the purposes of competition. Since the letters activate words of either language indiscriminately, and all words are fully interconnected, the BIA model assumes a full integration of the two languages, with no preference or tendency for one over the other. Thus, as demonstrated by Van Heuven et al., input from one language will usually spread at least some of its activation to the letters and words of the other language (for similar evidence in *spoken* word recognition, see Spivey & Marian, 1999, and Vandeberg, Guadalupe, & Zwaan, 2011).

The final level in the BIA model consists of only two nodes, one for each language. All the words for a given language serve to activate that language node, as well as inhibit the other language node. Just as the words in the Interactive Activation Model also send activation back to their constituent letters, the language

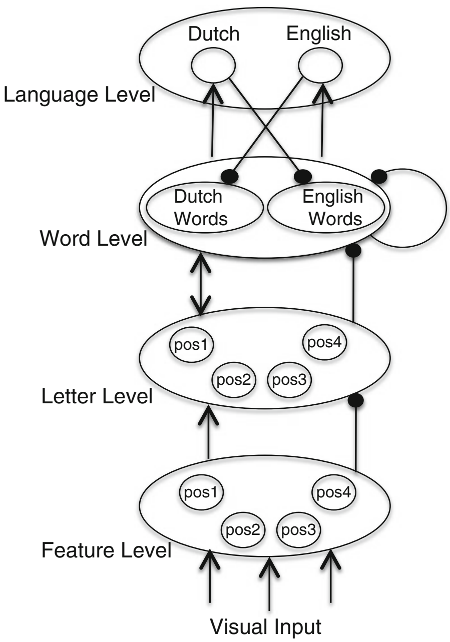


Fig. 9.3 Schematic diagram of the Bilingual Interactive Activation (BIA) network, with localist representations (Adapted from Van Heuven et al., 1998. Copyright by Elsevier)

nodes in BIA send feedback to the words beneath them, serving to give greater activation to a word that is recognized as part of a language, while simultaneously inhibiting possible competition from the alternate language. The English node and the Dutch node that receive input from their respective word nodes and then send inhibitory feedback to the other language's word nodes together serve an important purpose in this network. They allow the model to settle into different *language modes* based on the context of its input (e.g., Grosjean, 2001). If the model has been exposed to many words only from English, then it will be in a monolingual English language mode, and will be slow to process a Dutch word when it comes in. Similarly, being exposed to mostly Dutch words will result in slower processing of an English word. If it is exposed to alternating Dutch and English words (as in *code-switching* between two languages; Auer, 1999), it will behave as though in a bilingual language mode, ready to process words from either language. Bilinguals in the laboratory can have their language modes modulated in this way, and they will exhibit exactly the expected effects of easy cross-activation of both languages when in a bilingual mode and relatively little cross-activation of the irrelevant language when in a strong monolingual language mode (Marian & Spivey, 2003).

A challenge for any model of bilingualism is the existence of *cognates* and *homographs*. These are words that appear the same in multiple languages, such as "film," which exists both in English and in Dutch. When the word's meaning is the same in both, it is called a cognate. By contrast, words that appear in multiple languages but have different meanings in each are known as false-cognates, or interlingual homographs. An example of an interlingual homograph is the word "room," which in English refers to an enclosure in a building, and in Dutch refers to cream. There are two possible ways the BIA model could account for these words: either by having one word node for both languages, or by having two word nodes, one in each language. As it turns out, using only one word node for both languages results in the model recognizing those words faster than other words, which is contrary to the behavior of experimental participants. The two-node version, on the other hand, has the interesting problem of mutual inhibition—that is, both words are activated by the letter level, and then work against each other so that neither word is easily recognized. Using the language nodes, a context effect takes place such that the ambient language in the environment inhibits the words of the other language, and thus the appropriate word is recognized via this imbalance in top-down activation (Van Heuven et al., 1998). Evidence is mixed for both facilitation with cognates and inference with cognates, with the possibility of a distinct asymmetry between L1 and L2 processing (for review, see Midgley, Holcomb, & Grainger, 2011). Therefore, future versions of these models will need to take this into account.

Dijkstra and Van Heuven (2002) later expanded the BIA model by adding nodes for semantics, phonology, and orthography (see Fig. 9.4). Whereas BIA is a strictly orthographic model (meaning it operates on the visual structure of the word), this new model, BIA+, takes into account the processing of how a word, or parts of a word, may sound. In this way, it becomes possible to account for the differences in recognizing familiar or unfamiliar sounding words, the variations in how a word might sound depending on the language it is being read in, and in general the surprisingly

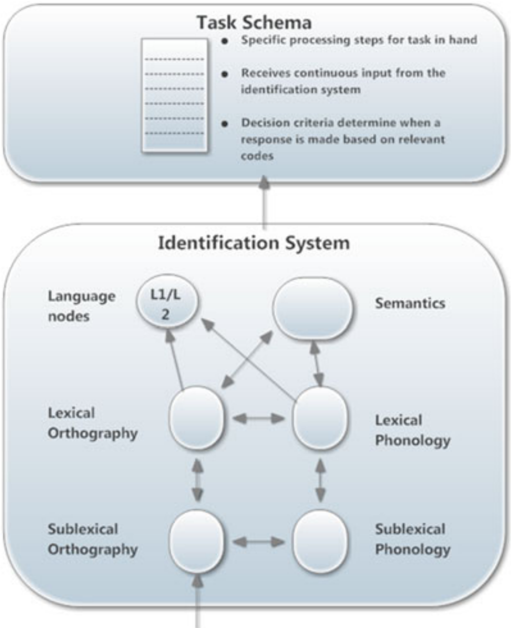


Fig. 9.4 Schematic diagram of the BIA+, with phonological constraints included (Adapted from Dijkstra & Van Heuven, 2002. Copyright 2002 by Cambridge Press)

strong role of phonology in written word recognition (Dijkstra, Grainger, & Van Heuven, 1999; Lukatela & Turvey, 1994; Van Orden, Pennington, & Stone, 1990).

The orthographic and phonological components are both included in the model, with each defining a sublexical component prior to the lexical one. In addition to these modifications, a task/decision system was added in order to be able to distinguish non-linguistic task-related effects from linguistic effects (e.g., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). Linguistic effects such as the sentence context may influence the word identification system, but non-linguistic information (such as participant expectations) are relegated to the task/decision system. While this model works for sublexical effects including word superiority in letter perception (Harley, 2001), it includes position-specific encoding, which means it cannot easily process words of different lengths.

Bilingual Single Network Model

In addition to those localist-coding networks that are directly inspired by interactive activation, where each linguistic unit (e.g., letter or word) is represented by a single neuron-like node, there are also distributed-coding networks of bilingual language processing, inspired more by the kinds of neural networks that employ distributed patterns of activation across many nodes to represent any given linguistic unit (e.g., Elman, 1990; French & Jacquet, 2004; Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989).

In the late 1990s, two different distributed approaches to the problem of bilingual word reading were being devised independently by Michael Thomas and by Robert French. In Thomas's *Bilingual Single Network* (Thomas & Van Heuven, 2005), rather than using distinct nodes associated with each word as in the previous models, the distributed method relies on particular patterns of activation within a network to represent each word. Part of the rationale for the creation of this kind of model was the issue of homographs and cognates, and how to account for the interference between multiple languages. Thomas worked off the *Single Network Hypothesis*, which asserts that interference effects are a consequence of attempting to store two languages in a common representational resource. Parts of the model were built in a similar manner to BIA, with position-specific encoding of letters as well as a binary value to indicate membership in one or the other language. Word recognition starts with the activation of letters in a word, passes to a distributed-coding hidden layer, and ultimately to a distributed-coding semantic pattern (see Fig. 9.5). Thomas trained the network on two artificial languages, via the backpropagation learning algorithm, which relies on thousands of learning instances to gradually adjust synaptic strengths of the network. The network was able to show behavior that reflected both the independence of lexical representations along with interference effects. Homographs were subject to within-language frequency effects, as has been shown previously to occur in human subjects (Gerard & Scarborough, 1989), and interlingual homographs were shown to have a disadvantage over cognates. This slower recognition of interlingual homographs has been identified under certain conditions in human data (Doctor & Klein, 1992), while other studies have found that human recognition of cognates is notably faster (Cristoffanini, Kirsner, & Milech, 1986; Gerard & Scarborough, 1989).

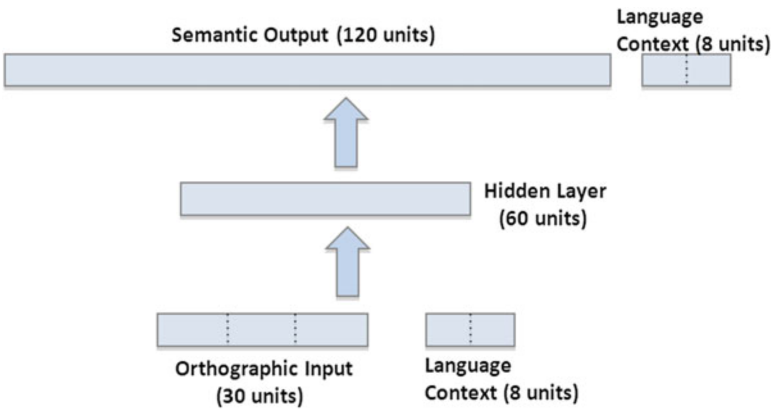


Fig. 9.5 Schematic diagram of the Bilingual Single Network Model, with distributed representations (Adapted from Thomas & Van Heuven, 2005. Copyright by Oxford University Press)

Bilingual Simple Recurrent Network

While the Bilingual Single Network includes binary nodes to explicitly encode which language the network is being exposed to at any one time, French's (1998) *Bilingual Simple Recurrent Network* relies solely on the statistical clustering of the inputs over time to allow the network to store distributed encodings of word meanings in separate regions of the hidden layer's state space. Inspired by Elman's (1990) simple recurrent network of monolingual language learning, the Bilingual Simple Recurrent Network receives inputs that are localist activations of individual words, 12 word nodes for Language Alpha and 12 word nodes for Language Beta (see Fig. 9.6). Any given word node then spreads its activation to the hidden layer, based on the strengths of its connections to those nodes. This pattern of activation in the hidden layer constitutes the way the network has encoded that word input, and the encoding is contextualized by the recurrent feedback from the activation pattern that the hidden layer had on the previous time step, stored in the "context" layer. This contextualized internalization of the current word input then spreads its activation to the output layer, thus constituting a prediction of some partially activated candidates for the likely next word in the incoming sentence. As in Thomas's Bilingual Single Network, the network's connections are trained via backpropagation, but unlike the Bilingual Single Network, the simple recurrent network

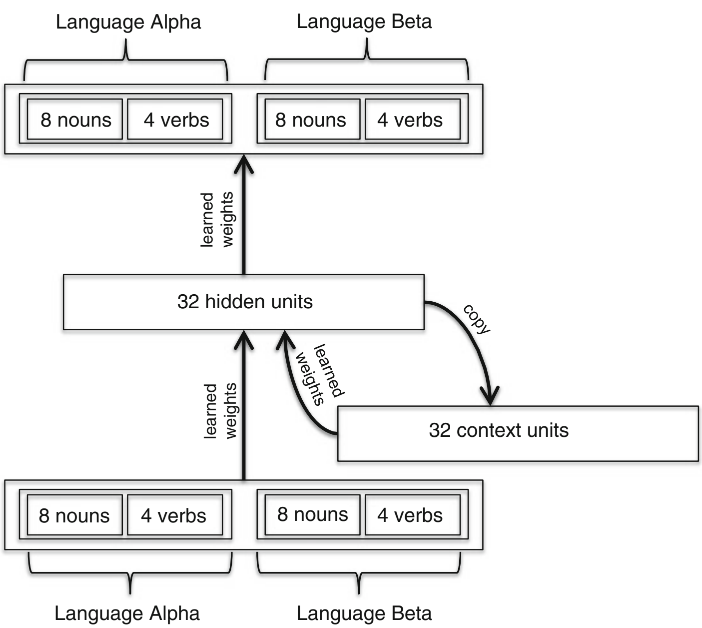


Fig. 9.6 Schematic diagram of the Bilingual Simple Recurrent Network, with distributed representations and a context feedback loop (Adapted from Thomas & Van Heuven, 2005. Copyright by Oxford University Press)

architecture allows it to use backpropagation in a less "supervised" manner. Because the output of the network is a prediction of the likely next inputs, the error signal necessary for training simply comes from waiting for the next actual input. Without using an explicit "teacher" to tell it what output it should produce for each input, the Bilingual Simple Recurrent Network can simply "eavesdrop" on a sequence of sentences, and its predictions of each next word will gradually get better (over the course of tens of thousands of iterations), because the pattern of connection strengths throughout the network is learning the sequential contingencies that make up the two languages.

By analyzing the state space of the hidden layer in these distributed-coding networks, one can see how, even though each node participates in coding for both languages to some degree, the set of activation patterns that one language employs is substantially non-overlapping with the set of activation patterns that the other language employs. This representation scheme allows these networks to simultaneously exhibit both language-independence findings and language-interference findings. Moreover, these distributed-coding models offer the promise of tracking the developmental process of gradually achieving proficiency in two languages, as seen in Li and Farkas's (2002; see also Zhao & Li, 2013) analysis of their Self-Organized Model of Bilingual Processing (SOMBIP), which is currently devoted to simulating spoken bilingual language processing, not yet reading.

There are, of course, some drawbacks to the current distributed-coding models of bilingual word reading that have been developed so far, which future extensions could perhaps improve. For example, the existing models are somewhat small in scale, have not been designed in such a way as to produce simulations of experimental laboratory data, and have not yet been analyzed for their insight into the developmental trajectory (as Li & Farkas, 2002, have done with their model of spoken bilingual word processing). While French (1998) was able to show that the Bilingual Simple Recurrent Network is capable of sorting out two languages from one another based entirely on the frequency of words that co-occur, it has not been directly compared to human performance in recognizing words. Future work on distributed single-network simulations of bilingual word reading is very much needed to fulfill the intriguing promise of these models and may benefit greatly from drawing inspiration from existing distributed-coding simulations of bilingual *spoken* word recognition (e.g., Li, Farkas, & MacWhinney, 2004).

Summary and Conclusions

There are multiple descriptive levels at which a set of cognitive phenomena may be modeled, and they do not always need to be placed in opposition to one another. The connectionist framework is frequently interpreted as a competing account of how cognition works, mutually exclusive with other accounts. However, this need not be so. Box-and-arrow models of bilingual word recognition (e.g., Kroll, Van Hell, Tokowicz, & Green, 2010), rule-and-symbol models (Doctor & Klein, 1992), and connectionist network models can actually coexist and learn from each other, as long as they are viewed as different levels of description that will each have their own respective domains of predictions that drive experimental testing forward. For example, treating a bounded region in the state space of a distributed neural network as roughly equivalent to a formal logical symbol is a statistical approximation that can, up to a point, accommodate a great deal of data (Dale & Spivey, 2005).

In fact, rather than taking place in theoretical battles, the lasting influence of the connectionist framework on all of cognitive science is probably better understood not as one that *fought and won some converts* but instead as one that *seeped into the mind-set of all cognitive scientists*—even those who were opposed to it. For example, as a result of three decades of connectionist influence, almost all cognitive scientists now embrace some form of interactivity between cognitive and linguistic subsystems. Moreover, as a result of three decades of connectionist influence, almost all cognitive scientists now embrace some form of integration between learning processes and innate constraints. Thus, regardless of what modeling paradigm one utilizes for one's research, there has been some cross-training going on between the various paradigms.

By the same token, it is probably best to encourage cooperation between localist connectionist accounts of bilingual word recognition and distributed connectionist accounts, rather than pitting them against one another. Bilingualism is a field that has benefited greatly from connectionist models of various types, being implemented in a fashion that allows both an existence proof of the theory of parallel distributed processing and relatively explicit predictions for each new experiment. In the case of localist interactive activation models, where each node has a label, and the connection strengths are set by hand, McClelland and Elman (1986) were very clear to note that *Each unit…stands for a hypothesis about the input being processed. The activation of a unit is monotonically related to the strength of the hypothesis for which the unit stands* (pp. 2–3). Thus, the localist nodes are clearly not intended as individual neurons devoted to entire concepts. Moreover, McClelland and Elman go on to add that *"In fact, interactive activation models like TRACE can be formulated in which each perceptual object is represented, not by a single unit, but by a pattern of activation over a collection of units"* (p. 78). Therefore, it is useful to treat each node in a localist network as a kind of shorthand for what, in the brain, is surely a distributed population of many neurons cooperating as an ensemble. Stone and Van Orden (1989) refer to this idea as *functional unitization*. The mechanism that implements the activation of a word representation may be a single unit in some simulated networks, but functionally it is approximately equivalent to a population of highly correlated neurons in a brain that work together to represent that word.

In the end, when one compares localist artificial neural networks (e.g., Grainger & Dijkstra, 1992) to distributed artificial neural networks (e.g., Thomas & Van Heuven, 2005), one need not necessarily see them as adversaries. Dijkstra et al. (2010) report a collection of ingenious behavioral experiments that appear to provide evidence in favor of a localist connectionist account of the bilingual language processing system. In a lexical decision task, greater orthographic and phonological similarity between an English target word and a known Dutch word produced a linear reduction in reaction times. Additionally, in a language-decision task (e.g., "which language does this word come from?"), greater similarity produced a linear increase in reaction times. Interestingly, in both tasks, genuine cognates exhibited a discontinuous deviation from the linear reaction time function, which the authors interpreted as naturally accommodated by the localist connectionist account, and inconsistent with a distributed connectionist account.

Rather than interpreting Dijkstra et al.'s (2010) findings as evidence that each word representation in the bilingual brain is implemented by a single neuronal unit, it may be useful to invoke Stone and Van Orden's (1989) functional unitization idea. Given that neurophysiological evidence strongly suggests that individual representations of objects and words are highly likely to be implemented by populations of neurons that become active in concert (e.g., Plaut & McClelland, 2010; Pouget, Dayan, & Zemel, 2000), Dijkstra et al.'s findings may indicate that, for bilingual word reading at least, these population codes are relatively sparse and have only partial overlap with one another. Indeed, among the range of cognitive processes implemented by the human brain, it could well be that some of them use representations that are highly functionally unitized (with very sparse population codes that are largely non-overlapping), and others use more broadly distributed population codes that exhibit substantial partial overlap with one another. This gives us a continuum between the localist connectionist camp and the distributed connectionist camp and may allow for more coexistence and cooperation among these differing modeling frameworks.

Ultimately, it is the role of all of these models to help tease apart what is undoubtedly a very complicated phenomenon. Each model is meant to embody a possibility; a potential way of explaining how humans acquire and manage multiple languages. It is important, however, to recognize that none of these models purport to fully explain how humans are able to utilize two languages, but rather that these are essentially suggestions to that effect. The more the model behaves like a human, demonstrating the same tendencies (such as faster recognition of words in the more recently used language), then the greater the probability that model is on the right track. The connectionist modeling approach to bilingual word reading, whether localist or distributed, is making steady advances in our understanding of how bilinguals use their two languages. The models are explicit enough to fit existing data and make clear predictions about new experiments, and can then be improved to fit the new data as well. A key important observation that the field has learned from these efforts is that bilingual word reading is neither the result of two completely independent lexicons, where one turns completely off while the other is active, nor the result of one monolithic network that has no architectural distinctions between the two language systems. The bilingual language processing system has an architecture that is somewhere in between those two extremes. As these modeling efforts progress over time, there continue to be strengths of one over another, but as a whole they continue to develop our understanding of just how complex bilingualism is and offer possibilities for discovering how it works.

List of Keywords

Bilingual Interactive Activation (BIA), Bilingual Simple Recurrent Network, Code-switching, Cognates, Connectionist models, Distributed representations, Emergentism, Functional unitization, Hidden layer, Interlingual homograph, Language decision task, Language mode, Localist Connectionist Models, Localist Interactive Activation Models, Localist representation, Neural networks, Parallel distributed processing, TRACE.

Review Questions

1. Have you ever tried learning another language? When reading in another language, do you have to translate it to English to understand it, or do you simply know what the words mean? How could a network model that translation process?
2. How quickly are you able to recognize words in other languages that use an alphabet similar to English as being foreign words, rather than words you are unfamiliar with?
3. Why do you think people who are bilingual can't always translate between languages on the fly, even though they can read or converse in either one? Is there a model that explains this better than the others?

Suggested Student Research Projects

1. Try testing a group of bilinguals on how fast they can recognize which language a word presented to them is in, and record their response time. If given several words in one language, does it make recognizing the other language slower?
2. Is reaction time different for people depending on how old they were when they learned a second language? To answer this question, choose about 20–30 high frequency English words (i.e., words with a frequency of 60 words per million or higher) from the database found in the website, http://expsy.ugent.be/subtlexus/, if English is the second language of your participants. Select the 20–30 words at random from the database that you can download. In addition to word frequency, you might want to also control for word length, and other variables such as concreteness (abstract vs. concrete words) and grammatical category (verb vs. noun). Remember, in order to establish causality, you have to control for as many variables or confounds (i.e., other third variables) as possible. You can select the two bilingual groups (early vs. late learners of L2) by simply asking them if they learned the L2 early (less than 12 years old) or late (after 12 years old). Once you have your groups and materials in order, visit http://step.psy.cmu.edu/scripts/

index.html and find a script on how to perform a lexical decision task. Scripts are fully functional and all you have to do is to modify and adapt them to your special needs. You will notice that each script is associated with a research article. Please read the article and follow the procedures as close as possible so that you can successfully develop and execute your experiment.

3. How much does visual similarity of languages impact the ability to switch between them? Are people able to switch more readily if languages have similar component features, such as English, French, and Spanish than visually diverse languages like Chinese, Korean, or Hebrew? Before you consider answering these questions with an experiment, perform a literature review and find out what the published literature has to say about this particular issue. What did you find out? How would you approach this issue with an experiment? Did you find a particular study that you would like to replicate to investigate this issue further?

Related Internet Sites

Debian Science: http://blends.debian.org/science/tasks/index
Nengo Neural Simulator: http://www.nengo.ca/
OpenModeller: http://openmodeller.sourceforge.net/
The Ultimate Neuroscience Software: http://neuro.debian.net/
TLearn: http://crl.ucsd.edu/software/

Suggested Further Reading

Grainger, J., & Jacobs, A. (Eds.), (1998). *Localist connectionist approaches to human cognition.* Hillsdale, NJ: Lawrence Erlbaum.

Heredia, R. R., & Altarriba, J. (Eds.), (2002). *Bilingual sentence processing*. Elsevier.

Altarriba, J., & Heredia, R. R. (Eds.), (2008). *An introduction to bilingualism: Principles and processes*. Taylor & Francis.

Heredia, R. R., & Altarriba, J. (Eds.), (2014). *Foundations of bilingual memory*. Springer.

References

Aisa, B., Mingus, B., & O'Reilly, R. (2008). The emergent neural modeling system. *Neural Networks, 21*(8), 1146–1152.

Anderson, J. A., Siverstein, J., Ritz, S., & Jones, R. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84*, 413–451.

Auer, P. (1999). *Code-switching in conversation*. New York: Routledge.
Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589–608.
Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology, 38*(3), 367–393.
Dale, R., & Spivey, M. J. (2005). From apples and oranges to symbolic dynamics: A framework for conciliating notions of cognitive representation. *Journal of Experimental & Theoretical Artificial Intelligence, 17*(4), 317–342.
Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.
Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496–518.
Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and language, 62*(3), 284–301.
Dijkstra, A., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum.
Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197.
Doctor, E. A., & Klein, D. (1992). Phonological processing in bilingual word recognition. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 237–252). Amsterdam: Elsevier.
Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*(2), 179–211.
Fodor, J. A. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese, 28*(2), 97–115.
French, R. M. (1998). A simple recurrent network model of bilingual memory. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 368–373). Mahwah, NJ: Lawrence Erlbaum.
French, R. M., & Jacquet, M. (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Sciences, 8*(2), 87–93.
Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(2), 305–315.
Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: Elsevier.
Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing. Explaining linguistics* (pp. 1–22). Malden, MA: Blackwell.
Harley, T. (2001). *The psychology of language*. Sussex: Psychology Press.
Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662–720.
Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638.
Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism (Cambridge, England), 13*(3), 373.
Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 59–85). Amsterdam: Elsevier.
Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks, 17*(8), 1345–1362.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: I. Evidence from associative priming by words, homophones, and pseudo homophones. *Journal of Experimental Psychology: General, 123*(2), 107–128.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism Language and Cognition, 6*(2), 97–116.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience, 23*(7), 1634–1647.

Munro, P. W. (1986). State-dependent factors influencing neural plasticity: A partial account of the critical period. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing* (Vol. 2, pp. 471–502). Cambridge, MA: MIT Press.

Plaut, D. C., & McClelland, J. L. (2010). Locating object knowledge in the brain: Comment on Bowers's (2009) attempt to revive the grandmother cell hypothesis. *Psychological Review, 117*(1), 284–290.

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience, 1*(2), 125–132.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275–280.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*(1), 60.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations*. Cambridge, MA: MIT Press.

Scheutz, M. J., & Eberhard, K. M. (2004). Effects of morphosyntactic gender features in bilingual language processing. *Cognitive Science, 28*(4), 559–588.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52*(2), 240–255.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568.

Spivey, M., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*, 281–284.

Stone, G. O., & Van Orden, G. C. (1989). Are words represented by nodes? *Memory & Cognition, 17*(5), 511–524.

Thomas, M. S. C., & Van Heuven, W. J. B. (2005). Computational models of bilingual comprehension. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 202–225). Oxford: Oxford University Press.

Vandeberg, L., Guadalupe, T., & Zwaan, R. A. (2011). How verbs can activate things: Cross-language activation across word classes. *Acta Psychologica, 138*(1), 68–73.

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458–483.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97*(4), 488–522.

Zhao, X., & Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism: Language and Cognition, 16*(02), 288–303.

Chapter 10
Second Language Sentence Processing: Psycholinguistic and Neurobiological Research Paradigms

Dieter Hillert and Yoko Nakano

Abstract Along with the introduction of different psycholinguistic and neurobiological research paradigms, we review some outcomes of second language (L2) research. In particular, we discuss the *probe recognition*, *cross-modal lexical priming*, *self-paced reading*, and *plausibility judgment* tasks, typically used to examine the temporal course of online sentence processing. Moreover, we present neurobiological methods such as *behavioral, electro- and magneto-physiological*, and *hemodynamic* measures. In considering the various types of research methods, we review the research on L2 sentence processing from extremely divergent perspectives. Finally, we discuss methodological issues to introduce the *status quo* of L2 in psycholinguistic and neurobiological domains.

Introduction

Sentence processing involves sequential or concurrent operations, in which an input string of auditory or visual information is segmented into small units (words). These units depend on each other to form larger units (phrases or clauses). When some are linked, they establish a dependency, that is, an asymmetric relationship between a head and its dependent (e.g., a verb and an object noun phrase [NP] or an NP and a clause modifying the NP) to construct a sentence structure.

D. Hillert, Ph.D. (✉)
School of Medicine, University of California, La Jolla,
San Diego, CA 92093-0505, USA
e-mail: dhillert@ucsd.edu

Y. Nakano
School of Human Welfare Studies, Kwansei Gakuin University,
1-155 Uegaharaichiban-cho, Nishinomiya City, Hyogo 662-8501, Japan
e-mail: y-k.nakano@kwansei.ac.jp

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_10

A sentence constitutes different types of units. Verb arguments, for instance, refer to constituents (e.g., object NPs) that are essential to form a sentence with the verb head, and they establish a dependency on their subcategorizing head (i.e., verb). Adjuncts are modifiers (e.g., relative clauses—a clause that modifies an NP), and a dependency is formed between the modifying element and the modified element (e.g., an NP).

Languages can be categorized into two types according to the position of heads—*head-initial* and *head-final languages*. In head-initial languages, the heads of a phrase and a clause tend to be in the initial position of the phrase and the clause. In contrast, in head-final languages, the heads tend to be at the end of a phrase and a clause. For instance, in the English verb phrase *kicked the ball*, the verb *kicked* is the head of the verb phrase. In contrast, in the Japanese verb phrase *booru-o ketta* "ball kicked," *booru* "ball" precedes the verb *ketta* "kicked." Since verbs hold the information about the structure whose head they become, the structure will be ambiguous until they appear. In other words, the sentence structures of head-final languages tend to be ambiguous until the appearance of the verb head at the end of a clause or a sentence while they are being processed. The structural differences of languages suggest that even if one model of sentence processing works on a particular language, it does not necessarily mean that it will work on other languages; hence, the models of sentence processing need to be tested on different types of languages.

To investigate how these syntactic dependencies are established, we need research techniques with which we can capture how these dependencies are formed during online processing. Thus, our psycholinguistic approach focuses on online methods, such as *cross-modal lexical priming* (CMLP) and *self-paced reading* (SPR) tasks.

One of the basic research topics in the field of sentence processing is how a dependency is established between the head and its subcategorizing argument NPs. For this purpose, sentences whose word order differs from the basic or canonical word order of a language—that is, the most common word order of a language—are often used because they contain a filler-gap dependency.

1. *Which woman$_i$ did a few boys approach __$_i$ to ask the way to the station?*

For instance, in sentence (1) the *wh*-phrase (i.e., the phrase that begins with wh-words like, *what, when, where, which*, and *who*), *which woman* is the object of the verb *approach*, but it is not in the syntactic position of the canonical word order. Since the right side of the verb, in which *which woman* originates, is blank, the phrase is indicated by an under-bar with an index, (__$_i$). *Woman$_i$ and __$_i$* are noted with the same index, *i*, which indicates that two items are related. Syntactic theory in the framework of generative grammar assumes that sentences are hierarchically structured and that a constituent that is dislocated into a different position leaves behind a trace, or a phonologically null copy of itself (Chomsky, 1995) at the position where it was located (__$_i$). Note that, although a trace is not exactly the same as a copy in generative grammar, the term *trace* may have been used conventionally (Chomsky, 1999). Psycholinguists often refer to the dislocated constituent as the filler and to the hypothesized trace position or copy as the gap (Fodor, 1978). In this vein, some previous psycholinguistic first-language (L1) studies assumed that the gaps of fillers are created at the purported base position (e.g., *Trace Reactivation Hypothesis*; e.g., Bever & McElree, 1988; Love & Swinney, 1998). For instance, English wh-phrases are fillers except when they are the subjects of a clause, because wh-phrases must move to the clause initial position in English (e.g., *which woman* in sentence 1 above). Therefore, when someone reads sentence (1) and encounters the wh-phrase *which woman*, he or she immediately predicts the presence of a gap for *which woman*; hence, he or she keeps the filler *which woman* in his or her memory, awaiting a potential position for positing a gap. According to the trace reactivation hypothesis, positing a gap means reactivating the filler in the reader's memory.

In contrast, some researchers assumed that a filler is directly associated with its subcategorizing verb without mediation of gaps (*Direct Association Theory*; e.g., Pickering, 1993; Pickering & Barry, 1991), or that a filler is semantically and directly associated with the verb (Carlson & Tanenhaus, 1988; Tanenhaus, Boland, Carlson, & Garnsey, 1989). It should be emphasized that the two views are not mutually exclusive; instead, each view can be interpreted as a description of different subprocesses that can concur in sentence processing (Nicol, 1993).

Behavioral studies on the late second language (L2) processing of filler-gap constructions have investigated several questions: Do L2 learners establish filler-gap dependencies in the same way that native speakers do? Do L2 learners create a gap, as reported in previous findings (Felser & Roberts, 2007; Marinis, Roberts, Felser, & Clahsen, 2005)? When is the filler semantically associated with its subcategorizing verb, and when is the filler associated with its subcategorizing verb via a mediation of gaps in L2 processing (e.g., Williams, 2006; Williams, Möbius, & Kim, 2001)? To what information are L2 learners sensitive in processing filler-gap constructions (e.g., Omaki & Schulz, 2011; Williams, 2006; Williams et al., 2001; see Felser, Cunnings, Batterham, & Clahsen, 2012 for eye-tracking experiments; see also Dallas & Kaan, 2008, for a review).

Methods and Studies

Equipment

The equipment necessary to run online behavioral experiments includes an experiment builder software program designed to present stimuli with the precision of milliseconds, a computer system equipped with this software, and hardware accessories for the stimulus software, including a response box with two or more keys. Well-known freeware programs are DMDX, PsyScope, Linger, OpenSesame, and Psychophysics Toolbox, which are based on MATLAB. Commercial programs include Superlab, E-prime, Presentation, and Experiment Builder. Some of these specify the required features of a computer, including the operating system, memory capacity, and sound and video cards. The software program for the presentation of stimuli enables researchers to measure online the response time to the stimuli in milliseconds.

Using a button box that is compatible with the software program minimizes the residual time and thus ensures that the data obtained in the tasks we describe are more reliable and precise than data collected without a button box.

Probe Recognition

After the segment-by-segment presentation (auditory or visual) of an experimental sentence, a probe (e.g., word, phrase, or picture) is presented at the end of the sentence. The participant is asked to judge whether the probe is part of the sentence and then press the *yes* or *no* button to indicate the response. The duration between the presentation of the probe and the button press is measured (Bever & McElree, 1988). The gap position needs to be processed before the probe is displayed at the end of the sentence; hence, the *probe recognition task* is not temporally sensitive to the linguistic region of interest. According to Just and Carpenter (1980), several types of cognitive processes, including syntactic and semantic processes, occur regarding the consistency of interpretation for individual referents of the sentence, as well as of the preceding texts (the end-of-sentence wrap-up process). This means that a filler is also retrieved from memory during the wrap-up process, and it facilitates a probe recognition, regardless of the presence of a gap; hence, the facilitation effect of the probe recognition could be due to the reactivation of a filler or the wrap-up process. Thus, the wrap-up effect may become a confounding factor that makes the interpretation of the results difficult.

Cross-Modal Lexical Priming

A typical design of the CMLP (see also Chap. 6) task used in the previous studies on filler-gap dependencies is as follows. Participants listen to an auditorily presented stimulus and simultaneously judge a probe or a target visually presented on a monitor. The judgment can be made on words (e.g., *blouse*) or non-words (*flouse*; Clahsen & Featherston, 1999; Nakano, Felser & Clahsen, 2002; Love & Swinney, 1998; Nicol & Swinney, 1989), or the animacy of pictures (Felser & Roberts, 2007). Probes are either semantically unrelated to the filler of the gap position (the control condition) or semantically related to the filler (i.e., the experimental condition). Alternatively, the probe type can be identical to the filler of the gap. Priming occurs when a preceding stimulus facilitates the participant's response to a word or concept. For instance, when a probe (e.g., *nurse*) is preceded by a semantically related stimulus (e.g., *doctor*), shorter lexical decision latencies are obtained, compared to a preceding stimulus (e.g., *butter*) that is semantically unrelated to the target stimulus. This effect is known as *the priming effect*.

In previous studies, the CMLP was utilized to investigate the reactivation of fillers at the hypothesized gap position. The stimulus sentences that contained a filler-gap dependency were auditorily presented via headphones. The participants were instructed to make judgments on the lexicality or animacy of the probes, which were visually displayed while the auditory sentence presentation continued. In this experiment, the probes or targets were pictures. For instance, Felser and Roberts (2007) presented experimental sentences, such as (2):

2. *Fred chased the squirrel to which the nice monkey explained the game's*$_{\#1}$ *difficult rules_* $_{\#2}$ *in the class last Wednesday* (note: the antecedent of the *wh*-pronoun *whom* is *squirrel*).

The picture probes were either identical to the filler (*squirrel*) or semantically unrelated to the filler (*toothbrush*). If the filler is retrieved from memory, the presentation of a semantically associated or identical probe could trigger a priming effect, regardless of the position in the sentence, that is, at both probe points, depicted by subscripts #1 and #2. The magnitude of the priming effects would be larger at #1 than at #2 if a gap were not created at the hypothesized trace position (#2) because of the decline of the activation level. In contrast, if a gap were created at the hypothesized trace position, the activation level of the filler would increase at the gap, which appears after the control position #1; hence, the magnitude of priming would be larger at point #2 than point #1.

Felser and Roberts (2007) found a priming effect at the purported trace position (#2) but not at the control position (#1) in a native-English-speaker group. In contrast, non-native speakers (Greek speakers with advanced L2 English competencies) revealed priming effects at both positions but no significant difference in the priming magnitude. The priming effects found in the native-speaker group were interpreted as the active creation of a gap by the native speakers, whereas no indication of positing gaps was found in the L2 learner group. Clahsen and Felser (2006a, 2006b) proposed a hypothesis for L2 processing, based on previous studies in various psycholinguistic subfields, including the present study, and they referred to it as the *Shallow Structure Hypothesis* (SSH). Briefly, this hypothesis suggests that language learners who started learning a new language after puberty could construct argument-predicate semantic dependencies, but they are less sensitive to syntactic information than native speakers of the language, and they have difficulties in constructing hierarchical structures that are as complex as those composed by native speakers.

Self-Paced Reading

In the SPR paradigm, sentences are segmented into the linguistic units of interest (e.g., word or phrase) and are presented unit by unit on a computer monitor. Participants read the displayed segment as fast and as accurately as possible and then press a key or computer button to trigger the display of the next unit. The participant reads the displayed units (e.g., a sentence) at their own self-pace. The time taken to read each unit is measured and recorded in a memory device, which is referred to as reading latencies. The stimuli are typically presented from left to right in the *moving-window* presentation (see also Chap. 5), in which the units previously presented on the monitor disappear when the next unit appears. Either the raw reading latencies or the residual reading times undergo statistical analyses. Residual reading times are distinguished by the raw data and the predicted time. They can be obtained in two steps: first, by computing the linear equation to predict the reading time as a function of word length, and second, by subtracting the predicted time from the raw data. Residual reading times allow the adjustment of the nonlinearity of the data (Trueswell, Tanenhaus, & Garnsey, 1994).

The SPR paradigm enables us to measure online access to a particular unit or segment in a sentence by means of recording the reading latencies while the participant is processing the segment. Researchers can compare the reading latencies of a region that includes a critical word or segment under the control and experimental conditions. It is assumed that longer reading latencies in the experimental condition as compared to the control condition reflect difficulty in processing the sentence. A self-paced listening task is an alternative method for younger participants with limited literacy (e.g., see Felser, Marinis, & Clahsen, 2003). This listening task has been used to investigate the *filled-gap effect*. This effect occurs when a listener anticipates filling a position that has not yet appeared in the form of a gap, but the position turns out to be already filled by another constituent.

3. (a) *My brother wanted to know if Ruth will bring us home to Mom at Christmas.*
   (b) *My brother wanted to know who Ruth will bring __ home to Mom at Christmas.*
   (c) *My brother wanted to know who Ruth will bring us home to __ at Christmas.*

In one study, Stowe (1986, p. 234) presented sentences, as shown in (3), by using a word-by-word SPR task to native speakers of English. The results showed that the reading latencies for (3c) were longer than the latencies for (3a) or (3b) at the object position of the transitive verb *bring*. Because of the transitivity of the verb *bring*, the reader expects the appearance of the object position. The *wh*-phrase *who* is a potential object of the verb *bring*; hence, it is plausible that it triggered the gap creation at the object position. In fact, the appearance of *us* in (3c) indicates the incorrectness of the interpretation; namely, the predicted gap has already been filled with *us*, leading the reader to a subsequent reanalysis, which is an example of the filled-gap effect. In contrast, in (3b), the purported gap position was not filled with an NP, and the position could be filled. Sentence (3a) does not have a potential filler object; therefore, no incongruence occurred between a created gap and the word that has already filled the position. These results also indicate that although readers could wait for the appearance of the actual gap position, they actively created a gap as soon as they found a potential gap position. The type of processing observed in (3) is called the *active filler strategy*.

Another example illustrates the SPR study that investigated the establishment of filler-gap dependencies. There can be more than one gap in a filler-gap dependency because in some syntactic theories, it is assumed that a filler moves in a cyclic manner from the base position and lands on a particular position of a sentence and then moves to a different position. Through these movements, the constituent leaves more than one copy of itself behind. For instance, in sentence (4), the filler *who* leaves two traces—an intermediate trace e'_i and the trace e_i at the base position. In an SPR experiment, Marinis et al. (2005) also tried to find evidence for the intermediate trace.

4. (a) *The nurse who*$_i$ *the doctor argued* e'_i *that the rude patient had angered* e_i *is refusing to work late*.
   (b) *The nurse thought the doctor argued that the rude patient had angered the staff at the hospital*.

It is assumed that sentences such as (4a) contain an intermediate trace of a *wh*-phrase *who*. The hypothesized position for the intermediate trace (e'_i) is between the verb *argued* and the complementizer *that*. It is predicted that the appearance of *that* triggers the creation of the intermediate gap for *who*. In contrast, because in (4b) no *wh*-phrase appears, no intermediate gap will be created by the appearance of the complementizer *that*. Therefore, longer reading latencies are predicted at *that* in the sentence containing hypothesized intermediate traces (sentence 4a), compared to the sentence with no hypothesized intermediate traces. Marinis et al. (2005) found longer reading latencies for the complementizer in the native-speaker group but not in any of the L2 groups in the study (Greek, German, Chinese, or Japanese). The position of the reading latencies was identical between (4a) and (4b); the complementizer *that*, and the phrase before and after it were also identical—*the doctor argued that the rude patient had angered*. The only difference is the presence of the intermediate trace e'_i. Therefore, longer reading latencies at the complementizer *that* in (4a) in the native-speaker group reflected the time needed to postulate the intermediate trace. In contrast, no difference was found in the learner group, which indicates that it did not postulate any intermediate trace.

Plausibility Judgment

The plausibility judgment (or the stop-make-sense [SMS]) task requires the participant to judge the plausibility of a sentence while performing word-by-word SPR. This informative task investigates the position where the sentence stops making sense or becomes implausible. The participant is asked to press a button as soon as possible when he or she feels the sentence is implausible or stops making sense. Thus, it is possible to determine the position at which the thematic argument structure of the verb is saturated by the filler (Boland, Tanenhaus, & Garnsey, 1990; Boland, Tanenhaus, Garnsey, & Carlson, 1995), as well as the position at which the semantic and pragmatic compatibility of the filler with the verb is evaluated (i.e., the semantic goodness-of-fit evaluation; Felser et al., 2012; see also Traxler & Pickering, 1996). For instance, Boland et al. (1990) presented sentences, such as *Which food* (*book*) *did the boy read in class*? to native speakers of English. The *wh*-filler *which food* is a semantically unlikely direct object of the verb *read* because the verb *read* assigns a thematic role, not to an edible object but to a readable object; hence, it was predicted that participants would press the SMS key for the unreadable and implausible object *food* at the verb *read*. Boland et al. interpreted these results as indicating that the filler was directly and immediately associated with the thematic role of the verb, without involving any gaps. Subsequent studies interpreted these results as reflecting complex semantic processes in associating the filler and the verb. The implausibility judgment task has also been used in L2 studies (Williams, 2006; Williams et al., 2001).

For example, Williams et al. (2001) compared the rates of SMS decisions at the verb in sentences, such as *Which river* (*girl*) *did the man push the bike into late last night*? The *wh*-phrase *which girl* is a plausible object of the verb *push* (the plausible-at-V condition), whereas *which river* is an implausible object of *push* (the implausible-at-V condition). They found a higher rate of SMS decisions at the verb in the implausible-at-V condition than in the plausible-at-V condition in both the native-speaker group and the groups of proficient L2 English speakers with the L1 background of a *wh*-movement language (German) and a *wh*-in-situ language (Chinese and Korean). According to Williams et al., the results indicated that both native and non-native speakers utilized the active filler strategy and created gaps. With regard to the reading-time data, both native and non-native groups read more slowly at the noun *bike* in the plausible-at-V than in the implausible-at-V conditions. However, no difference was found at the verb in either of the groups. Only the native speakers showed a slow-down at the post-verbal determiner (*<u>the</u> bike*) in the implausible condition as compared to the plausible condition. The tendency was reversed at the noun *bike*. The non-native-speaker groups showed no difference at the determiner between the conditions. The appearance of the determiner, after the verb, indicated that the potential gap position had already been filled by another noun phrase. Williams et al. (2001) argued that the native speakers' fast responses to the determiner could be ascribed to their sensitivity to the syntactic cue. That is, because of the appearance of the noun, the non-native speakers needed additional information by the appearance of the noun and the plausibility, in order to respond differently to the two conditions. Although both the native and non-native speakers used the active filler strategy and judged plausibility, the balance of the syntactic and semantic cues seemed to differ between native and non-native speakers. To address this issue further, Williams (2006) conducted an additional plausibility judgment study (Experiment 1). In this experiment, the distance between the post-verbal determiner and the noun was increased by inserting words (e.g., *the very nice bike*) to examine further the decision timings for implausibility. The results were similar to those in Williams et al. (2001). The implausible-at-V condition yielded more SMS decisions than the plausible-at-V condition did at the verb for both native and non-native groups, and both groups read more slowly at the intensifier (*very*) in the plausible-at-V condition than in the implausible-at-V condition. Williams (2006) argued that both native and non-native speakers employed the same syntactic processing strategy and were sensitive to plausibility. Williams also pointed out that the results could have been influenced by the unnaturalness of the task in two ways. First, participants encountered implausible sentences frequently during the task, in response to which the participants devised a strategy to delay decisions.

Second, SMS task sentences were presented word by word, and participants were required to make plausibility judgments incrementally. The incremental plausibility judgment is not forced in normal reading; hence, the results of the SMS task do not inform us about the processing when the incremental plausibility judgment is not required in natural reading.

Williams conducted a second experiment using an SPR task followed by comprehension and memory questions, which was free from the obligatory plausibility judgment. Although participants were different in the first and second experiments, they had comparable language proficiencies; however, the results of the two tasks differed. The non-native speakers read more slowly than the native speakers. The locus of the plausibility effect varied according to the participants. The participants in each group were divided into high- and low-memory subgroups according to their scores on the memory task. The high-memory native speakers revealed slower reading times at the determiner, and the low-memory native speakers showed slower reading times at the post-verbal noun in the plausible-at-V than in the implausible-at-V conditions. The high-memory non-native speakers showed slower reading times at the preposition in the implausible-at-V than in the plausible-at-V conditions. The low-memory non-native speakers did not show any plausibility effects. Williams (2006) argued that the native and non-native speaker participants processed the target sentences similarly, but the varying timings of the effects could be ascribed to individual differences in cognitive factors, such as working memory and motivation, which may or may not be present according to the task requirement. However, Felser et al. (2012) pointed out the possibility that the slower reading times in the L2 learner groups may not have been caused by the delay of the SMS decisions but the delay of the filled-gap effects. Indeed, it is difficult to distinguish the effect of syntactic gap-filling processes from the effect of semantic goodness-of-fit evaluation in Williams (2006) and Williams et al. (2001). Felser et al. (2012) further suggested that in Williams and colleagues' studies, the patterns of the SMS decisions were the same between native speakers and L2 learners. The L2 learners could respond immediately to semantic information of plausibility, and the L2 reading times were delayed, compared to native speakers' reading times, because the L2 learners were less sensitive than the native speakers were to structural information.

Sensitivity to Structural Information in L2 Processing

Clahsen and Felser (2006a, 2006b) reviewed a wide range of published studies, including L1 studies on adults and children and studies on late L2 learners, in which the aforementioned various research techniques were used. They proposed, as previously discussed, the SSH for L2 sentence processing. According to this hypothesis, L2 learners can form argument-predicate structural representations based on lexical and semantic information, but they are less sensitive than native speakers to syntactic information, and they have more difficulty in computing detailed hierarchical representations in real time. For instance, when an L2 learner reads or hears a sentence that contains a filler and its corresponding gap, he or she needs to construct a hierarchical structure that is detailed enough to find a gap position. However, if the learner is not able to construct detailed sentence structures, resulting in shallow structures, he or she cannot find any gap sites for the filler. Williams (2006) and Williams et al. (2001) argued that both native speakers and L2 learners syntactically process sentences in the same way but that semantic processes are affected by task-related and cognitive factors, such as memory capacity, which varies individually. Omaki and Schulz (2011) argued for the postulation of a gap in the case of L2 learners.

Omaki and Schulz (2011) utilized the implausibility paradigm to investigate gap creation by native speakers of English and Spanish speakers of L2 English. In addition, the experimental sentences contained a clause that began with a wh-phrase, such as *who* in (6c, d). It has been shown that a constraint can prohibit a constituent from moving out of a particular region (Ross, 1967). The region is metaphorically referred to as an *island*. The types of constituents that become islands vary, depending on the language. In English a wh-clause can be an island and is referred to as wh-island. For instance, although the object noun phrase (*which novel prize*) can move into the sentence initial position and form a wh-question, as in (5b), the same constituent cannot move out of the wh-phrase, as in (5d). The asterisk (*) indicates that the sentence is ungrammatical.

5. (a) *The professor won the novel prize in physics*.
   (b) *Which novel prizei did the professor win __i?*
   (c) *Mary admires the professor who won the novel prize in physics.*
   (d) **Which novel prize$_i$ does Mary admire the professor who won __$_i$?*

In Omaki and Schulz (2011), the four different types of sentences shown in (6) were presented in a word-by-word SPR task.

6. (a) Non-island, implausible: *The city that the author wrote regularly about was named for an explorer*.
   (b) Non-island, plausible: *The book that the author wrote regularly about was named for an explorer.*
   (c) *Island, implausible: The city that the author who wrote regularly saw was named for an explorer.*
   (d) *Islan*d, plausible: *The book that the author who wrote regularly saw was named for an explorer*.

The plausibility of the combination of filler (*the city* vs. *the book*) and a verb (*wrote*) and the constraint (non-/island constraint) were manipulated in the quadruplets. In (6b) the plausible filler (*the book*) can be associated with the verb *wrote*, whereas in (6a) the implausible filler (*the city*) cannot be associated with *wrote*. In (6c, d), *who* indicates the presence of a wh-island. This means that neither *the city* nor *the book* is moved out of the clause *who wrote*; hence, it cannot be associated with the verb *wrote*. In the native-speaker group, in the critical region *wrote*, the implausible non-island condition (6a) yielded slower reading times than the plausible non-island condition (6b).

The one or two regions that follow the critical region are referred to as spillover regions. It is often the case that the effect of a particular region appears in the regions that follow it; the delayed effect is metaphorically referred to as a *spillover effect*. In the experiment, the spillover region was *regularly*, and it indicated the same result in the critical region; namely, the implausible non-island condition (6a) yielded slower reading times than the plausible non-island condition (6b) did. In contrast, there was no difference between the plausible and implausible island conditions. The L2 learners showed the same pattern of results in the spillover region. The slower reading time was interpreted as indicating that both native speakers and learners actively generated a gap at the verb in the non-island conditions and that both participant groups experienced processing difficulty in the implausible non-island condition because of a plausibility mismatch. The lack of difference in the island conditions could be interpreted as indicating that the island constraints blocked the dependency formation in both native- and non-native-speaker groups.

Omaki and Schulz (2011) argued that not only native speakers but also late L2 learners could construct structural representations with rich grammatical details because the sensitivity to the relative-clause island constraints required the learners to construct hierarchical structure representations. Their findings, however, did not necessarily reject the SSH. Instead, they proposed a weaker view of the SSH, which assumes that L2 learners produce shallow structures more often than native speakers do under certain conditions, such as when the learner's L1 does not share some grammatical properties with the L2. They also suggested that L2 processing is cognitively demanding because several processes are concurrent; therefore, the parser tries to reduce the burden by adopting shallow structures.

As argued earlier, Williams (2006) and Williams et al. (2001) had difficulty in dissociating the effects of syntactic and semantic processes and in judging whether both processes occurred during the initial parsing or only one of them occurred. Note that Omaki and Schulz (2011) also had difficulty in distinguishing the effects caused by syntactic and semantic subprocesses, such as examining whether a verb and its arguments semantically and pragmatically matched well and fulfilling the number of arguments that a verb controls (Felser et al., 2012). As Pickering (1993) and Pickering and Barry (1991) pointed out, both gap creation and semantic association between the filler and the argument structure of the verb may occur at the offset of the verb. The different reading times between the non-island plausible and implausible conditions could imply the occurrence of the semantic subprocess, but they are not necessarily indicative of the postulation of gaps.

Felser et al. (2012) conducted two eye-tracking experiments, each of which examined the semantic goodness-of-fit evaluation for matching the filler object with its subcategorizing verb and the formation of a syntactic filler-gap dependency by postulating a gap that corresponds to a filler. The results of the two experiments differed between L1 speakers and L2 learners, suggesting different timings in utilizing different types of information between L1 speakers and L2 learners. The example sentences in (7) below were used for the plausibility effect as a diagnostic for the formation of a semantic dependency (Experiment 1), and those in (8) were used for the filled-gap effect as a diagnostic for the formation of syntactic filler-gap dependency

(Experiment 2). In both experiments, the participants were instructed to read a short text that constituted a lead-in sentence and a target sentence. The texts were displayed on the monitor, and when the participant pressed a button, a yes-or-no question about the text was displayed for two-thirds of the materials.

7. *The new shampoo was featured in the popular magazine.*
   (a) No constraint, plausible: *Everyone liked the magazine that the hairdresser read extensively and with such enormous enthusiasm about before going to the station.*
   (b) No constraint, implausible: *Everyone liked the shampoo that the hairdresser read extensively and with such enormous enthusiasm about before going to the station.*
   (c) Island constraint, plausible: *Everyone liked the magazine that the hairdresser who read extensively and with such enormous enthusiasm bought before going to the salon.*
   (d) Island constraint, implausible: *Everyone liked the shampoo that the hairdresser who read extensively and with such enormous enthusiasm bought before going to the salon.*

In Experiment 1, the target sentences contained a relative clause that began with *that*. The noun phrases (*the magazine* and *the shampoo*) that preceded *that* were fillers. The earliest potential gap position was immediately after the verb *read*. All the sentences were globally plausible, but the implausible sentences were locally implausible at the verb because of the mismatch between the filler and the type of filler required by the verb. Sentences (7a, b) contained no *wh*-islands, while sentences such as (7c, d) contained another relative clause embedded in the *that-relative* clause. The antecedent NP could not be extracted from the *wh*-clause in (7c, d). Felser et al. (2012) analyzed three types of measurements (i.e., first-pass reading times, regression path durations, and re-reading times). Briefly, first-pass reading time is *the summed duration of all initial fixations on a region until that region is exited to either the left or right*. Regression path duration is defined as *the sum of all fixations on a region until this region is first exited to the right*, and re-reading time is *the summed duration of all fixations on a region after it first exited to either the left or right* (Felser et al., 2012, p. 80). It is assumed that different measures reflect different cognitive stages of processing. First-pass reading times reflect the initial stage of processing, and regression path durations and re-reading times reflect later stages than first-pass reading times do (Pickering, Frisson, McElree, & Traxler, 2004). The L1 speakers showed the main effect of constraint (no-constraint and island constraint conditions) for the first-pass reading time. The first-pass reading time was shorter at the verb in the island constraint than in the no-constraint conditions, but no interaction of constraint and plausibility (the plausible and implausible sentences) was found. The re-reading time indicated the interaction of constraint and plausibility. The re-reading times for the implausible sentences were longer than for the plausible sentences in the no-constraint condition. The plausibility effect was not found in the constraint condition. The L2 participants showed the main effect of constraint as well as the interaction of constraint and plausibility.

Their first-pass reading time was shorter at the verb in the island constraint than in the no-constraint conditions. It was also shorter in the plausible condition than in the implausible conditions. There was no significant difference between the plausible and implausible sentences in the island constraint condition. In the spillover region, the effect of the participant groups was not found. The interaction between constraint and plausibility was found for regression path duration and re-reading time; the implausible sentences yielded longer reading times than the plausible sentences did in the no-constraint condition, but no such difference was found in the constraint condition.

The results differed between the L1 and L2 speakers. The L1 speakers' response to the syntactic constraint appeared in the first-pass reading times, their response to the plausibility appeared in the re-reading time, and the L2 speakers' responses to the syntactic constraint and to plausibility appeared in the first-pass reading time. The results indicated that the timings in responding to the plausibility and the syntactic constraint differed between the two groups of speakers. The L1 speakers followed the syntax-first strategy, whereas the L2 speakers responded to the semantic plausibility and syntactic constraint at the same time.

8. There are all sorts of magazines on the market.
   (a) No constraint, gap: *Everyone liked the magazine that the hairdresser read quickly and yet extremely thoroughly about before going to the beauty salon.*
   (b) No constraint, filled gap: *Everyone liked the magazine that the hairdresser read articles with such strong conclusions about before going to the beauty salon.*
   (c) Island constraint, gap: *Everyone liked the magazine that the hairdresser who read quickly and yet extremely thoroughly bought before going to the beauty salon.*
   (d) Island constraint, filled gap: *Everyone liked the magazine that the hairdresser who read articles with such strong conclusions bought before going to the beauty salon.*

The materials in Experiment 2 were the same as in Experiment 1, except the verb (*read*) was followed by an adverbial phrase (*quickly*) in the gap condition (8a, c) and by a noun phrase (*article*) in the filled-gap conditions (8b, d). If the participants tried to link the filler and the verb by creating a gap, they would see that the predicted gap position had already been filled. Therefore, the reading time would slow down because of the processing difficulty. The results of Experiment 2 were as follows: L1 speakers showed an interaction between gap (the filled-gap and gap sentences) and constraint (no-constraint or island constraint conditions). The filled-gap sentences were read more slowly than the gap sentences were only in the no-constraint condition, and no such difference was found in the island constraint condition. This pattern was found in the first-pass reading time, the regression path duration, and the re-reading time in the critical region and for the regression path duration and re-reading time in the spillover region. These results suggest that the wh-phrase *who* indicated the presence of a wh-island in the island constraint condition, and it blocked the creation of a false gap in both the gap and the filled-gap conditions. The results also suggest that because no wh-phrase blocked the creation of gaps in the no-constraint condition, a filled-gap effect was observed in (8b). In contrast to the L1 speakers, no interaction between gap and constraint was found for the L2 learners in the critical region. A significant interaction between gap and constraint was found only in the re-reading time in the spillover region.

The implausibility sentences produced longer first-pass reading times than the plausible sentences in the no-constraint condition for the L2 speakers in Experiment 1. The L2 speakers also read the filled-gap sentences more slowly than the gap sentences in the no-constraint condition, but a difference between the filled-gap and gap sentences was found in the island constraint condition only in the later processing stage (i.e., the re-reading time in the spillover region) in Experiment 2. In contrast to Williams (2006) and Williams et al. (2001), Felser et al. (2012)) argued that when the filler was associated with the verb, the semantic goodness-of-fit was evaluated at the initial stage, and the integration of the filler into a structure was conducted semantically. The L2 learners were also not sensitive to the structural information so that the gap-filling based on the structural information was not initially conducted. The plausibility effect was found later in the L1 speakers than in the L2 speakers in Experiment 1, but the filled-gap effect was found at the initial stage (i.e., first-pass reading time) in Experiment 2. These results indicate that the L1 speakers first posited gaps based on the structural information and later evaluated the semantic goodness-of-fit between the filler and the verb.

In Felser et al. (2012), sensitivity to the island constraint was found in both the L1 and the L2 speakers. Omaki and Schulz (2011) argued that in their study, the slow-down in the implausible condition, compared to the plausible condition in the no-constraint condition and the lack of the plausibility effect in the island constraint condition, indicated sensitivity to the *wh*-islandhood and the gap creation in the L2 learner group. The results were compatible with Felser et al. However, the timing of gap creation in the L2 processing is problematic in Omaki and Schulz because they did not directly test the gap-filling process by using a diagnostic such as the filled-gap effect.

Methodological Considerations

The tasks described so far have enabled psycholinguists to measure participants' response times during the time course of sentence processing. This property of time-sensitiveness meets the need for research to capture online operations. The cross-modal priming task measures the response times that correspond to different degrees of activation in the target and control items at a particular point during the online sentence comprehension process. The SPR task is sensitive to difficulty in online processing. The plausibility judgment paradigm is useful for identifying the location at which participants detect semantic plausibility while they are processing a sentence.

Every task, however, has some limitations. In the cross-modal priming task, it is difficult to analyze the complete time course of sentence processing. This task can detect the active representation of a gap only at the relevant probe points. This task requires a pair of words that elicit lexical decision latencies when they are presented in isolation. Because word associations and frequencies may vary, particularly between native speakers and L2 learners, it is difficult to counterbalance the lexical decision latencies of target words, which are semantically related and unrelated to the prime word presented in the sentence. The cross-modal priming task is a dual task; hence, there may be a case in which the task is cognitively too demanding for learners. Moreover, the SPR task has limited sensitivity to process difficulty and temporal resolution. For example, Miyamoto and Takahashi (2002) compared reading latencies in a pair of canonically structured and scrambled sentences (In Japanese, these are comparable filler-gap constructions). Miyamoto and Takahshi found significant differences in latencies in the pairs in which modifiers were inserted to increase the distance between filler and gap, but they found marginal significance in the pairs in which the distance between the filler and the gap was shorter. This means that the processing cost of the shorter condition was too small for the SPR task to detect. The sentences were segmented into units that had a certain length; hence, it was difficult to determine which part within a segment caused processing difficulty. Further, the SPR technique requires participants to read sentences in an *unnatural* way because participants are *forced* to read function words, which they tend to skip in normal reading (Rayner & Sereno, 1994). Moreover, the participants are unable to go back to the initial parts of the stimuli, and additional demands are imposed on their working memory (Dallas & Kaan, 2008). Finally, the segment that indicated longer reading latencies under one condition than in another does not necessarily identify the source of difficulty. The reading latency of a particular segment may reflect several effects, which are difficult to separate. For instance, participants continue to process a previous segment while they are reading subsequent segments, and the effect of a particular segment often appears downstream but not at the region of interest (i.e., the spillover effect; Harberlandt & Bingham, 1978; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). The plausibility judgment task has the same limitations; the presentation of stimuli is the same as in the SPR task. The task can also be unnatural with respect to two additional points: (1) The experimental materials used for the plausibility judgment task contain more implausible sentences, compared to normal reading and other reading studies; (2) the task also forces participants to evaluate plausibility incrementally (Williams, 2006).

The studies reviewed here suggest two critical points. One concerns the importance of using different techniques for investigating a particular phenomenon. If the results obtained from a few different experimental methods consistently support a particular hypothesis, the hypothesis is more reliable than that supported by the results obtained from only one experimental method. For instance, off-line tasks are not very informative about how dependencies are created during online processing, but they can indicate the final decision for a construction that includes a structurally ambiguous constituent. In the sentence, *John saw the girl of the mother who was holding a large umbrella*, the relative clause *who was holding a large umbrella* could modify both *the girl of the mother* and *the mother*; hence, the sentence is structurally ambiguous. The noun phrase most often chosen as the antecedent of the relative clause is determined by asking the participants whom they think is holding the large umbrella. The results indicate the final decision in choosing the antecedent of the relative clause. Online experimental methods, such as SPR tasks and eye-tracking techniques can reveal the online decisions made to choose the antecedent of a relative clause. Therefore, if the results of both off-line and online tasks are considered, the results will provide a more comprehensive picture (Lieberman, Aoshima, & Phillips, 2006). The second point is that not only is the choice of research technique informative with respect to the occurrence and timing of subprocesses in parsing but also it is important with regard to the combinations of different linguistic effects (Felser et al., 2012).

Moreover, the linguistic environment of L2 learners varies across countries. In most countries, L2 learners do not find many opportunities to use the target language outside their language classroom, whereas in some countries there are more opportunities to speak the second language outside that setting. Thus, the proficiency levels of L2 learners may affect their ability to comprehend sentences. Therefore, it is important to measure individual L2 competencies and to take this information into consideration when analyzing sentence processing data that are obtained using online methodologies. Furthermore, in the design of online L2 sentence processing studies, control tasks should be included in order to obtain a profile of L2 proficiency (e.g., placement tests).

Neurobiological Research Paradigms

The present section discusses how the human brain processes non-native or native-like (i.e., L2) languages as compared to native languages (i.e., L1). The particular configuration of research in L2 is that it is impossible to examine an L2 isolated and independent of a person's native language. It is exactly this configuration that raises numerous questions about the cortical structures and dynamics involved in sentence processing. For instance, to what extent does our brain process non-native sentence structures differently from native sentence structures? Does a possible processing difference between L1 and L2 depend on the degree of structural similarity and/or on a certain stage of brain growth? Is the number of languages our brain can handle limited? What are the benefits and/or the downside of speaking more than one language? Although we will address these and other questions in the following discussion, our focus here is on reviewing and discussing the temporal online parameters and spatial locations and connections involved in L2 as compared to L1 processing. As the present volume focuses on the introduction of methods used in language research, we will organize this section according to the *status quo* of the most common methods and techniques applied to examine the electrophysiological and neural activities involved in sentence processing. It is important to consider that we do not favor any particular method and technique, as all contribute to knowledge gain about the neural correlates of language processing. From a methodological viewpoint, the *what*-question precedes the *how*-question, that is, first we ask what we would like to investigate, and then we ask which means are available to investigate our statements and hypotheses. As the present volume is about the means, the methods and techniques available, we provide in the following an overview of current neurobiological approaches complemented by variants thereof. Finally, it should be noted that the techniques are identical for native and second language research.

Methods and Studies

Before the introduction of broadly used electrophysiological and neuroimaging techniques in the 1980s and 1990s, observations and analyses of language behavior in neurologically impaired bilingual patients (i.e., lesion studies) served as the main source for drawing conclusions about the bilingual brain. This occurred not only because of theoretical interests, but was a clinical necessity. More than half of the world population can be considered multilingual, and therefore patients suffering from bilingual language disorders is not an exception, but represents the majority of cases. The systematic diagnosis of L2 disorders in aphasia began with the use of the *Bilingual Aphasia Test* (Paradis & Libben, 1987). Specific psychometric and linguistic criteria were set for adapting the English version to other languages. Beyond standard tests, researchers evaluated language disorders in a customized fashion by presenting test material in a paper-and-pencil (off-line) format. Thus, this neurolinguistic method described language disorders in relation to clinical symptoms and/or syndromes and linked these patterns to the lesion site assessed by X-ray computed tomography (CT scans). It is apparent that this dual approach has its limits, as it neither informs about the specific cortical regions or circuitries involved in L2 processing nor does it consider other cognitive functions such as working memory, temporal parameters, cognitive costs, and world knowledge representations. Thus, it is extremely difficult to draw general conclusions about the language–brain relationship by observing and analyzing the recovery process after aphasia. However, some facts should be mentioned in this context.

Behavioral Measures

In Fabbro's (2001) study, for example, the recovery patterns of 20 right-handed bilingual Italian-Friulian aphasic patients, who acquired their second language in young childhood (5–7 years of age), revealed the following: approximately 65 % showed parallel recovery in both languages, 20 % were more impaired in their L2, and 15 % were more impaired in their L1. Interestingly, Fabbro could not determine a specific factor responsible for the recovery patterns; neither the variables lesion type or site nor aphasic syndrome or pre-onset usage of L1 and L2 (to name just a few) were responsible. In general, it can therefore be concluded that a combination of multiple factors seems to be responsible for the individual recovery process. Another finding is what is often referred to as *pathological code switching* (or language interference); that is, sometimes aphasic patients seem to suffer from an impaired attention control of switching between both languages. For instance, lexical units of L2 cannot be inhibited and are produced although the listener does not understand this language (e.g., Mariën, Abutalebi, Engelborghs, & De Deyn, 2005). These code switching disorders have been associated with deep left frontal lesions. Here, we would need to consider also that the chance of linguistic interference between two languages is higher the more similar the languages are. For example, one might expect more instances of interference if the relevant language pair is Spanish and Italian rather than Spanish and Urdu. As the present chapter focuses on (morpho) syntactic processing in bilingual speakers, let us look at two additional examples. In Fabbro's (2001) study, agrammatic Italian/Friulian aphasic patients showed in general a parallel recovery process for both languages, but behaved different with respect to omitting pronouns. This is not surprising if we take into account the typology of both languages. Italian is a pro-drop language (much like Spanish or Japanese), but not Friulian or English. For instance, in Italian you will say *bevo vino* (drink wine), whereas the verb inflection "-o" indicates first person singular, a grammatical role also expressed by the pronoun "io" (as in "I drink"). Thus, if a pronoun will be dropped in Friulian, it is obviously a grammatical error, but this error cannot be detected in Italian as the pronoun omission is grammatically permitted and actually preferred. Similarly, English is a weakly inflected language; it has no grammatical gender (though not in Old English). Most Slavic and other languages have more than two grammatical genders. Romance languages typically use two different grammatical genders, feminine vs. masculine, but there are often exceptions, and often linguists are required to account for specific morpho-syntactic patterns of a particular language. For instance, Spanish uses in addition to feminine and masculine markers, pronouns that do not have a gendered noun as an antecedent but are neuter and refer to a whole idea, clause, or objects not mentioned in the discourse (e.g., *ello, esto, eso*, and *aquello*). The reader might want to realize that the observational method relies heavily on the behavioral-linguistic analysis, while the associated neural correlates can only be broadly defined. It is desirable that the behavioral approach uses a typologically relevant analysis of the observed L2 patterns. The exact description of the typological findings can be considered as a prerequisite for preparing customized stimulus material in those studies that use sophisticated technology to reveal the neural correlates of L2 processes. Although the behavioral approach primarily serves as a control for the main experiment, it represents an essential and very important method of controlled testing of language processing.

In this vein, an attempt has been made to link the behavior of outstanding personalities with exceptional skills to cortical properties that are different from those of the *average person*. In the domain of language, we refer here to the postmortem brain examination of the German sinologist/linguist Emil Krebs (1867–1930), who, according to family reports, "mastered" more than 68 languages verbally and in writing and had knowledge of about 120 languages. While there are good reasons to doubt that his language skills reached the online fluency level of 68 different native speakers, we can be certain that he was an extreme polyglot. In other words, his meta-linguistic knowledge and his ability of phonological modulation were exceptionally good. Cytoarchitectonic or anatomical differences between Krebs' brain and 11 control brains were analyzed by means of cortical measurements (morphometry) and multivariate statistical analysis (Amunts, Schleicher, & Zilles, 2004). The authors concluded that Krebs' brain shows a local microstructural specialization (as compared to the control brains) for Broca's area (speech-related brain area): a unique combination of interhemispheric symmetry of BA 44 and asymmetry of BA 45 with respect to the right hemisphere (areas BA 44 and 45 are anatomical correlates of Broca's speech region). These findings are difficult to interpret, as a unique *exceptional brain* cannot be compared. However, let us assume for a moment that indeed a correlation between linguistic behavior and cortical structure exists in the case of Emil Krebs. Still, we cannot conclude that the cortical differences are actually related to linguistic computations *per se* or to cognitive operations supporting or providing the base for these computations. For instance, it is unclear whether cortical differences are related to high demands on working memory functions, to operations associated with controlled switching among different languages (as required for translations), or to the amount of lexical information processed, or whether the results are coincidental and unrelated to his linguistic behavior. However, in assuming that any highly repeated cognitive activity results in cytomorphological changes, much like people train their leg muscles to run faster, a correlation might be plausible in the case of Emil Krebs, but conclusions about neural correlates of a specific linguistic behavior remain highly speculative. Today, more direct methods are available to reveal the neural substrates of L2 processing. Let us turn therefore to electrophysiological and neuroimaging methods and studies that provide new insights regarding the neural correlates of bilingual processes.

Electrophysiological and Magnetophysiological Measures

Event-Related Potentials

The most popular noninvasive method to measure electrophysiological activity of the brain is called event-related potentials (ERPs). It can be considered as functional electroencephalography (EEG), as electric cortical activity is measured in response to a cognitive-behavioral task. ERPs reflect thousands of parallel cortical processes, and correlation of the electric signal to a specific stimulus requires many trials, so that random noise can be averaged out. ERPs provide an online measurement of the brain's activity and may reveal responses that cannot be exclusively detected by behavioral means. The most-known ERP components are the *early left anterior negativity* (ELAN), the N400, and the P600. ELAN is a negative μV response that peaks at approximately less than 200 ms after presentation of a phrase structure violation (e.g., *Sam played on the *wrote*), and the N400 is a negative response to a semantic violation at approximately 400 ms after the onset of the stimulus presentation (e.g., **Sam ate the shoes*); the P600 is a positive response (also called *syntactic* *positive shift*, SPS) that peaks at approximately 600 ms after stimulus presentation and can be measured in sentences requiring revision of the initial parse (e.g., garden-path sentences), at gap-filling dependencies, and when morpho-syntactic violations (e.g., number, case, gender) are encountered.

Magnetoencephalography

Magnetoencephalography (MEG), first reported by Cohen (1968), has a temporal resolution and generates evoked responses much like EEG/ERP. The magnetic components are labeled according to temporal latency. For example, the M100 is elicited at approximately 100 ms post-stimulus presentation of a particular stimulus, usually tones, phonological information, or words. The M400, which corresponds to the N400 found with ERPs, is generated in the context of semantic processing. However, magnetic fields are less distorted than EEG and therefore have a better spatial resolution. While EEG is sensitive to extracellular volume currents elicited by post-synaptic potentials, MEG is sensitive to intracellular currents of these synaptic potentials. EEG can detect activity in the sulci and at the top of the cortical gyri, but MEG detects activity mostly in the sulci. (A sulcus is a depression or groove between two cortical convolutions.) In contrast to EEG, MEG activity can be localized with more accuracy. MEG is often combined with functional magnetic resonance imaging (fMRI) to generate functional cortical maps.

Selected Studies

Depending on a series of L2 factors (e.g., language proficiency, age of L2 acquisition, and structural similarities between L1 and L2), various findings have been reported. To begin with, the data reported do not support the account of a critical period for language acquisition. However, before addressing this very important issue, let us look closer at some interesting electrophysiological findings with respect to L2 acquisition.

In Weber-Fox and Neville's (1996) seminal study, a difference was found in late and early L2 learners. While all groups (i.e., native speakers, late and early L2 speakers) showed an N400 effect, they reported that late L2 English speakers (less than 11 years of age) showed a delayed N400 of 20 ms as compared to the other groups. In Hahne and Friederici (2001) study, late L2 (Japanese-German bilinguals) and monolinguals showed a similar N400 effect for semantically incorrect sentences. However, the N400 effect lasted approximately 400 ms longer in bilinguals than in monolinguals. The authors considered the possibility that this delay might have reflected the attempt of late L2 speakers to integrate the critical word in the sentence context, as reduced lexical knowledge may have prevented a fast decision comparable to native speakers (see also Mueller, 2005; Sanders & Neville, 2003). Thus, the N400 effects found are quite similar among L1 and L2 speakers. The differences are mostly related to changes of latency and amplitude in late L2 speakers.

In the case of morphologically complex words, Russian late L2 speakers of German showed an ERP waveform with two phases much like L1 speakers (Hahne, Mueller, & Clahsen, 2006). While incorrect participles elicited an early anterior negativity and a P600, incorrect plurals solely generated a P600. This finding is in line with production proficiency levels, as L2 speakers perform worse on plurals than on participles, probably due to differences in rule complexity. Thus, these data indicate that even late L2 speakers can reach native-like, automatic computations of morphologically complex words. A study by Rossi, Gugler, Hahne, and Friederici (2006) shows that age of acquisition is not necessarily the leading factor, but proficiency is more important. They found for late high-proficient L2 speakers of German or Italian and respective monolinguals comparable ERPs (ELAN, negativity, P600) for active voice sentences and agreement violations (ELAN, P600). In contrast, low-proficient L2 speakers elicited similar patterns for phrase structure violations, but only a P600 (not an ELAN) for agreement violations. Moreover, the low-proficient L2 speakers showed a delayed P600 with reduced amplitude.

Fine-grained differences in syntactic L1 and L2 processing were reported in a series of MEG studies with Japanese (relatively) late L2 English learners (average age across studies: 25–28 years; Kubota, Ferrari, & Roberts, 2003, 2004; Kubota, Inouchi, Ferrari, & Roberts, 2005). The first study tested case violations checked phrase-internally (9a) or checked phrase-externally (9b).

9. (a) **I believe he to be a spy*.
   (b) **I believe him is a spy*.

Only the M150 (ELAN-like response at approximately 150 ms post-stimulus) was reported for the phrase-internal checking violation in L1 speakers. L2 speakers seemed unable to process this structure in an automatic fashion. The second study tested violations of noun phrase raising (10a) and Case filter (10b; i.e., every overt noun phrase must have a Case).

10. (a) **The man was believed (t) was killed*.
    (b) **It was believed the man to have been killed*.

Here, the case filter violation did not elicit an M150 response, but the noun phrase raising violation did. Both L1 and L2 speakers showed this response pattern, indicating high-order syntactic sensitivity in L2 speakers. The third study examined infinitive (11a) and gerund complement violations (11b).

11. (a) **He postponed to use it*.
    (b) **He happened using it*.

Again, the gerund complement violation resulted in an M150 response for L1 and L2 speakers but the infinitive complement violations did not. Overall, these results show that only certain syntactic structures can be processed in an automatic (online) fashion much like native speakers. Numerous MEG bilingual studies are published referring to different linguistic levels (for a review, see Schmidt & Roberts, 2009).

Hemodynamic Measures

Magnetic Resonance Imaging

The most popular neuroimaging technique among researchers is magnetic resonance imaging (MRI). The invention of MRI did not arrive in one step and is the result of a series of accomplishments in physics. A description of the methods and mechanisms behind MRI is beyond the scope of this chapter, and the reader will be referred to adequate tutorials (e.g., Pooley, 2005). However, let us briefly summarize some important facts about these important but still developing noninvasive neuroimaging techniques. The most common kind of MRI is known as blood oxygenation level-dependent (BOLD) imaging and is credited to Ogawa, Lee, Nayak, and Glynn (1990). Neurons receive energy in the form of oxygen by means of hemoglobin in capillary red blood cells. An increase of neuronal activity results in an increased demand for oxygen, which in turn generates an increase in blood flow. Hemoglobin is unaffected by the magnetic field (diamagnetic) when oxygenated but strongly affected (paramagnetic) when deoxygenated. The magnetic field is generated by an MRI scanner, which houses a strong electromagnet. For research purposes, the strength of the magnetic field is typically 3 T (1 T = 10,000 G) and is 50,000 times greater than the Earth's field. It is predicted that the spatial resolution at the cell level requires high-field magnets (far greater than 10 T; Wada et al., 2010). This difference in magnetic properties causes small differences in the MR signal of blood depending on the degree of oxygenation. The level of neural activity varies with the level of blood oxygenation. This hemodynamic response (HDR) is not linear. The onset of the stimulus-induced HDR is usually delayed by approximately 2 s because of the time it takes the blood to travel from arteries to capillaries and draining veins. There is typically a short period of decrease in blood oxygenation immediately after neural activity increases. Then, the blood flow increases not only to meet the oxygen demand, but to overcompensate for the increased demand. The blood flow peaks at around 6–12 s before returning to baseline. In contrast to a relatively good spatial resolution between less than 1 mm, the temporal resolution has its limits. However, let us look in the following at some studies using fMRI to investigate L2 processing.

Selected Studies

Some fMRI studies were designed to find an answer for the basic question of whether L1 and L2 would activate the same or different cortical regions according to age of acquisition. Kim et al. (1997) studied *early* (mean age 11.2 years) and *late* (mean age 19.2 years) bilingual speakers. The age of L2 acquisition was defined with respect to age when conversational fluency was reached in the L2. The (healthy) participants were asked to silently generate sentences according to imagined events. The authors reported spatial differences in Broca's area in late bilinguals for processing L1 and L2, but early bilinguals activated for both languages two non-overlapping subregions of Broca's area. No differences were reported for Wernicke's region. Dehaene et al. (1997) reported sentence processing differences between L1 and L2 English-French speakers, where the L2 speakers recruited more right hemispheric activations. Only early bilinguals who acquired both languages at birth showed an overlap of activation for L1 and L2 (see also Perani et al., 1996; Saur et al., 2008). Two other studies revealed no difference in cued word generation and sentence judgment tasks by early (younger than 6 years of age) and late (older than 12 years of age) Mandarin-English bilinguals (Chee et al., 1999; Chee, Tan, & Thiel, 1999). However, the variable *age of acquisition* might not actually be the critical variable, at least at the level of sentence comprehension (see for example, Heredia & Cieślicka, 2014). Instead, the variable *fluency* (often to some extent inter-related to age of acquisition) seems to be important as highly fluent bilinguals activate similar left temporal lobe areas for L1 and L2 (Perani et al., 1998), but not less-fluent bilinguals (Perani et al., 1996). Very interesting findings stem also from a positron emission tomography (PET) study. PET scans were popular before MRI technology became fully established. PET is an imaging test that uses a small amount of radioactive substance (called a tracer). This neuroimaging technique has been superseded by MRI technology, although it is sometimes used in identifying brain receptors (or transporters) associated with particular neurotransmitters (although not applied for this reason in Price, Green, & von Studnitz, 1999 study). In this study, neural activity was measured during reading in German and English and translating words from German into English or vice versa (Price et al., 1999). The L1 of the six participants was German, and all acquired English as L2 at approximately 9 years of age. Compared to reading, the translation task activated cortical regions outside of the typical language areas, which involved the anterior cingulate and bilateral subcortical structures (putamen and head of the caudate nucleus). Translation involved less automatized circuitries but a higher effort of coordination. In addition, during translation, control functions showed higher activation of the supplementary motor cortex, cerebellum and the left anterior insula. During language switching (not translation), an increase of activation was found in Broca's area and in the bilateral supramarginal gyri. Thus, many neural activities related to processes between L1 and L2 occur outside of the typical language circuitries. Another bilingual fMRI study examined how L1 English speakers' process visually presented simple declarative sentences and signed sentences in comparison to signers of American Sign Language (ASL). The classical Broca-Wernicke circuit was activated in both languages, but in contrast to native English speakers, reliable activation was found in native signers (deaf or hearing) in posterior right hemisphere areas. This study confirms the particular role of the right hemisphere in visuospatial processing (Bavelier et al., 1998).

Let us look now more closely at syntactic processing in bilinguals, a cognitive domain typically supported by Broca's region in L1 speakers. In Suh et al.'s (2007) fMRI study, it was shown that for both languages (Korean-English), among other areas, the left inferior frontal gyrus (IFG) and the (bilateral) inferior parietal gyri were activated when late bilinguals were asked to read center-embedded (12a) and conjoined sentences (12b).

12. (a) *The director that the maid introduced ignored the farmer.*
    (b) *The maid introduced the director and ignored the farmer.*

However, in the left IFG (but not in any other areas) activation was higher for embedded vs. conjoined sentences in L1 but not for L2. The authors concluded that the same cortical areas are recruited for syntax for both languages, but the underlying neural mechanisms were different. These data are in direct contrast to the findings of those of Hasegawa, Carpenter, and Just (2002), who reported that neural activation increased in L2 as compared to L1 due to sentence complexity (negated vs. affirmative sentences). Suh and colleagues assumed that in L1, less complex sentences might be processed in an automatic fashion while more complex sentences are not automatized and thus involve a higher cognitive demand. In L2, however, this difference cannot be detected, as processing of different sentence structures would not have been automatized. This is a plausible interpretation. In the present case, syntactic complexity correlates with higher cognitive demands and multiple linguistic and/or pragmatic aspects can be the source of increased neural activation.

A recent study that used magnetic resonance diffusion tensor imaging (MR-DTI; see Basser, Pajevic, Pierpaoli, Duda, & Aldroubi, 2000) revealed white matter difference in L1 and L2 speakers (Mohades et al., 2012). White matter connections can be better analyzed with DTI and fiber tractography than with standard MRI. The DT-MRI method measures in all three dimensions *in vivo* and noninvasively the random motion (diffusion) of hydrogen atoms within water molecules. Water resides in tissues, which consist of a large number of fibers such as brain white matter. DT-MRI renders in 3D complex information about how water diffuses in tissues. The participants of this study were native speakers and *simultaneous* and *sequential* bilinguals (mean age: 9.5 years). Sequential bilingualism refers to acquiring the L2 after 3 years of age, and in simultaneous bilingualism both languages are acquired from birth onward (L1 was either French or Dutch, and L2 was a Romance or a Germanic language). One of the findings is that simultaneous bilinguals had higher mean fractional anisotropy (FA) values for the left inferior occipito-frontal fasciculus tracts (which connect anterior regions of the frontal lobe with posterior regions in the temporal occipital lobe) than monolinguals. However, the comparisons for the fiber projection anterior corpus callosum to the orbital lobe showed a lower mean FA value in simultaneous bilinguals as compared to monolinguals. In both cases, the sequential bilinguals had intermediate values as compared to the other two groups. FA is a measure for fiber density, axonal diameter, and myelination in white matter. It is therefore plausible to assume that the acquisition of two native languages at birth is beneficial for stronger and faster anterior-posterior fiber connections supporting language processing. However, as the myelination process of the fiber tracts is not complete in childhood, it might be that this outcome reflects only a particular time window of white matter development. We cannot exclude the possibility that no significant FA differences will be measured for the anterior-posterior connection in adult monolinguals and bilinguals. If the fiber system is fully developed, a ceiling effect might be reached. Therefore, we do not exclude the assumption of a lifetime learning process that can modify or change already-established properties of fiber connections. However, a post-puberty modification involved presumably different neural modifications from those in infantile brain development. The second interesting finding reported by Mohades and colleagues, namely, lower mean FA value for simultaneous (early) bilinguals regarding the corpus callosum to orbital lobe connection, is in line with the results that early bilinguals tend to be less left-sided lateralized for language than monolinguals or late bilinguals (Hull & Vaid, 2006; Josse, Seghier, Kherif, & Price, 2008). Additionally, an increase in the size of the corpus callosum seems to correlate with a higher degree of left lateralization for language. These and other findings directly verify the assumption that the specific language acquisition process shapes the fiber system that is responsible for connecting different language-relevant regions. In other words, cortical regions become language sensitive in a specific manner, as the fiber system connects these regions according to linguistic input received.

Methodological Considerations

Some neurolinguistic findings show that late L2 speakers activate different cortical areas for L1 and L2. In contrast, there is clearly a tendency that early L2 speakers recruit the same cortical areas for L1 and L2. This general outcome is difficult to interpret: Do early L2 speakers rely on a single language system opposite to late L2 speakers, who have different computational systems for L1 and L2? How many different language systems are then cortically represented in a different way in non-early-polyglot speakers? We do not have access to sufficient specific data to draw more general conclusions. L2 speakers vary in proficiency and fluency, use languages with different degrees of similarity, and have experiences with different communication styles and domains, for example. Thus, it is not surprising to assume that every individual brain organizes language(s) in a different way. Often the differences found for early and late L2 speakers have been attributed to a critical period of language acquisition.

The concept of a *critical period* (in contrast to a *sensitive period*) refers to a phase in the life span of an organism in which it develops or acquires a particular skill. If the organism is not exposed to the relevant stimuli during this critical phase, it is difficult or even impossible to use these skills later in life. For example, the common chaffinch must be exposed to the songs of an adult chaffinch before adulthood, before it sexually matures, to be able to acquire this intricate song. A critical period for language acquisition has been claimed by Lenneberg (1967; see also Pinker, 1994). Lenneberg argued that the critical language period is between 5 years of age and puberty, and referred to the observation that feral (e.g., *Genie*; see Rymer, 1993) or deaf children have difficulties acquiring spoken language after puberty. Moreover, Lenneberg assumed that children with neurologically caused language disorders recover significantly better and faster than adults with comparable impairments. This argument is, however, not well supported. First, feral or deprived children vegetate in an inhuman environment, which has severe consequences for physiological, psychological, cognitive, and social developments in general. It seems quite naive to assume that the dramatic impact of deprivation can be reversed or should not influence learning (including language) after the child has been rescued. Second, one cannot draw direct comparisons between a neuropsychological recovery process and a typical acquisition process in children. One might argue that there is a sensitive period for recovery from neurological language disorders, but at the same time it cannot be concluded that the same process applies to typically developing children. Neural structures (re)organize throughout the life cycle, and it is not surprising that, during the formation of neurons and connectivity in infancy and early childhood, irreversibility of disorders is most promising and gradually decreases the more neural circuits become wired. However, this genetically determined neural developmental process does not represent a period of language recovery, as neural recovery occurs throughout the life cycle. New neurons are continuously developing throughout adulthood and are integrated in existing neural formations. If the assumption of a critical recovery period were true, aphasic patients would not be able to recover at all or with minimal success. However, the clinical reality shows the opposite; though recovery takes more time than at a young age, neural plasticity provides good recovery at any stage of the life cycle if the cortical damage does not exceed a certain degree of severity (Heiss, Thiel, Kessler, & Herholz, 2003).

Certainly, our daily observations tell us that young children acquire cognitive skills in a playful manner as compared to adults, whose learning process is apparently more effortful. However, does this imply that adults cannot reach the fluency or proficiency of a second language that young children do? The answer must be strictly denied. Everyone at any age can reach L1 fluency level in L2. Our brain is not an organ whose functionality declines with the onset of adulthood. Brain plasticity and adult neurogenesis is a dynamic process and facilitates the acquisition of L2 proficiency in adulthood. Many variables would need to be considered to explain why an individual acquires L2 knowledge in a specific manner. In general, it needs to be considered that it is difficult to capture neural activities requiring similar processing resources in L1 and L2. As pointed out before, morpho-syntactic and phonological rules are different among languages, and *comparable* structures in L1 and L2 may recruit different cognitive demands because of different degrees of automatized processes. These differences may also be reflected in recruiting non-overlapping, different neural correlates, and thus it cannot be strictly concluded that specific linguistic structures are processed by L1 and L2 in different cortical regions. For example, studies involving late bilingual twins (13 years of age) suggest that the same neural regions are involved during grammatical processing in the L1 as well as in the L2. The twins' native language is Japanese, but they were trained during a period of 2 months on English verb conjugations. Pre- and post-training fMRI studies revealed increased activity in the left dorsal IFG, which correlated with their behavioral performance. Despite significant proficiency differences in L1 and L2 with respect to the verb generation of past tense, the same cortical region was activated (Sakai, Miura, Narafu, & Muraishi, 2004). Similarly, when *grammatical rules* were examined in a non-natural, foreign language that included rules that were inconsistent with those of natural languages, only the language-consistent rules activated Broca's area (Musso et al., 2003; Tettamanti et al., 2002). This is confirmed by a recent fMRI study showing neural convergence in highly proficient bilinguals with respect to sentence comprehension and verb/noun production tasks (Consonni et al., 2013). Taken together, anatomical studies support the following conclusion: If the L2 proficiency level matches native-level proficiency, common neural activities can be found in the left frontotemporal language circuit; if the L2 proficiency level is clearly lower compared to L1, additional cortical resources are recruited in the prefrontal cortex.

Summary and Conclusions

In this chapter, we have presented a wide range of different methods used to examine the cognitive and neural foundations of L2 processing. In the field of experimental psycholinguistics, special methods have been developed to tap online, moment-by-moment into the (re)activation patterns of lexical information during sentence comprehension. These online methods are important for measuring automatic linguistic computations. While the application of a single method depends on the specific issue to be examined, it is generally recommended that more than one method is used in a single study. One of the reasons is that method-specific factors can be better controlled, which in turn allows interpretation of the data from different empirical and theoretical perspectives. Moreover, researchers should be encouraged not only to rely on specific psycholinguistic methods, but also to consider customizing established methods for special needs.

In the field of cognitive neuroscience, various complex methods and techniques are applied to reveal the neural correlates of cognitive processing. Thus, the approach is less theory driven, but attempts to shed light on those neurobiological circuitries and cortical structures that serve as a scaffold in language processing. The introduction of different electro- and magneto-physiological and neuroimaging methods, respectively, demonstrates certain inherent technical limitations. However, the development of the neurobiological research paradigm is a highly dynamic, progressing field. The focus is on how to improve the temporal and/or spatial resolution to track language processing in a time span of milliseconds as well as at a neuromolecular level. Thus, a neurobiological approach is less concerned with finding evidence for a particular linguistic model, but tries to reveal the underlying cortical structures supporting language processing. However, future studies may find a synthesis between these different paradigms to link fine-grained L1 and L2 computations, respectively, to specific neural circuits and ultimately to biochemical conditions.

List of Keywords

Active filler strategy, Adjuncts, Age of acquisition, Aphasia, Bilingual Aphasia Test (BAT), Blood oxygenation level dependent (BOLD), Broca's area, Cross-modal lexical priming (CMLP), Direct Association Theory, Event-related potentials (ERPs), Fiber tractography, Filler-gap dependency, First pass duration, Fractional anisotropy (FA), Generative grammar, Hemodynamic response (HDR), Inferior frontal gyrus (IFG), M100, M150, Magnetic resonance diffusion tensor imaging (MR-DTI), Magnetic resonance imaging (MRI), N400, Online processing, P600, Pathological code switching, Positron emission tomography (PET), Priming effect, Probe recognition, Reading latencies, Re-reading time, Regression path duration, Relative clause, Sequential bilinguals, Self-paced listening task, Shallow Structure Hypothesis (SSH), Stop-make-sense (SMS) task, Syntactic positive shift (SPS), Syntactic theory, Trace Reactivation Hypothesis.

Review Questions

1. Basic word order varies according to languages. In some languages, a filler and its corresponding gap site always appear before their subcategorizing verb but in other languages, a filler will be encountered first, its subcategorizing verb appears next, and, finally, a gap site. Are the processes involved in the establishment of a filler-gap dependency different if the basic word orders are different?
2. It is common knowledge that if one starts learning a new language after puberty, it is difficult to achieve native-like proficiency in this language. What are the possible causes?
3. Typically sentences are embedded in a text. The restrictive use of the relative clause (e.g., when a comma does not occur before "to which") in sentence (13) below implies that *the nice monkey* explained the game's difficult rules to another *squirrel*. For instance, in the *text below there are two squirrels. If a listener hears this sentence, he or she knows that the nice monkey* had explained the game's difficult rules to one of the squirrels by the time she/he hears the sentence. Does the listener still need to associate *the squirrel* and *explained* and reactivate the antecedent at the gap site?

   Fred and a monkey were playing a new game with their friends. In the game, they were chasing each other. Two squirrels came to join the game, but they didn't know the game's rules. The rules were difficult and took time to explain. Unfortunately, a bell rang, telling them to go home. Later, a nice monkey explained the rules to one of the squirrels in the class last Wednesday, and Fred explained the rules to the other squirrel during lunchtime last Thursday. On the weekend, everybody got together and started playing the game.

13. *Fred chased the squirrel to which the nice monkey explained the game's*$_{\#1}$ *difficult rules_* $_{\#2}$ *in the class last Wednesday.*

4. Williams (2006) suggested that resources such as memory capacity affect the experimental results. It has been proposed that working memory is used to temporarily retain information and then use it during sentence processing. Nakano, Felser, and Clahsen (2002) found that the capacity of individuals' working memory varied, and it influenced the magnitude of priming in their cross-modal priming experiment. That is, the participants with larger working memory capacities showed a priming effect at the gap site, but the participants with smaller working memory capacities showed no priming effect at the gap site. Do these differing results indicate that the groups' mechanisms for sentence processing are different?
5. Which factors may contribute to the findings that a bilingual speaker processes L1 and L2 differently or similar?

Suggested Student Research Projects

1. Describe an experimental design to investigate the establishment of a filler-gap dependency by using one of the methods in languages other than English, including auditory languages, such as Spanish and Chinese, and if possible, in visual languages, such as American Sign Language, and Japanese Sign Language.
2. Extend bilingual research to figurative language.
3. Determine whether a regional dialect behaves like an L2.
4. Research whether a form of bilingualism can be found in non-human species (e.g., songs of birds or whales).
5. Speaking more than one language is beneficial. Describe the benefits.
6. It is well known that children learn a second language more easily than adults. Please discuss reasons for this phenomenon.

Related Internet Sites

Lexical Acccess: https://en.wikipedia.org/wiki/David_Swinney
Multilingualism: https://en.wikipedia.org/wiki/Multilingualism
Wh-movement: https://en.wikipedia.org/wiki/Wh-movement
Word-Sense Disambiguation: https://en.wikipedia.org/wiki/Word-sense_disambiguation

Suggested Further Reading

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Review Neuroscience, 15*(5), 336–345.

Hillert, D. (2014). *The Nature of Language. Evolution, Paradigms, Circuits*. New York: Springer.

Hillert, D. (Ed). (1998). *Sentence processing: A crosslinguistic perspective* (Syntax and Semantics v. 31). San Diego, CA: Academic Press.

Acknowledgement With special permission from Springer, selective paragraphs of this chapter are published in Hillert (2014).

References

Amunts, K., Schleicher, A., & Zilles, K. (2004). Outstanding language competence and cytoarchitecture in Broca's speech region. *Brain and Language, 89*(2), 346–353.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine, 44*(4), 625–632.

Bavelier, D., Corina, D., Jezzard, P., Clark, V., Karni, A., Lalwani, A., et al. (1998). Hemispheric specialization for English and ASL: Left invariance-right variability. *NeuroReport, 9*(7), 1537–1542.

Bever, T., & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry, 19*, 35–43.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language, 29*, 413–432.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from *wh*-questions. *Journal of Memory and Language, 34*, 774–806.

Carlson, G., & Tanenhaus, M. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Syntax and semantics* (Vol. 21, pp. 263–288). London: Academic Press.

Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *The Journal of Neuroscience, 19*(8), 3050–3056.

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., et al. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron, 23*(1), 127–137.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (1999). *Derivation by phrase*. In MIT Occasional Papers in Linguistics 18. Cambridge, MA: MITWPL.

Clahsen, H., & Featherston, S. (1999). Antecedent priming at trace positions: Evidence from German scrambling. *Journal of Psycholinguistic Research, 28*, 415–437.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences, 10*, 564–570.

Cohen, D. (1968). Magnetoencephalography: Evidence of magnetic fields produced by alpha-rhythm currents. *Science, 161*(3843), 784–786.

Consonni, M., Cafiero, R., Marin, D., Tettamanti, M., Iadanza, A., Fabbro, F., et al. (2013). Neural convergence for language comprehension and grammatical class production in highly proficient bilinguals is independent of age of acquisition. *Cortex, 49*(5), 1252–1258.

Dallas, A., & Kaan, E. (2008). Second language processing of filler-gap dependencies by late learners. *Language and Linguistics Compass, 2*, 372–388.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., et al. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport, 8*(17), 3809–3815.

Fabbro, F. (2001). The bilingual brain: Bilingual aphasia. *Brain and Language, 79*(2), 201–210.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research, 23*, 9–36.

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in non-native sentence processing. *Studies in Second Language Acquisition, 34*, 67–98.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's processing of ambiguous sentences: A study of relative clause attachment. *Language Acquisition, 11*, 127–163.
Fodor, J. D. (1978). Parsing strategies and constraints on transformation. *Linguistic Inquiry, 9*, 427–473.
Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*(2), 123–141.
Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience, 18*(1), 121–134.
Harberlandt, K. F., & Bingham, G. (1978). Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behavior, 17*, 419–425.
Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage, 15*(3), 647–660.
Heiss, W. D., Thiel, A., Kessler, J., & Herholz, K. (2003). Disturbance and recovery of language function: Correlates in PET activation studies. *NeuroImage, 20*, 42–49.
Heredia, R. R., & Cieślicka, A. B. (2014). Bilingual memory storage: Compound-coordinate and derivatives. In R. R. Heredia & A. Jeanette (Eds.), *Foundations of bilingual memory* (pp. 11–39). New York: Springer.
Hillert, D. (2014). *The nature of language. Evolution, paradigms and circuits*. New York: Springer.
Hull, R., & Vaid, J. (2006). Laterality and language experience. *Laterality, 11*(5), 436–464.
Josse, G., Seghier, M. L., Kherif, F., & Price, C. J. (2008). Explaining function with anatomy: Language lateralization and corpus callosum size. *The Journal of Neuroscience, 28*(52), 14132–14139.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.
Kubota, M., Ferrari, P., & Roberts, T. P. L. (2003). Magnetoencephalography detection of early syntactic processing in humans: Comparison between L1 speakers and L2 learners of English. *Neuroscience Letters, 353*(2), 107–110.
Kubota, M., Ferrari, P., & Roberts, T. P. L. (2004). Human neuronal encoding of English syntactic violations as revealed by both L1 and L2 speakers. *Neuroscience Letters, 368*(2), 235–240.
Kubota, M., Inouchi, M., Ferrari, P., & Roberts, T. P. L. (2005). Human magnetoencephalographic evidence of early syntactic responses to c-selection violations of English infinitives and gerunds by L1 and L2 speakers. *Neuroscience Letters, 384*(3), 300–304.
Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
Lieberman, M., Aoshima, S., & Phillips, C. (2006). Native-like biases in generation of wh-questions by non-native speakers of Japanese. *Studies in Second Language Acquisition, 28*, 423–448.
Love, T. E., & Swinney, D. A. (1998). The influence of canonical word order on structural processing. In D. G. Hillert (Ed.), *Syntax and semantics* (pp. 153–166). London: Academic Press.
Mariën, P., Abutalebi, J., Engelborghs, S., & De Deyn, P. P. (2005). Pathophysiology of language switching and mixing in an early bilingual child with subcortical aphasia. *Neurocase, 11*(6), 385–398.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
Miyamoto, E. T., & Takahashi, S. (2002). Source of difficulty in the processing scrambling in Japanese. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 167–188). Stanford, CA: CSLI.
Mohades, S. G., Struys, E., Van Schuerbeek, P., Mondt, K., Van De Craen, P., & Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Research, 1435*, 72–80.
Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*(2), 152–174.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buechel, C., et al. (2003). Broca's area and the language instinct. *Nature Neuroscience, 6*(7), 774–781.
Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace position in Japanese long-distance scrambling. *Journal of Psycholinguistic Research, 31*, 531–571.
Nicol, J. L. (1993). Reconsidering reactivation. In G. T. M. Altmann & R. C. Shilcock (Eds.), *Cognitive models of speech processing: The second sperlonga meeting* (pp. 321–347). Hillsdale, NJ: Lawrence Erlbaum.
Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5–19.
Ogawa, S., Lee, T.-M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine, 14*(1), 68–78.
Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition, 33*, 563–588.
Paradis, M., & Libben, G. (1987). *The assessment of bilingual aphasia*. Hillsdale, NJ: Lawrence Erlbaum.
Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., et al. (1996). Brain processing of native and foreign languages. *Neuroreport, 7*(15–17), 2439–2444.
Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain, 121*(10), 1841–1852.
Pickering, M. (1993). Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes, 8*, 163–196.
Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6*, 229–259.
Pickering, M., Frisson, S., McElree, B., & Traxler, M. (2004). Eye movement and semantic composition. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: ERP, eye-tracking and beyond* (pp. 33–50). Hove, England: Psychology Press.
Pinker, S. (1994). *The language instinct*. New York: Morrow.
Pooley, R. A. (2005). Fundamental physics of MR imaging. *Radiographics, 25*(4), 1087–1099.
Price, C. J., Green, D. W., & von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain, 122*(12), 2221–2235.
Rayner, K., & Sereno, S. (1994). Eye movements in reading: Psycholinguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–81). San Diego, CA: Academic Press.
Rayner, K., Sereno, S. C., Morris, R. J., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*, S121–S149.
Ross, J. R. (1967). *Constraints on variables in syntax*. PhD dissertation, MIT.
Rymer, A. (1993). *Genie: A scientific tragedy*. New York: Harper Perennial.
Sakai, K. L., Miura, K., Narafu, N., & Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex, 14*, 1233–1239.
Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research, 15*(3), 214–227.
Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., et al. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences, 105*(46), 18035–18040.
Schmidt, G. L., & Roberts, T. P. L. (2009). Second language research using magnetoencephalography: A review. *Second Language Research, 25*(1), 135–166.
Stowe, L. A. (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes, 1*(3), 227–245.

Suh, S., Yoon, H. W., Lee, S., Chung, J.-Y., Cho, Z.-H., & Park, H. (2007). Effects of syntactic complexity in L1 and L2: An fMRI study of Korean-English bilinguals. *Brain Research, 1136*(1), 178–189.

Tanenhaus, M., Boland, J. E., Carlson, G., & Garnsey, S. M. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research, 18*, 37–50.

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., & Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *NeuroImage, 17*, 700–709.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*, 454–475.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic effects in parsing: Thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.

Wada, H., Sekino, M., Ohsaki, H., Hisatsune, T., Ikehira, H., & Kiyoshi, T. (2010). Prospect of high-field MRI. *IEEE Transactions on Applied Superconductivity, 20*(3), 115–122.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*(3), 231–256.

Williams, J. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition, 9*, 71–88.

Williams, J., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509–540.

Chapter 11
The Electrophysiology of the Bilingual Brain

Jon Andoni Duñabeitia, Maria Dimitropoulou, Margaret Gillon Dowens, Nicola Molinaro, and Clara Martin

Abstract In this chapter, we focus on how bilinguals cope with reading individual words and sentences in their different languages, and on how electroencephalographic (EEG) recordings could be used to explore the time course of the cognitive processes underlying bilingual comprehension of visually delivered linguistic stimuli. *Event-related brain potentials* (ERPs) comparing native and nonnative written language processing have been repeatedly used to reveal effects that occur very early in the stream of processing and that are essential for a correct understanding of the cognitive processes leading to efficient word and sentence processing by multilingual readers. We summarize the most relevant studies from this field and offer a list of recommendations for researchers aiming at using EEG recordings as a tool to investigate the complex pattern of feed-forward and feedback interactive activations flowing along the visual recognition system that ultimately lead to efficient bilingual reading.

Introduction

As active members of developed modern societies, we are constantly processing information that is delivered to us through the senses in multiple formats. Thus, an individual standing in any street is bombarded with hundreds of simultaneously delivered stimuli that need to be either discarded or processed. The human brain has, however, developed a series of complex mechanisms to filter and sort out

J.A. Duñabeitia, Ph.D. (✉) • M. Dimitropoulou, Ph.D. • N. Molinaro, Ph.D. •
C. Martin, Ph.D.
Basque Center on Cognition, Brain and Language,
Paseo Mikeletegi 69, 2, San Sebastián–Donostia P 20009, Spain
e-mail: j.dunabeitia@bcbl.eu; m.dimitropoulou@bcbl.eu; n.molinaro@bcbl.eu;
c.martin@bcbl.eu

M.G. Dowens, Ph.D.
School of English, The University of Nottingham,
Room 429, Admin Building, 199 Taikang East Road, Ningbo 315100, China
e-mail: Margaret.Dowens@nottingham.edu.cn

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_11

multiple items of information, selectively deciding which are to be further processed and which ones rejected. In fact, the human cognitive system has evolved in such a way that it prioritizes the processing of linguistic information from the social context over other types of stimuli that lack intentional communicative purposes. Thus, human beings in modern societies can be seen as receivers of constant streams of linguistic information, mainly in a visual or auditory format, that is intentionally or unintentionally filtered out by the cognitive system, processed, and then reacted to in a communicative manner. Given the amount of linguistic information present in the modern social environment, it could be said that it is not so easy to be a citizen of the linguistically demanding world, nowadays.

Given the perceptual requirements of our linguistically taxing modern societies, it is predictable that the number of languages used in each specific social context will exponentially increase the cognitive demands on the human being reacting and interacting in this context. Learning to see and understand words and sentences written in different languages possibly requires from the multilingual individual an extra effort that monolingual individuals do not have to make. Nonetheless, the compensation for this extra effort is considerable, allowing multilingual individuals to react to linguistic information from different sources and informants and in different formats (i.e., in different languages) that are unintelligible to monolinguals, and increasing their communicative skills and access to information. Although it may not be easy to become a multilingual citizen of the linguistically complex multilingual world today, it is certainly worth the effort.

In this chapter, we focus on how bilinguals read words and sentences in their first (L1) and second (L2) languages by focusing on the electrophysiological signatures of bilingual reading. There are, of course, many different definitions and profiles of bilingualism, each referring to different dimensions and factors of language acquisition and usage, such as the age of acquisition (AoA) of the L2 (e.g., early vs. late bilinguals), the sequence of L1 and L2 acquisition (e.g., simultaneous vs. sequential bilinguals), the levels of proficiency in each language (balanced vs. non-balanced bilinguals), or the context in which the L2 has been learned (*natural immersion* vs. *classroom learning*). Across cultures and societies, it is easy to find individuals who have acquired their two languages very early in life during childhood (e.g., bilingual societies like the Basque Country, Catalonia, or Wales). These individuals are typically considered early bilinguals, due to their early L2 AoA (as opposed to late bilinguals who have acquired their L2 later in life, during adolescence or adulthood). Bilingual societies are therefore characterized by the existence of early bilinguals (simultaneous bilinguals, in most cases, given that the two languages are often acquired simultaneously during childhood), that in general terms are balanced in their degree of knowledge and use of the two languages. However, it should be considered that most modern societies are not purely bilingual in essence, and that the presence and knowledge of an L2 largely varies across cultures. This way, the majority of bilinguals in the world correspond to sequential bilinguals (i.e., individuals who have acquired the L2 after having acquired and consolidated the mother tongue), who are in most cases non-balanced in their degree of knowledge and dominance of their languages (i.e., the L1 is the most dominant, used, and known one).

Still, this admittedly simplistic reductionism of the typology of bilinguals does not capture the entire complexity of the linguistic reality, and we anticipate that the idiosyncrasy of each bilingual group needs to be correctly identified and defined in order to understand the scope of the findings here reviewed. In this chapter, we attempt to be explicit with regards to these important factors, indicating the type of L1 and L2 users who have been tested in each study. Nonetheless, the reader should note that here we will use the terms “bilingual” and “multilingual” in a general manner, to refer to different profiles of L2 users with different levels of proficiency, ranging from L2 learners with relatively low proficiency in their newly acquired language to bilinguals with native or native-like skills in both languages.

In this chapter, we focus on how bilingual and multilingual individuals cope with reading in their different languages, referring to the cognitive processes underlying bilingual comprehension of visually delivered linguistic stimuli. These will be examined from an electrophysiological perspective, as there are convincing arguments for adopting this approach. Human cognition can, of course, be studied from multiple perspectives and different technical and methodological approaches, and electrophysiology is not the only way to examine how the mind/brain functions. However, to understand the cognitive processes involved in bilingualism, it is important to understand the nature of the neural activity occurring within the cerebral cortex. That this is the case can be seen by a quick Internet search using keywords related to bilingualism and electrophysiology or electroencephalography. This results in an impressive number of articles reporting highly influential results that are currently guiding scientific inquiry in the field of bilingual reading comprehension.

That there is such a striking amount of research exploring the electrophysiology of the bilingual brain is hardly surprising. Electroencephalographic (EEG) recordings are a relatively inexpensive yet extremely effective and rich source of information about the stream of processes that constructs cognition in the human brain. Together with magnetoencephalographic recordings (MEG) and eye-tracking techniques, EEG is one of the most appropriate techniques for identifying the fast time course of linguistic processes, given its exquisite temporal resolution. The extensive use of EEG recordings to explore and understand the cognitive processes underlying bilingual language comprehension is, therefore, grounded in decades of intensive EEG research on monolingual language processing.

EEG recordings allow researchers to tap into the mechanisms of human cognition by registering the neural activity of thousands of millions of neurons, principally the pyramidal neurons that largely make up the cortex, due to their synchronous activity and similar orientation. By means of electrodes strategically located on the scalp, this electrical activity of populations of neurons is registered and quantified over time. Then, by averaging the EEG associated with the processing of a specific type of stimuli (e.g., items from one experimental condition) recorded over a number of trials, the signal-to-noise ratio is improved, allowing for a fine-grained analysis of the processes that hold constant across trials and eliminating the electrical activity corresponding to unstable or noisy processes that are not relevant from a cognitive point of view (e.g., some markers of biological activity, such as electromyographic responses). In this way, for half a century, EEG researchers have focused on *event-related brain potentials* (ERPs), conceived as the averaged electrophysiological response to a specific type of stimuli.

Using this technique, large amounts of evidence have been gathered about the nature of bilingual language processing. In particular, due to the high temporal resolution of ERPs, they allow for a detailed analysis of the time course of the cognitive processes underlying bilingual reading processes which are difficult to track using purely behavioral techniques that only consider final reaction times, as illustrated in the seminal study by Thierry and Wu (2007). These authors, exploring automatic translation processes in Chinese–English bilinguals, illustrated the importance and utility of electrophysiological recordings for close examination of the reading-related processes that take place in the multilingual brain. The results of this study, described below, have been extremely influential in the field for a number of reasons. Apart from the theoretical importance of the findings, this study was able to perfectly exemplify how ERPs can uncover critical effects that could otherwise be overlooked if only behavioral results are considered. The scientific value and importance of ERPs indeed lies in their ability to reveal effects that may occur very early in the stream of processing and to differentiate these effects from those occurring later (i.e., information that cannot be fully captured by traditional behavioral techniques).

In this chapter, we summarize and synthesize some of the most relevant findings from the bilingual EEG literature, paying special attention to how bilingual individuals read words and sentences in their native and nonnative languages. Given the extensive use of EEG recordings in this research area, the present chapter is not the first synthesis of decades of research on bilingual reading (e.g., Van Hell & Kroll, 2012; see also Van Heuven & Dijkstra, 2010). However, the length, depth, and structure of the current chapter offer additional information specifically about bilingual word and sentence reading. Furthermore, this chapter is primarily aimed at researchers who are tackling their first theoretical and practical experience with this technique.

Reading Individual Words

Although being exposed to individual words without a context is not the most common reading situation for a bilingual reader, research on how readers process words in isolation provides a unique way to examine the mandatory steps of word processing that take place during lexical access. In fact, by examining how bilinguals recognize individual words in their L1 and L2, researchers have the opportunity to further explore (a) how the adult brain acquires and integrates new lexical representations (e.g., those representations belonging to the L2) and (b) how these new representations interact with the existing ones (i.e., L1 vs. L2 interactions) at different levels of word processing.

The use of electrophysiological measures to study the processing of individual words during reading has been extremely useful in isolating the underlying cognitive processes leading to efficient word recognition in monolingual reading, as well as for identifying the time course of these processes (e.g., Grainger, Kiyonaga, & Holcomb, 2006; Petit, Midgley, Holcomb, & Grainger, 2006). By combining ERP recordings with the visual presentation of words, many studies have shown that, at least for readers of alphabetic languages, visual word recognition is first mediated by a series of mental operations devoted to processing the low-level visual features of letters. After this, the visual word recognition system proceeds to the identification of the letters composing the words and to the mapping of the graphemes with their corresponding sounds, leading to an initial activation of the whole word form at the orthographic and phonological levels. Finally, the reader's brain accesses the conceptual level of representation, thus activating the word's semantic representation (see Grainger & Holcomb, 2008).

The use of ERPs in the study of bilingual visual word recognition (L1 and L2 processing), though much more scarce and recent, has so far indicated that bilingual reading is not essentially different from monolingual reading at many levels of processing and that the basic steps sketched above take place in a similar manner and time. However, this increasing number of bilingual ERP studies has provided additional insights into when cross-language interactions take place, when the differences and the similarities between L1 and L2 reading are manifested by bilinguals of different characteristics (e.g., L2 AoA, different L2 proficiency levels, different scripts), and when the specific language (L1 or L2) of a visually presented word is perceived and processed (see Moreno, Rodriquez-Fornells, & Laine, 2008; Van Heuven & Dijkstra, 2010, for recent reviews on ERPs and bilingual language processing).

In the following sections, we review the ERP evidence so far on these aspects of bilingual individual word reading. This review will be organized based on whether this evidence has been acquired in experimental settings involving the presentation of words belonging to one of the two languages of interest (single-language context) or in experimental settings involving the presentation of words from both languages of a bilingual (dual-language context). This distinction is considered to be critical when establishing the patterns of L1 and L2 visual processing as well as their interactions at the word level.

On the one hand, studies with experimental paradigms in which all the critical items belong to one language provide more reliable information on the comparison of L1 and L2 word processing, given that the only way in which some characteristics, properties, or word forms of the nontarget language can modulate target language processing is by implicitly activating representations of the nontarget language. Critically, such single-language experimental contexts are also considered to be ideal for revealing the extent of automatic co-activation of the representations of the two languages of a bilingual, since they do not explicitly provide the bilinguals with cues to intentionally or explicitly activating the task-irrelevant (i.e., nontarget) language. In contrast, studies in which bilinguals are presented with words belonging to their two languages are especially well suited for identifying cross-language interactions at different levels of word processing.

Bilingual Word Processing in Single-Language Contexts

The majority of studies aimed at establishing how bilinguals visually process words have focused on how bilinguals process words whose lexical representations (orthographic or phonological) are partially or totally shared across their two languages, as compared to the processing of words whose lexical representations do not share any segments across the two languages. By investigating the differences in the processing of these word types, researchers have explored whether or not single-word processing is affected by the presence of complete or partial cross-language formal overlap in single- and dual-language contexts (see Guo, Misra, Tam, & Kroll, 2012; Peeters, Dijkstra, & Grainger, 2013, for reviews). Evidence about whether words of both languages automatically activate each other has been crucial in establishing whether lexical activation is language specific or not. Furthermore, by using electrophysiological measures, researchers have been able to identify the word processing stages at which these effects emerge.

Words with Semantic and Formal Overlap Across Languages

Some words in a given language have a translation equivalent in another language with largely or completely shared ortho-phonological, lexical, and semantic representations (i.e., identical cognates as *piano* in English and Spanish, and nonidentical cognates as *guitar* in English and *guitarra* in Spanish). In contrast, other word pairs of translation equivalents do not have a similar-looking ortho-phonological representation, while they still refer to the same concept (i.e., non-cognates; e.g., *house* and its Spanish translation *casa*; e.g., Dijkstra, Grainger, & van Heuven, 1999; Lemhöfer & Dijkstra, 2004; Lemhöfer, Dijkstra, & Michel, 2004). In the strictest definition of cognates, the two versions of a cognate pair share both orthographic and phonological segments, as in the English–Spanish examples above. However, there are also cognate pairs across languages that do not share their scripts, in which case the overlap is limited to the phonological level (i.e., phonological cognates such as the English word *cannon* and its Greek translation *κανόνι*; Voga & Grainger, 2007).

With regards to the processing of cognates in comparison to non-cognates, numerous behavioral studies have established that when reading, bilinguals process cognates faster than non-cognates and that this benefit, known as the *cognate facilitation effect*, is larger for identical than for nonidentical cognates and larger as a matter of increased overlap across the two readings of a cognate pair (e.g., Cristoffanini, Kirsner, & Milech, 1986; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). Moreover, cognate effects are more likely to emerge when processing L2 words rather than L1 words (De Groot, Delmaar, & Lupker, 2000; Van Hell & Dijkstra, 2002). In a recent ERP study, Midgley, Holcomb, and Grainger (2011) examined when in the time course of word processing this benefit appears for cognates (as compared to non-cognates), and whether this pattern is constant for L1 and

L2 words. The authors presented English and French partial cognate and non-cognate words organized in blocks by language to a group of English (L1) learners of French (L2). Participants were asked to silently read words for comprehension and to perform a *go*/*no-go* semantic categorization task. Midgley et al. found that for both L1 and L2 items, the differences in the processing of cognates and non-cognates were located mainly in the N400 time window. The N400 component is a classic ERP component conceived as a negative-going deflection peaking at around 350–400 ms after target word onset that has typically been linked to lexical access and semantic processing. In this study, the N400 peaks elicited upon presentation of L1 non-cognate items were more negative-going as compared to those elicited when participants were processing L1 cognates, and this was also the case for the N400 effects elicited by L2 cognates as compared to non-cognates. This pattern suggests that in both L1 and L2, the form-to-meaning mapping is costlier for non-cognates than for cognates. It also provides evidence for language-independent lexical activation, since both L1 and L2 cognate recognition benefited from the ortho-phonological overlap with their nontarget language counterparts. While the L2 cognate advantage was a clearly expected outcome based on the existing behavioral evidence in L2 word processing, the appearance of a cognate effect in L1 words was partially unexpected. Interestingly, while the L2 N400 cognate effect was robust, starting after the first 400 ms of word processing and lasting until the later 500–800 ms time window, the L1 N400 cognate effect had a shorter duration (from 200 to 500 ms). The difference in the timing of the L1 and L2 N400 cognate effects was interpreted by the authors as reflecting a difference in the cross-language activation patterns underlying the cognate effects in the dominant and the non-dominant language. Midgley et al. proposed that, upon presentation of an L1 cognate, its L2 translation becomes partially activated, which in turn sends activation back to its L1 counterpart via direct links between the two word forms, thus facilitating its processing. Within the framework of the *Revised Hierarchical Model* (RHM; Kroll & Stewart, 1994; Kroll, Van Hell, Tokowicz, & Green, 2010), such L2-to-L1 links at the lexical level between translation equivalents are established early during L2 acquisition, as a result of the common association of the new L2 words to their L1 translations. In line with the RHM's proposal that the links of L1 words to their corresponding meaning are stronger than the links of the L2 words, the later-arising and longer lasting L2 N400 cognate effect would be reflecting the fact that the processing of L2 cognates is benefitted by the activation received from their L1 translation via the activation of the shared meaning taking place at a later moment in time.

More recently, Peeters et al. (2013) investigated the processing of orthographically identical French–English cognates (e.g., *message*) both at the behavioral and at the electrophysiological level. This subtype of cognate has been thus far understudied, since when reading identical cognates it is unclear whether bilinguals identify them as belonging to their dominant or to their non-dominant language. In order to identify the representational status of identical cognates, Peeters et al. took advantage of the fact that the N400 has been found to be sensitive to word frequency manipulations and examined the processing of identical cognates, by performing an orthogonal word frequency manipulation of both readings of cognates.

They included identical cognates with high or low lexical frequencies in both French and English as well as identical cognates whose lexical frequencies differed across their French and English readings, and compared them against English non-cognate words. Late French–English bilinguals were asked to perform English lexical decisions while ERPs were recorded. (Please note that throughout this chapter, for any language combination such as French–English, the language listed first is the L1 [French] and the second one is the L2 [English]). The behavioral responses showed facilitation effects for all identical cognates, which were larger for cognates with low English word frequency. Furthermore, the authors found word frequency effects for words with high English frequency as compared to low English frequency. The electrophysiological data were in the same direction, showing cognate facilitation effects (i.e., more negative-going waves for non-cognates), as well as a widely distributed English word frequency effect (i.e., more negative-going waves for identical cognates with low English frequency as compared to those with high English frequency) and a shorter French word frequency effect in the N400 epoch. Furthermore, cognates produced more positive waveforms in the 600–900 ms time window than control words. Critically, the ERP data allowed the authors to discriminate between word frequency effects in English (L2) and French (L1) in the processing of identical cognates. They found an English (L2) N400 word frequency effect for cognates with low English and high French frequency vs. those with high English and low French frequency, suggesting that, despite the fact that bilinguals had acquired English late in life and were clearly more proficient in French than in English, they were more influenced by the L2 reading of the cognate and its lexical characteristics (e.g., word frequency). The authors interpreted these findings as providing evidence in favor of the existence of common orthographic and semantic representations shared by identical cognates but two distinct phonological and morphemic representations, one for each of the two readings of a cognate.

Words with Formal Overlap and Different Meanings Across Languages

In contrast to cognate words, which have a cross-language counterpart with which they share their meaning as well as part of their lexical representation, there are other words that have extensive cross-language overlap at the lexical level but that correspond to different concepts. Research on the processing of words exclusively overlapping at the ortho-phonological level has been extremely informative with regard to (a) the extent to which pure cross-language orthographic or phonological overlap influences L1 or L2 processing, and (b) whether words of both languages compete with each other during lexical access, thus providing crucial information on whether the bilingual lexicon is functionally unique or language specific. Based on whether these words share part or their entire orthographic representation or their phonological representation with words from the other language of a bilingual, we could consider the following types of lexical entries: *interlingual homophones* (i.e., words with largely or completely overlapping phonological representations but different orthographic representations across two languages (e.g., *cow/kau/* in English and *kou/kAu/*, meaning “cold” in Dutch)), *interlingual homographs* (i.e., words from two languages sharing their orthographic and phonological representations (e.g., *red* in English and *red* meaning “net” in Spanish)), and *interlingual orthographic neighbors* (i.e., words of the same length sharing all but one of their letters with a given word of the other language of a bilingual (e.g., the orthographic neighbors of the word *cat* are the words *bat*, *fat*, *mat*, *cab*, etc.; see Coltheart, Jonasson, Davelaar, & Besner, 1977)).

Most of the studies examining interlingual homophones with behavioral measures have thus far used experimental paradigms involving the presentation of both words of the homophonic pair, or have tested pseudo-homophones (namely, nonwords sharing their sound with existing words of the nontarget language but differing in their spelling; Brysbaert, Van Dyck, & Van de Poel, 1999). These studies have primarily used priming paradigms and they have consistently shown facilitative effects on the processing of target words of either the L1 or the L2 when preceded by homophonic words or nonwords of the nontarget language (masked or unmasked), as compared to when primed by control words of the nontarget language (e.g., Dimitropoulou, Duñabeitia, & Carreiras, 2011a). Masked priming paradigms are characterized by the initial presentation of a pattern mask for around 500 ms, followed by the brief presentation of a prime item (for around 50 ms) which can be either related or unrelated to an explicitly presented target item. Under these presentation conditions, participants are typically unaware of the existence of the primes and consequently, any effect observed on the processing of the targets is considered to be automatic and unconscious in nature. In contrast, unmasked or explicit priming paradigms lack this brief and masked prime presentation procedure, making the relationship between primes and targets explicit to the participants. The consistent pattern found in priming studies with homophones has supported the view that the activation of the phonological code is extremely fast and that the phonological representations of the words interact with each other, irrespective of whether they belong to the same language or not. Furthermore, the fact that these cross-language phonological facilitation effects have not been found to depend on the relative frequency of the test items has led researchers to propose that the effects are more likely located at the sub-lexical phonological level (e.g., Dimitropoulou et al., 2011a; Duyck, 2005; Duyck, Diependaele, Drieghe, & Brysbaert, 2004; Van Wijnendaele & Brysbaert, 2002). In the few studies examining the processing of interlingual homophones in the presence of only one of the words composing the homophonic pair (single-word contexts), the results are less straightforward. In two of the first single-presentation studies on cross-language homophones, results showed a disadvantage in processing for interlingual homophones, as compared to control words or to words having either orthographic or orthographic and semantic overlap with words of the nontarget language (Dijkstra et al., 1999; Doctor & Klein, 1992). However, in later studies this pattern has not been replicated, leading to somewhat conflicting results (see Dijkstra et al., 2004; Haigh & Jared, 2007).

Recently, Carrasco-Ortiz, Midgley, and Frenck-Mestre (2012) used ERPs to establish the pattern of effects caused by interlingual homophones under single-word presentation conditions (either inhibition or facilitation) and to further examine whether cross-language phonological overlap influences the lexical selection process. These authors recorded ERPs while French–English bilinguals performed a go/no-go semantic categorization task on English (L2) words which either had a homophonic pair in French or not. The authors focused on the N400 component, since under single-word presentation conditions this component is sensitive to ease of lexical access (Federmeier & Kutas, 1999; Kutas & Federmeier, 2000), while its amplitude has been found to be proportional to the effort made in integrating the word's phonological information (Holcomb, 1993). A monolingual English group was also tested for control purposes. As expected, only the bilingual group showed a reduced N400 amplitude for interlingual homophones as compared to control English words, thus providing support for the hypothesis that cross-language phonological overlap facilitates L2 processing even in a pure monolingual (L2) context. As such, these findings were interpreted in the same line as the previous behavioral reports of facilitative cross-language phonological effects (e.g., Haigh & Jared, 2007). Given the reduction of the N400 for interlingual homophones, the authors proposed that there were no signs of any inhibition at the lexical level as a result of the cross-language overlap across homophones.

The pattern of behavioral effects found when the critical words are interlingual homographs completely sharing their orthographic and at least part of their phonological representation across languages also represents a source of conflict. So far, inhibitory, facilitative, and null effects have been reported for interlingual homographs as compared to control words (e.g., Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Lemhöfer & Dijkstra, 2004; Von Studnitz & Green, 2002). Dijkstra et al. (1998) found that bilingual lexical decisions to Dutch–English interlingual homographs were costlier as compared to English control words when these words were included in an English-only experimental context, suggesting that even in apparently monolingual contexts, the words from both languages compete for selection. Furthermore, this inhibitory effect was larger when homographs had a high lexical frequency in Dutch (the nontarget language) and a low English frequency, as compared to when the Dutch frequency was relatively low, providing evidence of the co-activation of the two versions of the homographic pair. Kerkhofs, Dijkstra, Chwilla, and De Bruijn (2006) carried out a similar manipulation of the lexical frequency of interlingual homographs in an English (L2) semantic priming lexical decision task, while recording both behavioral and electrophysiological responses with a group of late and highly proficient Dutch–English bilinguals. The authors found inhibitory effects for interlingual homographs in the reaction times, as well as increased N400 amplitudes as compared to control words. They also explored whether the magnitudes of the N400 semantic priming effects were modulated as a function of language (L1 vs. L2) and cross-language similarity. They found that L2 words primed by semantically related L2 primes showed smaller N400s than when primed by unrelated primes. Notably, this semantic priming N400 effect interacted with the frequency of the interlingual homographs, providing strong evidence in favor of a language nonselective activation account, as suggested by the *Bilingual Interactive Activation Model*s (BIA, BIA+; Dijkstra & van Heuven, 2002). It also indicated that these cross-language interactions have a lexico-semantic component and that words of the nontarget language are processed as potential candidates for lexical selection.

The view that nontarget language activation proceeds up to the semantic level has been directly supported by a recent ERP study also testing interlingual homographs. In this study, Hoshino and Thierry (2011) presented English prime-target pairs that were either semantically related or unrelated and asked English–Spanish bilinguals to perform a go/no-go semantic relatedness judgment task upon target presentation only when these targets were displayed in red color. The authors analyzed the ERP responses only to the targets that were free from motor artifacts (i.e., the targets displayed in black). Critically, targets in these trials were English–Spanish interlingual homographs (e.g., *pie*, meaning "foot" in Spanish), and primes were English non-cognates that were either semantically related or unrelated to the English target language meaning of the interlingual homograph (e.g., *apple-pie* vs. *rug-pie*) or semantically related or unrelated to the task-irrelevant Spanish meaning of the interlingual homograph (e.g., *toe-pie* vs. *stove-pie*; see also Martín, Macizo, & Bajo, 2010). The authors found smaller N400 amplitudes on targets preceded by semantically related as compared to unrelated primes, and this semantic relatedness N400 effect was present for both the English (target language) and the Spanish (nontarget language) meaning of the interlingual homograph. Nevertheless, in a later 500–650 ms time window corresponding to the *late positive component* (LPC), there was a more positive-going waveform when targets were preceded by primes related to the English meaning of the interlingual homograph, as compared to when they were preceded by unrelated primes, but this semantic relatedness LCP was absent when the semantic relationship involved the Spanish meaning of the homograph. The LCP has been associated with more explicit processing and with the reevaluation of the stimuli (e.g., Martin, Thierry, & Démonet, 2010; see also Müller, Duñabeitia, & Carreiras, 2010). The overall pattern of results suggested that, even if both meanings of the interlingual homograph were activated, the meaning corresponding to the nontarget language was inhibited after 400 ms, while the one corresponding to the task-relevant language was consciously and intentionally processed up to a later stage.

Finally, effects revealing lexical competition across languages have also been repeatedly found in studies testing cross-language neighbors. In this area of research, behavioral studies have shown that the number of orthographic neighbors a given word has, both within and across languages, modulates its processing. Van Heuven, Dijkstra, and Grainger (1998) found that response times to English words were faster when these words had many English orthographic neighbors. However, responses to English targets with many Dutch orthographic neighbors were slowed down, as compared to English words with a low number of orthographic neighbors in Dutch (see also Grainger & Dijkstra, 1992). This critical set of findings showed that, upon presentation of a given word to a bilingual, activation proceeds to orthographically similar words of the target language as well as of the nontarget language, and words from both languages compete for lexical access and selection. Using electrophysiological recordings, Midgley, Holcomb, van Heuven, and Grainger (2008) examined the pattern of ERP effects generated by a manipulation of the cross-language orthographic neighborhood of French (L1) and English (L2) words. Relatively proficient French–English bilinguals were presented with French or English words (blocked by language) and were asked to perform a go/no-go semantic categorization task. The critical words had either many or few orthographic neighbors in the nontarget language, while the density (i.e., the number of orthographically similar words) of the within-language orthographic neighborhood was controlled for. Their results showed increased N400 amplitudes for both English (L2) and French (L1) words with dense neighborhoods in the nontarget language, as compared to words with a small number of cross-language neighbors, confirming the behavioral findings of competition of cross-language neighbors and further locating these at the time of lexico-semantic word processing. Nevertheless, this orthographic density N400 effect was more pronounced for the target words in the L2.

Words with Distinct Lexical Representations Across Languages

The ERP studies described above were all intended to provide an exclusively monolingual experimental context, in order to avoid any intentional activation of the nontarget language of bilinguals and to provide conclusive evidence of automatic and implicit cross-language activation. Nevertheless, given the extensive amount of orthographic and/or phonological features of the specific types of words used as critical materials across the two languages of interest (i.e., cognates, interlingual homophones, homographs, and cross-language orthographic neighbors), it could be argued that their mere presentation could have evoked the activation of the task-irrelevant language. In other words, the use of words with formal overlap across languages could be functioning as an artificial dual-language context. Thierry and Wu (2007; see also Wu & Thierry, 2010) were able to circumvent this limitation and examine whether the task-irrelevant native language of bilinguals is active during reading in the L2 by using a cross-script language combination (Chinese–English) with an implicit manipulation in an exclusive L2 (English) context with non-cognate items. Specifically, late and proficient Chinese–English bilinguals were presented with English word pairs upon which they had to perform a semantic relatedness task, while both behavioral and electrophysiological measures were being collected. Critically, the Chinese translations of half of the prime-target pairs, both semantically related and unrelated, shared a character. The authors found an N400 effect of semantic relatedness, with smaller N400 amplitudes for semantically related pairs, as compared to semantically unrelated word pairs. While the presence of the "hidden" overlapping Chinese character did not modulate participants' behavioral performance, it modulated the ERPs recorded. Pairs with the critical hidden Chinese character repetition yielded smaller N400 effects than those without it. This finding demonstrated that the Chinese non-cognate translation was active while reading exclusively English (L2) words. It furthermore provided strong evidence supporting the view that bilinguals keep both languages active during reading, irrespective of the presence of cues of cross-language activation.

Even if cognates are quite common across most alphabetic languages, the vast majority of words composing each language are non-cognates. Along these lines, Midgley, Holcomb, and Grainger (2009a, 2009b) examined the ERP correlates of the L1 word processing advantage over L2 word reading in three groups of English learners of French who differed in their levels of L2 proficiency. Participants were presented with English and French non-cognate items, blocked by target language, and performed a go/no-go semantic categorization task by pressing a button upon presentation of an animal name. The authors analyzed the EEG responses on the critical non-animal trials. When late and relatively low proficient English–French bilinguals were tested, differences between L1 and L2 words appeared at the N400 time window, when the settling of a form-meaning mapping is proposed to take place (e.g., Van Petten & Kutas, 1990). At posterior sites, the N400 effect was larger for L1 words, while at anterior sites the effect arose 150 ms later for L2 words. When late and non-proficient French–English bilinguals completed the experiment with a different set of non-cognates, a similar pattern of L1 and L2 N400 differences emerged, with L2 words producing a prolonged N400 peak. However, when a more proficient group of French–English bilinguals was tested, the latency shift in the N400 for L2 words as compared to L1 words was reduced as compared to the less proficient groups. This pattern suggests that the N400 effects are sensitive to the relative language dominance of the bilinguals, as well as to their level of L2 competence during single-word reading.

Bilingual Word Processing in Dual-Language Experimental Settings

Studies on bilingual visual word recognition in which words from both languages of a bilingual are present allow manipulation of cross-language relationships of interest. In this way, researchers ensure that bilinguals will activate both languages at the intended level, thus obtaining information on the structure of the cross-language sub-lexical, morphological lexical, and semantic links (e.g., Duñabeitia, Dimitropoulou, Morris, & Diependaele, 2013). It has been mainly in the last decade that these dual-language experimental settings have been combined with the use of ERPs to examine the time course of the processes underlying bilingual visual word recognition and the cross-language interactions taking place at the word level. Apart from studies investigating cross-language interactions at distinct levels of processing, there are a growing number of bilingual ERP studies addressing a different aspect of bilingual word processing: the computation of L1 and L2 language membership during reading and the relative degree of L1 and L2 activation. Critically, this type of evidence can be exclusively obtained in experiments in which participants are exposed to a dual-language setting.

Cross-Language Lexico-Semantic Interactions During Word Reading

ERP evidence from studies testing the cognitive processes underlying cross-language lexico-semantic interactions is crucial in establishing the structure of the bilingual lexico-semantic representational system. Interestingly, to isolate cross-language interactions exclusively motivated by the semantic relationship between words of both languages, most researchers have opted for examining the processing of words that apart from their semantic relationship do not have any formal cross-language overlap, such as cross-language non-cognate semantic associates (e.g., *window* in English and *casa*, the Spanish word for "house") or non-cognate translations (e.g., *casa-house*; see Perea, Duñabeitia, & Carreiras, 2008, for review). Only a small number of studies have addressed the issue of whether cross-language lexico-semantic interactions are modulated by cross-language ortho-phonological overlap.

Rodriguez-Fornells, Rotte, Noesselt, Heinze, and Münte (2002) as well as Martin, Dering, Thomas, and Thierry (2009) used ERPs to examine whether or not the non-attended language is semantically active when bilinguals process words of the target language, and reported a highly contrasting pattern of effects. In the first of these studies, Rodríguez-Fornells et al. presented early and relatively balanced Spanish–Catalan bilinguals with a stream of Catalan and Spanish words and nonwords. The authors performed a lexical frequency manipulation on both Catalan and Spanish items. Participants were asked to provide a motor response each time an existing Spanish word appeared on the screen and to disregard Catalan words and nonwords. The ERPs showed a lexical frequency N400 effect only for the Spanish target words, with reduced N400 amplitude emerging for high frequency Spanish words, as compared to low frequency Spanish words. Interestingly, no such frequency effect was observed for the non-attended Catalan words, leading the authors to propose that the words of the nontarget language were rejected before accessing their meaning. In other words, the findings by Rodríguez-Fornells et al. supported the view that there is no nontarget language lexico-semantic activation when bilinguals consciously process words of a given language. Martin et al. (2009) addressed the same issue in two experiments and obtained a very different pattern of effects. The authors presented early and fluent English–Welsh bilinguals with English and Welsh words. In the first experiment, bilinguals were asked to indicate whether English words had more or less than five letters and to disregard the Welsh words. In the second experiment participants performed the same task but this time, they were asked to make the same judgment on the length of the Welsh items and to disregard the English words. Martin et al. manipulated the language membership and the semantic relationship of the words, thus creating within and across-language word pairs that were either semantically related or unrelated. The behavioral results confirmed that the letter-length judgment task had effectively oriented participants' attention away from the semantic relationship across words, since there were no behavioral effects of semantic relatedness in either of the two experiments. Still, in both experiments, significant N400 effects of semantic relatedness were found, with smaller N400 amplitudes to English and Welsh targets preceded by semantically related words as compared to when they were preceded by unrelated words. Notably, these N400 semantic priming effects were present for both the target and the non-attended languages. The set of findings reported by Martin and colleagues shows firstly that at least early and highly proficient bilinguals process L2 words up to the semantic level. Secondly, these findings suggest that both languages of a bilingual are active and effectively processed up to the conceptual level, even when no attentional resources are placed upon them, contrasting with the claims of Rodríguez-Fornells et al.

In a different line of studies aimed at testing explicit bilingual lexical access, Guo et al. (2012), Palmer, Van Hooff, and Havelka (2010), and Yudes, Macizo, and Bajo (2010) combined the translation recognition task with ERPs to examine different aspects of the time course of semantic activation and its interaction with the formal properties of the words. Yudes et al. (2010) recorded the behavioral and electrophysiological responses of late and fluent Spanish–English bilinguals, who were asked to decide whether English (L2) cognate and non-cognate target words were the correct translation of Spanish (L1) primes. Results showed that the cognate status of the words affected responses, with cognates yielding faster and more accurate responses as well as reduced N400 amplitudes, as compared to non-cognates. These findings suggest that the formal overlap across cognate translations facilitated their co-activation at the lexico-semantic level.

Palmer et al. (2010) used the same task in combination with ERP recordings to examine the time course of the co-activation of non-cognate translations across the two translation directions (i.e., from L1 to L2 and vice versa), with late but proficient English–Spanish and Spanish–English bilinguals. As expected, the results showed that incorrect translations led to larger N400s than correct translations, while they also found a modulation of the N400 effect as a result of the translation direction. The N400 difference between correct and incorrect translations was larger in the L1-to-L2 translation direction than in the opposite translation direction, replicating the asymmetric pattern of translation effects repeatedly reported at the behavioral level. The authors interpreted their findings as providing evidence in favor of the RHM's claim that the L1 translation is automatically activated when processing L2 words (e.g., Kroll & Stewart, 1994).

Although they also used the translation recognition task, Guo et al. (2012) exclusively focused on the trials involving incorrect translations to investigate the time course of cross-language lexico-semantic activation with two groups of proficient Chinese–English bilinguals. The authors manipulated the formal (phonological) or semantic relationship of the incorrect Chinese translation with the correct translation of the target English words, to examine whether the activation of the L1 translation equivalent is influenced by formal or semantic factors during L2 processing. In their first experiment, the authors used a long 750 ms stimulus onset asynchrony (SOA) and in their second experiment they used a shorter SOA (300 ms). In both experiments, the authors found comparable interference effects for Chinese words either semantically or phonologically related to the correct Chinese translation of the target English (L2) words, as compared to their corresponding control Chinese words.

Critically, for the long SOA (750 ms) conditions, ERPs revealed a different time course for semantic and phonological interference effects. Chinese distractors that were semantically related to the translation of the English targets elicited smaller N400 as compared to the unrelated controls, and significant differences were also observed for the LPC that varied as a function of the polarity of this component (anterior vs. posterior LPC). Form-related distractors elicited a larger P200 and LPC than control words. Briefly, the P200 is a positive-going electrical potential peaking at about 200 ms that is modulated by the degree of overlap between visually presented items at multiple levels (e.g., ortho-phonological, semantic). At the shorter SOA (300 ms), the pattern of ERP effects was slightly different. While semantic interference effects were obtained in the N400 and the LPC time windows, the form interference effect was only obtained in the later LPC time window. The overall pattern of effects indicated that even highly proficient bilinguals automatically activate the L1 translation of L2 words. Importantly, the fact that a P200 effect was found for L1 distractors formally related to the L1 translation of the L2 targets demonstrated that this activation of the L1 translation takes place before L2 words are semantically processed.

Unlike these studies where participants were required to overtly process the relationship between translation equivalents, Alvarez, Holcomb, and Grainger (2003) and Geyer, Holcomb, Midgley, and Grainger (2011) opted to examine the ERP pattern of mental translation while bilinguals were asked to perform a task that did not direct their attention to the processing of translation pairs. In these studies, bilinguals were presented with a stream of words in both languages and researchers manipulated the relationship and the language of words presented in subsequent trials, in such a way that participants could be presented with the same word twice (i.e., within-language repetition), with non-cognate translation equivalents, or with unrelated words from the same or from different languages. This design allowed the authors to directly contrast the effects caused by within-language repetition to those caused by cross-language repetition (i.e., translation), as well as the ERP effects associated to a language change across two successive trials (see below for further discussion on language-switching effects). Alvarez et al. tested relatively low proficient English–Spanish bilinguals and asked them to perform semantic categorizations on the words presented. Geyer et al., on the other hand, tested very proficient Russian–English bilinguals who were asked to perform a generalized lexical decision task. In both studies, within-language repetition effects emerged in the 150–300 ms time window, with more negative-going peaks for within-language unrelated words than for within-language repetitions. Alvarez et al. also found such an effect for words preceded by their non-cognate translation as compared to their cross-language controls. Within the N400 time window, within-language repetition effects as well as translation effects were obtained for both L1 and L2 items in both studies. Still, while in the study by Geyer and colleagues target language did not interact with relatedness either in the within or in the cross-language conditions (i.e., a symmetric pattern of within-language repetition and translation N400 effects), in the study by Geyer and colleagues the magnitudes of the N400 effects were modulated by the target language. Specifically, the Spanish (L2) N400 within-language repetition effect was larger than the English (L1) one. With regard to the N400 translation effect, larger N400 differences were obtained for L1 words preceded by their L2 translation than vice versa. The asymmetric pattern of translation effects found by Alvarez et al. was interpreted as resulting from the clearly different levels of proficiency in the L1 and L2 of the participants, while the symmetric pattern of effects reported by Geyer et al. was thought to be related to the more balanced levels of proficiency of the Russian–English bilinguals tested.

Although these ERP studies have been very informative with regard to the patterns of cross-language lexico-semantic interaction underlying the processing of translation equivalents, a number of recent ERP studies have opted for subliminally presenting one of the two translation equivalents, in order to ensure that the effects obtained are not contaminated by cognitive processes related to the conscious perception of the lexico-semantic relationship between the translations (see Altarriba & Basnight-Brown, 2007). These studies followed the masked priming procedure, which, in its standard version, involves the presentation of a pattern mask (e.g., #####), followed by a brief presentation of a prime word (around 50 ms), which is immediately replaced by a target word (see Forster & Davis, 1984). Under these conditions, participants are not able to consciously perceive the prime, thus making it impossible to intentionally process its relationship with the target. In an adaptation of the masked priming task to study translation equivalents (i.e., the masked translation priming paradigm), the prime is either the translation of the target or an unrelated control word of the nontarget language. In the last two decades, numerous behavioral studies have examined masked translation priming effects with non-cognate translations and have established a clear asymmetric pattern, at least with late and non-balanced bilinguals performing lexical decisions on the targets: faster responses are found to the L2 targets when preceded by their L1 translation, as compared to those preceded by an unrelated L1 word, while this facilitation is smaller and more elusive in the opposite translation direction (e.g., Gollan, Forster, & Frost, 1997; Jiang, 1999; Jiang & Forster, 2001; see Dimitropoulou et al., 2011a, for review).

Following the increasing amount of interest in masked translation priming effects, recent ERP studies have used the masked priming paradigm to test the time course of the activation of non-cognate translation equivalents under strategy-free conditions (Duñabeitia, Dimitropoulou, Uribe-Etxebarria, Laka, & Carreiras, 2010; Hoshino, Midgley, Holcomb, & Grainger, 2010; Midgley et al., 2009a, 2009b; Schoonbaert, Holcomb, Grainger, & Hartsuiker, 2011). Non-cognate masked translation priming ERP effects have thus far been reflected in two components: the N250 and the N400. Within the monolingual masked priming ERP literature, more negative-going waves within both the N250 and the N400 time windows have been associated with more effortful target processing. With regard to the N250 effect, this processing has been proposed to tap into the mapping of pre-lexical to whole-word form representations, whereas the N400 effect has been proposed to tap into the form-to-meaning mapping process (see Grainger & Holcomb, 2009, for review).

Midgley et al. (2009a, 2009b) were the first to study ERP non-cognate masked translation priming effects by recording electrophysiological responses in L1 (English) and L2 (French) words preceded by their non-cognate translations or by unrelated masked primes. The authors also included a within-language repetition priming condition, along with its within-language control, to directly compare the ERP modulation caused by within and across-language repetitions. Participants were relatively proficient unbalanced French–English bilinguals who were asked to perform a semantic categorization task with L1 and L2 targets. In the cross-language conditions, they obtained a clearly asymmetric pattern of effects, with significant N250 and N400 modulations emerging only for L1 primes and L2 targets (i.e., more negative-going waveforms in the unrelated as compared to the translation condition). Within the masked translation priming framework, the N250 component was proposed to reflect the first instance of the automatic co-activation of the lexical representations of translations, while the N400 effect was interpreted as the electrophysiological marker of the match between the semantic representations of the prime and target. Furthermore, Midgley et al. found that, unlike the masked translation priming effects, the masked identity priming effects were symmetrical across languages (L1-L1 and L2-L2). Hoshino et al. (2010) fully replicated Midgley et al.'s findings using the same design and task but with a cross-script language combination (Japanese and English). Hoshino et al. also reported a significant modulation of the N/P150 ERP component (a component that is assumed to reflect early processes involved in mapping visual features onto higher level form representations), with L2 targets producing more negative-going waves when preceded by unrelated L1 primes, as compared to when preceded by their L1 translation pairs. The authors related this effect to an additional advantage in the processing of the L1 masked primes due to the visual cue provided by the change in script (i.e., primes were written with Kanji characters and targets with Roman letters). The non-cognate ERP masked translation priming asymmetry has also been reported with the lexical decision task used in the majority of the behavioral non-cognate masked translation priming studies. However, in this study by Schoonbaert et al. (2011) testing English–French bilinguals, the usual asymmetric pattern was observed in the presence of significant N250 and N400 effects in the backward translation direction (from the L2 to the L1), as well as larger forward (from the L1 to the L2) translation priming N400 effects. Surprisingly, the asymmetry reported in the N250 time window showed an inverse pattern, with a larger effect appearing in the backward translation direction. Nonetheless, it should be noted that this study did not use classic masked priming procedures as the prime presentation times used were relatively long (120 ms).

In contrast to these asymmetric patterns mentioned above, Duñabeitia, Dimitropoulou et al. (2010) results are the only ones to report symmetric ERP non-cognate masked translation priming effects. In their study, they tested a group of Spanish–Basque simultaneous and balanced bilinguals of similar characteristics to the ones tested in a parallel behavioral version of this study (see Duñabeitia, Perea, & Carreiras, 2010). Participants performed a semantic categorization task on both Spanish and Basque targets. Fully replicating the behavioral findings obtained with the Spanish–Basque bilinguals, the results of this ERP study showed significant bi-directional N400 masked translation priming effects which were of similar magnitude across the two translation directions. However, unlike other

ERP masked translation priming studies, Duñabeitia et al. did not find any N250 modulation in the cross-language translation conditions. Nonetheless, it should be noted that this study is the only ERP study so far to test balanced simultaneous bilinguals in a masked translation priming paradigm and that this could be the reason underlying the differences between this study and other studies testing unbalanced sequential bilinguals.

Processing Language Membership During Word Reading

Identifying language membership is not a straightforward process and recent research has shown that both the linguistic profiles of the readers and the specific characteristics of the target words play a crucial role during this process (see Casaponsa, Carreiras, & Duñabeitia, 2014, for review). De Bruijn, Dijkstra, Chwilla, and Schriefers (2001) investigated the processing of the language membership of Dutch–English interlingual homographs by using a multiple priming paradigm in combination with ERP recordings. In this paradigm, two word primes preceded each target. Dutch–English bilinguals were presented with the interlingual homograph as the second word of the triplet, preceded either by a Dutch or by an English word. The manipulation of the language membership of the first word of the triplet aimed at providing either an L1 or an L2 context that could bias the processing of the language membership of the subsequently presented interlingual homograph (e.g., Dutch [L1] context: *ZAAK–ANGEL–HEAVEN* or English [L2] context: *HOUSE-ANGEL-HEAVEN*, where *ANGEL* corresponds to an interlingual homograph). The authors also manipulated the semantic relationship between the interlingual homograph and the subsequently presented target in such a way that the target could either be semantically related to the English meaning of the interlingual homograph (e.g., *HOUSE-ANGEL-HEAVEN*) or unrelated to either the English or the Dutch meaning of the homograph (e.g., *HOUSE-ANGEL-LAUNCH*). Participants were asked to perform a generalized lexical decision task upon target presentation (i.e., are all the items of the triplet Dutch and/or English words?), while behavioral and electrophysiological responses were being recorded. The behavioral and ERP responses showed significant semantic priming effects, with shorter reaction times and smaller N400 amplitudes found for targets preceded by semantically related English words as compared to unrelated words. However, this effect was not modulated by the language context provided by the initial word of the triplet, either at the behavioral or at the electrophysiological level. This pattern indicates that the prior presentation of an L1 word is not enough to suppress the activation of the L2 meaning of the interlingual homograph.

In contrast to the study by De Bruijn et al. (2001), which addressed semantic processing of words without a clear language membership (interlingual homographs), ERP studies examining the way in which bilinguals identify the two languages during bilingual visual word recognition have focused on the processing of the language membership of items that are not shared across languages (i.e., non-homographic non-cognate words). In fact, most of the existing information in this regard has been obtained from studies investigating the cost observed when bilinguals switch between languages, referred to as the *code-switching cost* (e.g., Alvarez et al., 2003; Chauncey, Grainger, & Holcomb, 2011). The seminal studies exploring this cost were reported in the word production domain (e.g., Costa & Santesteban, 2004), in which the code-switching cost has been classically interpreted as the result of the inhibition needed to suppress the nontarget language following a code-switch (see Kroll, Bobb, Misra, & Guo, 2008, for review). However, code-switching costs have also been reported when bilinguals switch between languages during reading (e.g., Grainger & Beauvillain, 1987; Thomas & Allport, 2000). In the comprehension modality, where word processing proceeds in a bottom-up way, there is an on-going debate regarding the locus of the code-switching cost and the computation of the language membership of words. On the one hand, the BIA model proposes that the language membership of a given word is readily computed as part of the lexical processing of the word (e.g., Grainger & Dijkstra, 1992), while, on the other hand, the BIA+ model proposes that the language tag is provided after the activation of the lexical representation, at a later task-dependent processing stage (e.g., Dijkstra & van Heuven, 2002).

Alvarez et al. (2003) were the first to report such visual code-switching costs with ERPs in the word comprehension domain. As previously described, these authors presented late, relatively non-proficient English–Spanish bilinguals with mixed lists of English and Spanish words while participants performed a semantic categorization task. When examining the processing of the language switches by comparing the ERP responses to words preceded by words of the same language with the responses to words preceded by words of the other language, the authors found greater negativities in the 500–700 ms time window for words involved in a language switch. However, in the earlier N400 time window this code-switching cost was only obtained in one of the directions at test (namely, when the switch was from the L1 to the L2). This pattern of asymmetric ERP code-switching costs across the two language switching directions suggested that with relatively non-proficient unbalanced bilinguals, the language switch in the L2-to-L1 direction takes longer to emerge. Furthermore, given the fact that the L1-to-L2 code-switching cost appeared in the N400 time window, when form-to-meaning integration is proposed to take place, the authors concluded that their findings supported the view that the inhibition associated with the code-switching cost is applied at the lexical level, as the BIA model proposes (e.g., Van Heuven et al., 1998).

Geyer et al. (2011) followed the same experimental procedure, but this time with a generalized lexical decision go/no-go task, to test whether the same pattern of asymmetric ERP effects as reported by Alvarez et al. (2003) would be obtained with highly proficient late Russian–English bilinguals who were living in an L2 context. Although the authors' primary focus was on translation effects, their design also allowed for the investigation of code-switching costs with bilinguals with a higher level of L2 competence than those tested by Alvarez et al. Geyer and colleagues predicted that, given the high level of competence in the second language of their bilinguals, L1–L2 processing differences would be diminished and thus a more symmetric pattern of code-switching costs would be obtained across the two switching directions, in terms of both the timing and the magnitude of the effects. The results confirmed their prediction. Small but significant code-switching costs appeared in the N400 component and in the 500–850 ms time window, which did not interact with the direction of the language switch. The symmetric pattern of code-switching costs found with highly proficient bilinguals suggested that these effects depend on the relative level of L1 and L2 competence, as the BIA model proposes (Grainger & Dijkstra, 1992). Moreover, the fact that these ERP code-switching costs were found in trials that did not require an overt response was thought to further support the BIA model's account of code-switching costs, since unlike the Inhibitory Control (IC) assuming that bilingual language processing is in a way analogous to non-linguistic physical actions that consist of different mental task schemas that compete with each other to reach a goal, and unlike BIA+ models, the BIA model proposes that these processing costs should not be affected by task demands.

Even though the language switches were not predictable in these studies, since they randomly appeared within the experimental list, participants were able to intentionally process the language membership of the test items, since they were explicitly and overtly presented. In other words, the obtained pattern of code-switching costs could have been considerably contaminated by strategic and attention-related processes. As in the case of the studies investigating the processing of translation equivalents, researchers exploring switch cost effects in language comprehension have also taken advantage of the implicit presentation provided by the masked priming paradigm to study the electrophysiological correlates of the code-switching cost in bilingual visual word recognition. Similarly, Chauncey, Holcomb, and Grainger (2008; see also Chauncey, Holcomb, & Grainger, 2011) tested French–English bilinguals with an intermediate level of English proficiency. Participants performed a go/no-go semantic categorization task on French or English targets preceded by unrelated masked primes either in English or in French. Chauncey et al. found evidence for automatic code-switching costs across primes and targets in both the N250 and the N400 time windows as compared to non-switching conditions. These effects depended on the directions of the language switches, with L1 primes and L2 targets eliciting larger negativities in the N250 time window as compared to non-switch trials, and L2 primes preceding L1 targets eliciting more negative-going peaks within the N400 range (see also Midgley et al., 2009a, 2009b). Given the appearance of code-switching costs in the early N250 component (typically linked to the interface between sub-lexical and whole-word representations; Grainger & Holcomb, 2008, 2009), the authors proposed that in trials involving a prime-target language switch, the activation of the language node corresponding to the prime inhibits the lexical representations of the target's language, giving rise to this early switch-cost effect, in line with the interpretation of code-switching costs offered by the original BIA model (e.g., Van Heuven et al., 1998). Following this early stage, the inhibitory activation is then propagated to the later N400 component, when the form–meaning integration takes place. The asymmetric pattern of code-switching costs across the two switching directions was also taken as evidence in favor of the BIA model, since it was thought to reflect the fact that bilinguals with a higher L1 than L2 proficiency level take longer to process L2 primes and consequently to activate the L2 language node, hence causing a delayed language switching cost in the L2-to-L1 switching direction. The findings reported by Duñabeitia, Dimitropoulou et al. (2010) testing simultaneous and balanced Basque–Spanish bilinguals partly supported this view. Their results replicated the overall pattern of masked priming ERP code-switching effects first reported by Chauncey, Grainger, and Holcomb (2008). Significant N250 and N400 masked code-switching costs were obtained, but these effects were highly similar for both Basque and Spanish targets. The fact that the asymmetric pattern of code-switching costs found with unbalanced bilinguals disappeared when native-like balanced bilinguals were tested, as in the case of Duñabeitia et al., indicates that the relative level of proficiency of the bilinguals and the AoA of each language critically modulate these effects.

Reading Sentences

L2 Sentence Reading

The main research question in sentence reading in a second language is whether L1 and L2 processing mechanisms are completely different and separate, or, in contrast, whether L1 and L2 processes are completely or partially overlapping. Several studies have shown that even highly proficient L2 readers have persisting difficulties in syntactic processing (Johnson & Newport, 1989; Weber-Fox & Neville, 1996), while others report cases of L2 readers who have reached a native-like processing level (Birdsong, 1992). It is important to keep in mind that similarities in behavioral measures between L1 and L2 processing do not necessarily mean that the underlying cognitive and neural processing is equivalent in a first and a second language. As argued above, neurophysiological measures can contribute essential evidence to this debate, and recording electrophysiological activity can provide detailed information on the timing and degree of activation of neural networks. Comparing ERP activity in L1 and L2 reading thus provides important information on the similarities and differences between L1 and L2 processing, even where the final behavioral outcome is identical in the two situations. Consequently, it is not surprising that the ERP technique has been a popular approach to investigating the question of convergence or divergence of L1 and L2 sentence processing mechanisms.

In their seminal study, Ardal and colleagues (1990) explored ERPs elicited by semantic violations in sentences. They showed that the classic N400 effect was delayed in bilinguals. Interestingly, this was the case in bilinguals reading in L2, but also (to a lesser extent) in bilinguals reading in their first language. In this study, the N400 component was not modulated by AoA of L2 (earlier or later than age 11). In contrast, one of the most widely cited ERP experiments on bilingual sentence reading published by Weber-Fox and Neville in 1996 showed a different pattern of results. They tested the influence of L2 (English) age of acquisition on sentence reading.

The sentences read were either correct or included semantic or syntactic violations. They showed that semantic processing was slower in bilinguals who acquired their L2 after the age of 11, as compared to monolinguals and bilinguals who had acquired L2 earlier in life (N400 effects for semantic violation delayed in time). The N400 effect in this context is classically defined as the magnitude of the difference in amplitude between the N400 elicited by a semantic violation and that elicited by a semantically correct word in the same position (Kutas & Hillyard, 1980). Syntactic processing was affected by delays in L2 learning as short as 1–3 years and was revealed by a difference in the morphology of the ERP components reflecting syntactic processing (N125, left-lateralized negativity and P600). The authors argued that delays in L2 exposure might be associated with a reduced left hemisphere specialization in language processing (see Proverbio, Leoni, & Zani, 2004, for similar arguments of lesser degree of hemispheric lateralization during L2 processing).

These two early studies show the variability in results obtained in L2 sentence processing. However, interestingly, in spite of the considerable variability in the results—depending on the languages tested, the paradigms used in the studies and other multiple factors—the considerable literature on bilingual sentence reading has led to several consistent observations. In the following sections, we review these findings.

The majority of ERP studies on bilingual sentence reading have explored semantic and morphosyntactic processing. There are also many parallel ERP studies in the auditory modality (testing auditory sentence comprehension instead of sentence reading; see Friederici, Steinhauer, & Pfeifer, 2002; Hahne, 2001; Hahne & Friederici, 2001; Isel, 2007; Mueller, 2006; Mueller, Hahne, Fujii, & Friederici, 2005; Ojima, Matsuba-Kurita, Nakamura, Hoshino, & Hagiwara, 2011; Rossi, Gugler, Hahne, & Friederici, 2006; Sanders & Neville, 2003), but in the current chapter we will only focus on sentence reading. In most of these studies, L2 and L1 readers were presented with correct sentences, sentences containing semantic violations (e.g., *Peter likes to eat eggs and* <u>*socks*</u> *for breakfast*, violation underlined) and/or sentences containing syntactic or morphosyntactic violations (e.g., *Peter like to eat eggs for breakfast*, *Peter wants a egg for breakfast*). The ERP waves elicited by the critical words of the sentences (specific words making the sentence semantically and/or syntactically correct or incorrect) were compared across groups. As discussed in the following section, most studies have underscored the importance of two major factors in bilingualism: age of acquisition and proficiency. These factors are probably interrelated and we will see that both are important in determining the timing of semantic and morphosyntactic processing during sentence reading (see Moreno & Kutas, 2005).

Semantic Processing During L2 Sentence Reading

It has been repeatedly observed that the classic N400 effect associated with semantic processing is slightly reduced in amplitude and/or delayed in time when reading in the L2 as compared to the L1 (e.g., Ardal et al., 1990; Kutas & Kluender, 1991; Moreno & Kutas, 2005; Newman, Tremblay, Nichols, Neville, & Ullman, 2012;

Ojima, Nakata, & Kakigi, 2005; Proverbio et al., 2004; Weber-Fox & Neville, 1996, 2001; see Kutas, Moreno, & Wicha, 2009; Moreno et al., 2008; Mueller, 2005, for reviews). Some studies have revealed that the N400 semantic congruity effect is reduced and/or delayed even in the L1 of a bilingual, as compared to a monolingual (Ardal et al., 1990; Meuter, Donald, & Ardal, 1987; but see Proverbio, Cok, & Zani, 2002 for no monolingual vs. L1 bilingual differences). This result suggests that acquiring more than one language has consequences for semantic processing, even when reading in the native language. Proverbio et al. (2002) suggested that even if L2 semantic processing can reach a native-like level, differences between semantic processing in reading in L1 vs. L2 persist. The authors compared monolinguals and early, highly proficient bilinguals reading sentences where the last word was correct, semantically incorrect, or syntactically incorrect (e.g., *Something tickled me*, *I would like to read this dog*, *He ended up by forgetting his to go out*). They observed an N400 effect in the two groups when reading semantic violations. Nevertheless, the ERP response to semantic violations was left-lateralized in bilinguals and right-lateralized in monolinguals. So far, the results from semantic processing during L2 reading are relatively congruent, showing a reduced and delayed N400 effect. This delay in the N400 semantic effect is interpreted as the result of an extended lexical search and/or a lower degree of automaticity for L2 processing (Ardal et al., 1990; Weber-Fox & Neville, 1996).

In a recent study, Braunstein et al. (2012) did not only investigate L2 semantic processing during sentence reading by using semantic violations, but also using word cloze probability. Word cloze probability in a sentence is the probability that a speaker would complete the sentence by using a particular word (e.g., "road" would have a 90 % cloze probability if 90 % of the readers would complete the sentence *Peter was walking down the*… by using the word "road"). Braunstein et al. presented L2 readers with sentences ending with a semantically incongruent word, a highly expected word (i.e., high cloze probability), or an unexpected, but still semantically correct word (i.e., low cloze probability). Semantically incongruent and unexpected final words elicited N400 effects that were delayed in L2 relative to L1 readers. Thus, this experiment revealed the notion that semantic processing during L2 sentence reading varies as compared to L1 sentence reading, even in a "natural" context with no semantic violations.

As argued earlier, research on bilingual sentence reading has also benefited from experiments studying aspects of semantic processing other than brain reactions to semantic violations. Interesting paradigms already used in investigations of L1 sentence processing have been successfully applied to L2 reading. For instance, a productive line of research has been established on the active role of the reader when processing sentences in L2, by exploring anticipation and semantic integration. Anticipation processes refer to the active role of individuals when predicting upcoming linguistic information in a sentence. That is, comprehenders do not read or listen passively, but rather try to anticipate probabilistically the next words (or discourse topics) that are likely to appear in the sentence (e.g., DeLong, Urbach, & Kutas, 2005). This ability is fundamental in language processing, inasmuch as it reduces processing load and helps interlocutors to free resources to plan their utterances during a conversation, thus smoothing communication (Pickering & Garrod, 2007). By integration processes we refer to those processes that allow comprehenders to combine semantic information contained in the sentence with world knowledge (e.g., Hagoort et al., 2004). That is, comprehenders not only pay attention to the meaning of the words, but also analyze how such information matches their knowledge about the world. These integration processes are fundamental to a proper understanding of the communicative act, and therefore go beyond mere linguistic analysis. It is largely admitted that difficulties in L2 sentence reading stem, at least partly, from incomplete or faulty syntactic parsing (see below) or slow/more difficult semantic access (see previous discussion). This may lead to deficiencies in anticipation and integration processes as compared to L1 processing. However, as we will see below, the number of studies exploring these issues is markedly low, and more research is needed in order to assess whether (and how) these processes function in L2 sentence comprehension.

Active word anticipation during L2 sentence reading has been recently explored, but has to be further investigated since results so far are not conclusive. Martin et al. (2013) explored word anticipation in L2 sentence reading by comparing English natives (L1 readers) and Spanish–English bilinguals (L2 readers). Each sentence ended in an expected or unexpected noun (high vs. low word cloze probability). Importantly, in this study, no semantic violation was used (see also Braunstein et al., 2012). Contrary to L1 readers, L2 readers failed to show the N400 modulation elicited by word anticipation. The authors concluded that L2 readers do not anticipate upcoming words during sentence comprehension to the same extent as L1 readers do. However, strong conclusions cannot be drawn so far because the same research group has recently obtained discrepant results testing different samples of bilinguals. In another study, Foucart, Martin, Moreno, and Costa (2014) compared L1 and L2 readers in a similar sentence reading task in which sentences ended in expected vs. unexpected words (e.g., *She has a nice voice and always wanted to be a singer/an artist*). Spanish monolinguals (L1 readers), Spanish–Catalan early bilinguals (L1 readers), and French–Spanish late bilinguals (L2 readers) were compared. The classic N400 pattern revealing word anticipation was observed in the three groups of participants. The data revealed that (at least when the L1 and L2 are closely related) L2 readers are able to anticipate upcoming words in a similar way to L1 readers and that identical processes are involved. It is important to note that Martin et al. used phonological article–noun agreement in English as the critical manipulation, and that this phonological rule does not exist in Spanish (the participants' L1). Foucart et al. (2014) used gender agreement in Spanish as the critical manipulation, this rule being similar in French (the first language of L2 readers in their study). Thus, word anticipation during L2 sentence reading seems to be highly influenced by L1/L2 similarities. Further investigation is needed before drawing any strong conclusions since only two studies have explored word anticipation in L2 reading, thus far. Nevertheless, these recent results show the interest of using critical paradigms from research on L1 sentence reading in order to investigate further semantic processing during L2 sentence reading without resorting to semantic violations.

To summarize, most studies on semantic processing during sentence reading reveal similarities between L1 and L2 processing (Kotz & Elston-Güttler, 2004; Ojima et al., 2005). The main and most frequent observation of a delayed and reduced N400 effect in L2 relative to L1 processing suggests that differences between L1 and L2 semantic processing are mostly quantitative and not qualitative in nature. This quantitative difference might be due mainly to a slowdown or decrease in efficiency of semantic processing mechanisms in an L2 (see Kutas et al., 2009; Moreno et al., 2008; Mueller, 2005 for reviews). Note that whether L2 semantic processing is native-like or not seems mainly to depend on proficiency.

Syntactic and Morphological Processing During L2 Sentence Reading

In studies of L1 syntactic processing using violation paradigms, the most common ERP result observed is a biphasic pattern of early negativity (normally left-lateralized), followed by P600 effects. The P600 is a positive deflection in the waveform which peaks around 600 ms after stimulus onset, and the P600 effect is classically defined as the magnitude of the difference in amplitude between the P600 elicited by a syntactic violation and that elicited by a syntactically correct word in the same position (Hahne & Friederici, 1999). It is considered to reflect processes of reanalysis and syntactic repair (Friederici, 2002; Osterhout, Holcomb, & Swinney, 1994) and, more globally, the difficulty of syntactic integration (Kaan, Harris, Gibson, & Holcomb, 2000). The earlier left-lateralized negativity is commonly considered an index of early-stage syntactic processing mechanisms (Friederici, 2002). Although this is a common pattern for L1 morphosyntactic processing, for bilingual syntactic processing, the results have not been consistent. Ojima et al. (2005) observed that syntactic violations embedded in sentences elicited a left-lateralized negativity in English natives and Japanese–English highly proficient late bilinguals (tested in English, their L2). This left-lateralized negativity was not observed in Japanese–English low proficient late bilinguals also tested in L2. However, Chen, Shu, Liu Zhao, and Li (2007) reported a very different pattern of results revealing that subject–verb agreement violations elicited a late *anterior* negativity and no P600 effects in native Chinese–L2 English readers of intermediate proficiency. Steinhauer, White, Cornell, Genesee, and White (2006) also investigated ERPs during syntactic violation reading in L1 and L2 and found a different pattern of results again. The classic pattern of left-lateralized negativity followed by a P600 component was observed in native readers. However, low proficiency L2 readers elicited only a P600 effect, suggesting that L2 readers cannot automatically process syntactic violations in a native-like manner within the first 500 ms following the violation (Steinhauer et al., 2006; see Hahne, 2001; Weber-Fox & Neville, 1996 for similar arguments). Weber and Lavric (2008) tested English natives and German–English bilinguals reading English sentences containing morphosyntactic violations in the final word (e.g., *The door had been locked* vs. *The door had been locks*).

Morphosyntactic violations elicited the classic P600 effect in English natives, while German–English bilinguals reading in L2 showed the expected P600 effect, but also an N400 effect. This observation of an N400 modulation by morphosyntactic violations might suggest that processing morphosyntactic violations in an L2 relies on the lexico-semantic system (see also Tanner, McLaughlin, Herschensohn, & Osterhout, 2013). However, note that some studies on L1 sentence processing have also reported N400 modulations in response to morphosyntactic violations (e.g., Bentin & Deutsch, 2001; Severens, Jansma, & Hartsuiker, 2008).

As highlighted by the above studies, and as will be discussed below, the results observed in studies of L2 morphosyntactic sentence processing are not as consistent as those of semantic processing. We propose that this larger discrepancy in studies of L2 morphosyntactic processing relative to L2 semantic processing might be explained, at least partly, by the wider range of different processes and paradigms used in research on morphosyntactic processing. This area has explored a wide variety of paradigms (not only using syntactic violations) and has tested very different types of structure (subject–verb agreement, gender or number agreement, referential ambiguities, closed- vs. open-class word processing, for example), leading to a considerable variability in the phenomena observed. Interestingly, Osterhout et al. (2006) investigated syntactic violation processing in L2 English learners of French after 1, 4, or 8 months of instruction and compared it to processing in French natives. Before 4 months of instruction, subject–verb agreement violations in L2 readers elicited an N400 effect. In contrast, after 4 months of instruction, subject–verb agreement violations were processed similarly in L1 and L2 readers (similar P600 effects were observed in the different groups; see White, Genesee, Drury, & Steinhauer, 2007 for similar findings). However, determiner–number agreement violations elicited a P600 effect in L1 readers but not in L2 readers, even after 8 months of instruction. These results are congruent with our main assumption that processing of different morphosyntactic structures in L1 and L2 will vary.

In the above studies of morphosyntactic processing during L2 sentence reading, L1 and L2 readers were presented with correct sentences and sentences containing morphosyntactic violations. Interestingly, and unlike the field of semantic L2 processing, a large number of studies have also explored morphosyntactic processing using paradigms in which no violation is inserted in the sentences. In the following section, we review these studies on syntactic integration and ambiguities.

Regarding *function* word processing, several studies have suggested that a similar network is involved in L1 and L2 in the processing of function words, with delays reported in L2 readers (Weber-Fox & Neville, 2001). Briefly, function words in English are *closed-class* words that are primarily related to grammatical aspects of sentence processing (e.g., articles, determiners, conjunctions, prepositions). In contrast, *open-class* (*content*) words primarily convey referential meaning (e.g., nouns, verbs, adjectives). Neville, Mills, and Lawson (1992) reported that high-frequency closed-class and open-class words all elicited left anterior negativities that were larger and earlier for closed-class words than for open-class words. Those waves were, respectively, termed N280 and N350 by the authors (see also Brown, Hagoort, & ter Keurs, 1999). Later, in the 400–600 ms time window, closed-class words elicited a broad negative shift that was not observed when open-class words were displayed (Brown et al., 1999; Neville et al., 1992). Weber-Fox and Neville (2001) tested function word reading in bilinguals of various ages of acquisition and monolinguals. They observed the N280 component in all participant groups, the peak being delayed in bilinguals who learned English after the age of 7. The processing of open-class words was similar in all groups (similar N350 latencies and distributions; Weber-Fox & Neville, 2001). As for the processing of regular and irregular participles, Hahne, Mueller, and Clahsen (2003, 2006) observed that L2 readers use two different processing routes to process regular and irregular words, as is the case in L1 readers: rule-based decomposition might be used to process regular words, and lexical storage might be used to process irregular words. The authors concluded that relatively automatic morphosyntactic processes could be implemented in the L2 reader's brain (see Clahsen & Felser, 2006a). Note that this implementation certainly depends on the complexity of the morphosyntactic rule under consideration and the similarity of the rules in L1 and L2 (Mueller, 2005). The idea that some morphosyntactic processes can be implemented similarly in the L1 and L2 was also argued by Kotz, Holcomb, and Osterhout (2008), who suggested that L2 readers can process syntax in a similar way to L1 readers (see also Diependaele, Duñabeitia, Morris, & Keuleers, 2011, for a similar argument applied to the processing derivational morphology by bilinguals). The authors tested temporary syntactic ambiguity that was a language-specific phenomenon of English. They observed that the classic P600 effect elicited by temporary syntactic ambiguity was similar in English natives and Spanish–English highly proficient bilinguals.

Despite the high variability in results, some influential interpretations of L1/L2 differences in syntactic processing assume that the syntactic representations that L2 readers compute during sentence comprehension are shallower and less detailed than those of native readers (Clahsen & Felser, 2006b). It could also be that L2 readers underuse syntactic information in L2 processing (Marinis et al., 2005; Weber-Fox & Neville, 1996) and/or that difficulties in mapping discourse onto syntax constrain L2 performance (Hopp, 2009).

In summary, it could be argued that it is difficult to achieve L2 morphosyntactic processing in a native-like manner, and it is almost never the same as that of the L1 (Ojima et al., 2005; see Kotz, 2009, for a review). Note, however, that this idea of fundamental differences in language processing in the L1 vs. L2 is challenged by several studies showing similar electrophysiological signatures in syntactic processing during sentence reading in L1 and L2 (e.g., Bowden, Sanz, Steinhauer, & Ullman, 2007; Steinhauer, White, & Drury, 2009; see also Rossi et al., 2006 in the auditory modality). Similar P600 effects for L1 and L2 readers are reported in several studies suggesting that controlled syntactic reanalysis and repair can be acquired by L2 readers and achieved in a native-like way in some domains of grammar (Clahsen & Felser, 2006a; Kotz et al., 2008; but see Chen et al., 2007, and Ojima et al., 2005, for no P600 modulation by syntactic violations in L2 readers). Early, more "automatic" morphosyntactic processing mechanisms (reflected by left-lateralized negativities), however, seem to be much more difficult to acquire, especially where L2 proficiency is low. This absence of early anterior negativities seems to be a typical result in studies of syntactic processing in late and low proficiency L2 readers (Bowden et al., 2007; Hahne et al., 2003, 2006; Mueller et al., 2005; Steinhauer et al., 2006; Weber-Fox & Neville, 1996; see Kutas et al., 2009; Moreno et al., 2008; Mueller, 2005; Steinhauer et al., 2009 for reviews). However, it seems that L2 readers do not have systematic problems with all aspects of grammar, but more with real-time computation of complex hierarchical representations (Clahsen & Felser, 2006a). We can conclude, therefore, that L2 adult readers use lexico-semantic cues during sentence reading as native readers do, but are less able to cope with structurally based parsing strategies (see Clahsen & Felser, 2006b; Papadopoulou & Clahsen, 2003).

Influence of Proficiency and Age of Acquisition in L2 Sentence Reading

As stated above, Braunstein et al. (2012) explored the influence of L2 proficiency on semantic processing during sentence reading. They observed that the N400 latency for semantically correct but unexpected words was modulated by L2 proficiency. This was an indication that L2 proficiency influences the speed of semantic processing during sentence reading. The fact that L2 proficiency contributes to the speed of semantic processing has been also observed in previous studies (Ardal et al., 1990; Moreno & Kutas, 2005). However, other studies have suggested that the age of L2 acquisition, and not proficiency, is the factor determining L1/L2 processing differences (e.g., Proverbio et al., 2004). Hence, it still remains to be clarified whether proficiency and/or age of acquisition are able to capture and explain the differences between L1 and L2 sentence reading. In fact, these two variables are usually highly correlated, making it difficult to disentangle the contribution of one *vis-à-vis* the other (see Moreno & Kutas, 2005). Interestingly, Newman et al. (2012) also observed that the N400 amplitude to semantically appropriate words was larger for participants with lower English proficiency. Thus, it seems that language proficiency affects semantic processing in general, even during correct sentence reading. The proficiency of the language used in reading does not only affect semantic violation processing. It rather seems that proficiency modulates the main semantic processing system (see Newman et al., 2012). Note that previous studies have already revealed that proficiency in L2, but also in L1, modulates semantic processing (e.g., Moreno & Kutas). For instance, Pakulak and Neville (2010) observed that P600 amplitude was correlated with the individual's proficiency in L1.

The importance of proficiency in native-like acquisition of processing mechanisms has also been highlighted by studies on syntactic processing in L2. The early anterior negativity and P600 effects have been shown to be strongly affected by proficiency in the L2 (e.g., Bowden et al., 2007; Steinhauer et al., 2006; see also Hahne, 2001; Hahne & Friederici, 2001; and Rossi et al., 2006, for studies testing the auditory modality). In order to investigate effects of language proficiency, Weber-Fox, Davis, and Cuadrado (2003) tested syntactic and semantic processing during sentence reading in normal and highly skilled readers. They observed that the N280 elicited by closed-class words was delayed in normal readers relative to highly proficient readers. In contrast, the N350 component elicited by open-class words did not vary between normal and highly proficient readers. Interestingly, this pattern was consistent with previous observations with various groups of bilinguals (Weber-Fox & Neville, 2001). It seems that open-class word processing is independent of language proficiency (in native or non-native language). In contrast, the speed of closed-class word processing is a sensitive index of language proficiency (Weber-Fox et al., 2003). The following broad negative shift typically observed for closed-class but not open-class words (Brown et al., 1999; Neville et al., 1992) was also affected by language proficiency, but the interpretation of this finding is still unclear given the little evidence in the literature for such effects. Finally, the classic N400 component elicited by semantic violations was significantly smaller in amplitude in highly proficient L1 readers relative to normal L1 readers. This result, in accordance with previous literature, suggests that semantic processing is highly affected by language proficiency: the larger the proficiency (in L1 or L2), the smaller the reliance on sentence context for word recognition (Holcomb, Coffey, & Neville, 1992; Weber-Fox et al., 2003).

Age of L2 acquisition is also an important variable to take into account when investigating L2 sentence processing, with some studies providing support for the idea of a critical or sensitive period for L2 semantics and syntactic processing, ending more or less at the age of puberty (Clahsen & Felser, 2006a; Weber-Fox & Neville, 1996). It has been observed that L2 readers who acquire their second language before late childhood can often reach native-like L2 processing levels, while L2 acquisition later than that shows evidence of difficulties in semantic and/or syntactic processing or L2 processing that is not similar to L1 readers (Johnson & Newport, 1989; Ojima et al., 2005; Weber-Fox & Neville, 1996, 2001). Thus, it has been argued that the earlier the L2 is acquired, the more native-like the linguistic processing achieved (but see Birdsong, 1992; Bongaerts, 1999; Friederici et al., 2002; Rossi et al., 2006, for evidence against the critical period hypothesis). Some of the strongest evidence for a critical or sensitive period in L2 acquisition comes from studies of syntactic processing in L2 sentence reading. The fact that evidence of early automatic morphosyntactic processing mechanisms (reflected by left-lateralized negativities) is scarcely ever reported in late L2 readers, would seem to indicate that the development of some complex underlying syntactic processes requires prerequisites available only during childhood, in most cases at least.

Note that some authors consider that AoA and proficiency should not be considered as independent factors since they are usually highly correlated. Thus, the influences of age of acquisition and proficiency can be considered together, as proposed by Moreno and Kutas (2005). Accordingly, the latency of the N400 component varies with proficiency and age of acquisition. They report that the latency of the N400 effect in L2 readers is positively correlated with AoA and negatively correlated with fluency. Finally, one current view on proficiency and age of acquisition effects in L2 sentence reading is that age of acquisition would mainly affect syntactic, morphological, and phonological processing, while proficiency would better explain lexical and semantic processing (Hernandez & Li, 2007; Johnson & Newport, 1989; Weber-Fox

& Neville, 1996). It is important to note that AoA and proficiency are generally considered to be the two main factors influencing L2 sentence processing. Nevertheless, several other factors are important to take into account when investigating sentence processing in a second language. For instance, L1 to L2 similarity highly modulates sensitivity to grammatical processing in L2. In Tokowicz and MacWhinney's (2005) study, English–Spanish bilinguals were presented with grammatically correct or incorrect sentences varying in the type of syntactic processing under study. They compared a function that is similar in English and Spanish (noun–verb agreement; e.g., *El niño están jugando* [The boy (singular) are (plural) playing]) and a function that is different in English and Spanish (determiner–noun agreement; e.g., *Ellos fueron a un fiesta* [They went to a (masculine) party (feminine)]). The classic P600 effect was significant in L2 readers for subject–verb agreement violations (constructions that are formed similarly in L1 and L2) but not for determiner–noun agreement violations (constructions that differ between L1 and L2). The authors concluded that L2 readers process L2 syntax in a way that depends on the similarities between L1 and L2 (Tokowicz & MacWhinney, 2005).

Gillon-Dowens et al. (2010) also explored the influence of L1 to L2 similarities on electrophysiological correlates of L2 morphosyntactic processing. Spanish natives, serving as a control group, were presented with grammatically correct sentences and with sentences containing number or gender agreement violations in Spanish. The results showed the typical expected ERP pattern of an early anterior negativity followed by a P600 effect in response to both agreement violations. A group of late L1 English–L2 Spanish readers also showed a qualitatively similar early negativity-P600 pattern. More importantly, however, in this group, unlike the native Spanish group, quantitative differences (i.e., amplitude and onset latency differences) between the two types of agreement violations were observed, with greater amplitudes for number than for gender processing. Since number agreement is a syntactic feature of both Spanish and English and gender agreement is not a feature of English, the authors concluded that transfer processes from L1 to L2 significantly influenced syntactic processing (see Gillon-Dowens et al., 2011, for similar argument on L1 transfer observed in ERP data). Sabourin and Stowe (2008) also investigated electrophysiological correlates of L1 grammar transfer during L2 sentence processing. They presented Dutch natives with verbal domain dependency violations, or grammatical gender violations. L1 readers showed a significant P600 effect for both types of violations. German–Dutch bilinguals also showed a P600 effect for both violations. L2 readers who had a Romance language as their L1 showed a P600 effect only for verbal domain violations. The authors concluded that neural correlates for L1 and L2 are similar when morphosyntactic processing rules are shared between the two languages (e.g., verbal domain dependency in Dutch, German and Romance languages). However, rules that differ in L1 and L2 do not result in similar neural processes in the two languages (e.g., grammatical gender rules that differ in Dutch and Romance languages; Sabourin & Stowe, 2008). Foucart and Frenck-Mestre (2011) confirmed that L2 morphosyntactic processing is modulated by the similarity of the rules at play in L1 and L2. They studied various gender agreement violations in French natives and German–French advanced bilinguals. Foucart and Frenck-Mestre reported similar P600 effects in the two groups when agreement rules were similar in French and German, whereas no P600 effect was observed in L2 readers when agreement rules differed across languages. These studies demonstrate how the fine-grained temporal resolution of the ERP technique is currently being widely explored by cognitive neuroscientists investigating cross-linguistic influence effects (see also Gabriele, Fiorentino, & Bañón, 2013).

To conclude, we would like to point out a potential shortcoming of ERP research on bilingual sentence reading that could explain part of the inconsistent results observed up until now. Almost all ERP studies in the field have reported data averaged over trials and participants. However, variability across participants is very high when considering bilinguals. L2 processing skills are subject to strong individual variability that can lead to problems of interpretation of classic ERP components that are computed by averaging across individuals. This potential problem was extensively studied by Tanner et al. (2013). They revealed significant individual differences in L2 readers' ERP responses to morphosyntactic violations. Strikingly, the violations elicited N400 effects in some participants and P600 effects in others. Better accuracy in sentence acceptability judgments was associated with larger brain responses in both the N400 and P600 time windows. In a similar study, Tanner et al. (2013) replicated the observation that morphosyntactic violations elicit N400 effects in some L2 readers and P600 effects in others (despite homogeneous high L2 proficiency and long-term L2 exposure). In this study, they also showed that higher L2 proficiency is associated with larger brain responses. These results, in accordance with Osterhout et al.'s (2006) study, suggest that L2 learners go through distinct stages of learning, with inter-individual variability in the rate of progression through these stages (see White et al., 2007; Steinhauer et al., 2009, for similar proposal). More importantly, these studies also reveal that investigation of between-subject variability can sometimes be highly informative. Therefore, investigation of individual ERP responses should be combined with standard analysis of grand average waveforms in future research on L2 sentence reading. In the next sections we attempt to provide current and future users of EEG technology with a guide for good practices that can provide the field of multilingual reading with a snapshot of how the multilingual brain functions when processing words and sentences in the native and in the nonnative language(s).

Summary and Conclusions

The purpose of this chapter was to offer a complete description of the most relevant EEG studies on the field of bilingual written word and sentence comprehension. The time window in which a given letter string passes from being a mere sequence of graphemes to acquiring the word status takes around 300 ms (see Duñabeitia & Molinaro, 2013, for review), and recent evidence from monolingual and multilingual ERP studies has been extremely valuable in demonstrating how lexical representations from the first and second languages are acquired, consolidated and accessed by readers with different proficiency profiles in their L2. The initial part of this chapter covered this issue in depth. Following this initial section, we discussed evidence showing strong L1–L2 interactions at the ortho-phonological, lexical, and semantic levels of processing, both in single-language and in dual-language contexts. As seen in the first section of this chapter, the complexity of the possible cross-linguistic interactions at different levels of processing (e.g., cognates, interlingual homophones and homographs, cross-linguistic ortho-phonological neighbors) highly complicates the proposal of a unified account of bilingual lexical access that could be ultimately used to explain bilingual visual word recognition. Rather, EEG data on the processing of ortho-phonologically similar or dissimilar lexical representations suggest that any integrated theoretical account of bilingual visual word recognition would necessarily require the previous full description of the different time courses associated with the processing of each type of cross-linguistically interacting lexical representations, which would require the conscientious design of well-motivated individual experiments tapping onto the multiple stages yielding lexical access (i.e., orthography, phonology, morphology, semantics). In the second section of this chapter, we summarized the evidence from L1 and L2 sentence reading EEG experiments, trying to elucidate whether native and nonnative syntactic and semantic processing mechanisms partially overlap, or alternatively, whether they are different in essence. As in the case of studies on bilingual single-word processing, studies exploring semantic and morphosyntactic processes in a sentence context have revealed that both age of L2 acquisition and proficiency in the nonnative language are crucial in determining the timing of semantic and morphosyntactic processing during sentence reading. EEG data in this regard have consistently shown that the time course associated with syntactic and semantic processing of phrases is highly sensitive to the readers' profiles and that ERP components can vary qualitatively and quantitatively as a function of L2 AoA and proficiency. Finally, in the following section, we present a brief description of a series of good practices that could be used by researchers in this field to effectively use EEG recordings to assess how ortho-phonological, lexical, morphosyntactic, and semantic processing varies between the native and the nonnative language in bilinguals. In this section, we refer to specific aspects related to the design of an experiment, the data acquisition process, and the critical stage of data analysis, which will ultimately guide novice researchers in the process of becoming experts in the field.

Best Practices for Running ERP Experiments

ERPs calculated based on raw EEG data are electric voltage differences between an active electrode and a reference one, usually on the order of few microvolts. This measure provides researchers with a multidimensional dependent measure of brain activity, since ERPs can vary in:

(a) *Time*: ERPs have high temporal resolution that is reliable on the order of tens of milliseconds.
(b) *Amplitude*: ERP components can be larger or smaller (in terms of voltage differences) when comparing two conditions.

(c) *Polarity*: ERPs are positive/negative deflections, as compared to a zero-level baseline.
(d) *Scalp Topography*: ERPs can differ in their topographical distribution, possibly reflecting the recruitment of different brain substrates.

Experimental Designs

The researcher can observe variation of some ERP components on a lot of dimensions (based on the levels of the independent variable). These variations determine the so-called ERP effect that is assumed to reflect different patterns of brain activation in different experimental conditions/groups. Usually, the definition of an ERP effect is not that simple and requires a good knowledge of preceding ERP literature referring to a specific component. Even in this case, however, the experimental design employed in an ERP experiment is critical to interpreting a specific ERP effect. For this reason, it is important to define a simple and robust design, without too many conditions. This is even more important in the case of bilingual experiments, where usually the by-group comparison is mandatory, or at least, highly recommended (e.g., L1 vs. L2 readers). Experimental designs with a large number of manipulations embedded in complex factorial designs typically make the interpretation of the possible emerging interactions highly complicated, and in ERP studies, this becomes even more critical considering the aforementioned multidimensionality of the ERP dependent measure. This is the so-called *explosion of dimensions* that is responsible for the misinterpretation of the results in many ERP studies. Hence, the first general rule in designing an ERP experiment is to design well-motivated, self-explanatory, and robust experimental designs. Keep things simple. The specific ERP effect that the experiment is targeting should always be kept in mind. The experimenter should be able to advance predictions about a specific ERP component based on the available scientific literature. This is a critical issue, since a good set of predictions can considerably ease the interpretation of some specific ERP effects. An *exploratory expedition* without specific and well-motivated predictions is the worst possible approach for a young researcher.

After the definition of a sufficiently powerful and clear experimental design, the researcher has to select a good-enough number of observations to enter in each cell of this experimental design (namely, a sufficient number of items/participants per condition; for an in-depth discussion see also Picton et al., 2000). EEG measures are inherently noisy, since sensors (typically Ag/AgCl electrodes) do not only capture brain activity but also electromagnetic noise originating from both the lab environment and the participant under study. As stated in the introductory part of this chapter, ERPs derive from averaging (time intervals of EEG raw recordings after time-specific events) across multiple repetitions of the same class of stimuli. Through this averaging procedure, what is unrelated to the stimulus (i.e., noise) disappears, while the resulting ERPs would reflect brain reactions that are constant across those repetitions. Thus, before entering the lab for the experimental session, it is important to construct a material set with a high number of observations for each experimental condition, given that the larger the number of repetitions, the larger the amount of noise deleted from the ERP estimation. This is defined by the signal-to-noise ratio (SNR) estimate that reflects the level of desired signal with reference to the level of background noise. It has been demonstrated that the SNR for ERPs becomes higher while increasing the number of observations for a specific event. However, after a certain number of repetitions, the SNR curve reaches a ceiling level and does not increase significantly even if we increase the number of repetitions. The exact number of items per condition depends on the ERP component(s) of interest. As another general rule, we may consider that the earlier a component is, the larger the number of observations we should have. In typical psycholinguistic designs, a good number of items per cell can vary between 50 and 100 (also depending on the available material in the linguistic scenario of interest). It should be kept in mind that a large number of items per experimental condition will be very useful for an optimal evaluation of single-subject's ERPs. As already mentioned above, individual variability is an important factor in L2 learning studies (e.g., Tanner et al., 2013), and in order to be able to capture this variability, experimental designs must be powerful enough to allow single-subject analysis.

A similar argument can be applied to the selection of a reasonable number of participants per group. In a typical ERP experiment, for the grand-averaged data (ERPs averaged across multiple participants) the number of participants should not be less than 18–20. Of course, a higher number of participants will lead to a better (statistical) estimation of group-level effects. Nonetheless, experimental goals can differ, since some researchers can be more interested in group-level differences, while others may focus on individual variability. In both cases, we suggest collecting data at least from 30 participants per group, since the group-level estimation will be more (statistically) solid, and there will be more individual variability available to estimate which single-subject parameters (e.g., IQ, working memory span) reliably modulate an ERP effect of interest.

After the selection of proportionally balanced groups of participants and the identification of a solid set of materials, the experimenter will face the process of selecting both the online and the offline technical parameters related to EEG acquisition. In the two following sections we will discuss the *data acquisition* parameters for collecting EEG data and the *data analyses* parameters for extracting reliable ERP results.

Data Acquisition

Data acquisition mainly refers to the technical parameters we use for running an ERP experiment. The overall rule in this section is to acquire as much data as possible. However, how much data does it make sense to acquire? Currently, many EEG systems come with caps in which a pre-defined number of electrodes (typically 19, 32, 64, 96, or 128) are fixed to constitute a standardized pattern often referred to as the 10-20 or 10-10 systems (Picton et al., 2000). The selection of the number of electrodes to use has to be done considering that the larger the number of electrodes, the longer it will take us to mount the EEG cap for each participant before starting the experiment. In some experiments, it is worth using a higher number of electrodes since we aim to determine with great detail the scalp distribution of a specific ERP component. With a high-density array of electrodes (greater than 64) it will be possible to perform source analyses for a specific ERP deflection (keep in mind that reliable source reconstruction also requires a detailed 3D model of the anatomy of the participants' head). Otherwise, in typical cognitive neuroscience studies, a smaller number of electrodes (less than 64) are sufficient to obtain good ERP data. This is due to the fact that electric dipoles (the ones determining the familiar ERP deflections) do not have high spatial resolution when recorded over the scalp, and can be measured over multiple electrodes. In summary, it should be kept in mind that by increasing the number of electrodes, it will not increase sensitivity in the estimation of an ERP component.

One of the most important choices is the selection of the right position for the reference. In many studies, the best solution is to select the left mastoid as the online reference (to which all the scalp electrodes will refer; see Molinaro, Barber, & Carreiras, 2011, for a discussion of how ERP effects can depend on the reference choice). This solution is considered as the "standard," even if there is no counterargument in selecting alternative positions such as the right mastoid, the vertex electrode (Cz), or the tip of the nose. It is convenient in all cases to also record activity over the two mastoids to operate an offline re-reference of the data. This is the standard solution used in many studies on language processing, and it is important to use it for cross-study comparisons. On the other hand, using both mastoids as a reference (linked-mastoid solution) is not optimal (see Picton et al., 2000).

Finally, we will discuss the importance of a correct selection of the amplifier settings. These last technical parameters are crucial since this is one of the crucial steps in which there is no way back once we have made a choice. Two critical parameters are of special interest: the sampling rate and the online filters. As for the sampling rate (i.e., the number of points per second characterizing the voltage signal) the value should be larger than 250 Hz (250 points per second). This value also determines the temporal resolution at which you can estimate an ERP effect (in the 250 Hz case the resolution is of 1 data point every 4 ms). Our suggestion is to record the EEG signal at 500 Hz–1 kHz, since modern computers have enough memory resources to process such data. Having a high sampling rate is not too critical for ERP analyses (where the "real" temporal resolution for estimating a cognitive process is various tens of milliseconds). However, more sophisticated analyses can require a lot of data points (time-frequency decomposition for example, see below) to provide reliable information in short time periods. Using a low sampling rate will preclude us from reanalyzing ERP data with more sophisticated techniques.

EEG is made up of multiple signals overlapping at multiple frequencies (Hz, number of cycles per second for an oscillatory signal). In principle, ERP data mainly reflect brain activity in the EEG frequency range up to 30–40 Hz. Usually, this frequency limit is applied to "clean" the signal when it is too noisy and to provide smooth waveforms. It should be noted, however, that neural activity can oscillate at very high frequencies (intracranial recordings can measure activity at 200 Hz, the so-called ripples). Currently, sophisticated methods of EEG analyses have identified significant cognitive activity around ~100 Hz (upper values for gamma activity). For this reason, amplifiers are usually set to record activity from very low frequencies (0.01 Hz, oscillatory activity that correlates with the functional magnetic resonance imaging blood-oxygen level dependent or BOLD signal; Hipp, Hawellek, Corbetta, & Engel, 2012), to very high ones (400–500 Hz). Again, this wide frequency range captures much more than what a researcher would need for a regular ERP study, but it should be kept in mind that such recordings can be extremely useful in the near future if one aims toward approaching a set of data from multiple perspectives.

Finally, it is worth mentioning that data acquisition should be performed trying to keep all the electrode impedance values below 5–10 kΩ (indicating good electric contact between the electrodes and the scalp), and the experimental participant should be invited to stay quiet and relaxed during the recording session. A good acquisition session always implies less work during the following stages of data analysis. For this reason, the experimenter should be responsible for reducing the amount of non-electrophysiological activity in the EEG recordings.

Data Analyses

After data acquisition, the skeleton of a typical ERP analysis usually follows a few fixed steps with a not-too-flexible order: re-referencing, filtering, epoching, baseline correction, artifact evaluation, and average. This order can be changed, but some steps are a prerequisite for others, and this notion should always be kept in mind.

The process of re-referencing is critical to establish the reference electrode/s for all the active electrodes. Usually, in language experiments, there is the tendency to select a "balanced" reference (i.e., the algebraic average of the activity recorded from the two mastoids). In some cases, when the number of electrodes is high enough (greater than 64) an average reference can be computed reflecting an imaginary electrode that corresponds to the average activity of all electrodes. This location-independent solution represents a good alternative to other methods, but it can be biased by a few electrodes with high amplitudes when the number of sensors is too low.

ERP activity is usually measured in the frequency bands between 0.1 and 30 Hz. These frequencies are the ones with the most power and, for this reason, survive the averaging procedure. The selection of the lower threshold is critical and should not be too high, especially when analyzing long-lasting ERP components (with an overall estimated oscillation cycle of 1 s). In these cases, a high-pass filter slightly higher than 0.5 Hz is capable of washing out such components. This is due to the fact that filters do not usually cut exactly at the indicated frequency, but they can reduce the power in a larger frequency window. For instance, a 0.5 Hz high-pass filter can reduce at 50 % the power of frequencies up to 2 Hz.

The selection of the time interval of interest should consist of (a) a brief interval preceding the target event in which no difference between two experimental conditions is expected and (b) an interval after the target event that should be long enough for a good (visual and statistical) estimation of the modulation of the specific ERP component(s). In principle, the interval preceding the event (i.e., the baseline) should reflect no brain activity whatsoever. However, in many experiments, it is possible that some activity is going on before the presentation of a target stimulus (for example, in sentence processing experiments). In those cases, it should be explicitly stated why the activity in the pre-stimulus interval should not differ between conditions.

The baseline correction procedure tries to bring to the zero-voltage level the activity before the event that we just discussed. Usually, the baseline activity is subtracted (or divided) from the post-stimulus whole epoch. As previously indicated, there are many assumptions behind this procedure. For example, it is assumed that before stimulation there is no activity in the brain, but this is highly unlikely. Our brain is always working, and even if this pre-stimulus activity can be of no interest for the researcher, he/she should be aware of possible (expectation) effects originating before the presentation of a stimulus that can affect post-stimulus activity.

"Cleaning" the EEG signal from artifacts (eye movements and muscular activity in particular) is also a crucial step. In fact, ERPs are averages of activity across multiple segments of an EEG recording, and it is well known that algebraic means can be largely biased by outliers, so that it is possible to observe effects in the ERP waveforms that can reflect some residual artifactual activity (especially when some artifacts, like blinks, can be orders of magnitude larger than the brain activity of interest). The most common way to deal with an artifact is to remove the epochs containing it by visually inspecting all the epochs for each participant. This can be a lengthy procedure, but certainly it is the most reliable one. (Note that there is no computer algorithm as good at recognizing artifacts as the human eye/brain). Some automatic and semiautomatic procedures are available to reject "bad" epochs. These procedures provide good approximations for the ERP extraction, but they are not perfect.

A different approach aims at "correcting" artifacts instead of at deleting them (consequently saving a higher number of epochs for the analysis). The most efficient approach employs *Independent component analysis* (ICA) to detect components reflecting eye movements, and to regress them from the EEG signal (Delorme & Makeig, 2004). However, this procedure is not risk-free, since the critical activity of interest can be involuntarily removed from the data. Our suggestion is to employ correction procedures only in cases in which it is really necessary, and for the whole dataset (i.e., for all participants). In fact, these procedures alter the reality of the signal and should be used only as a last resort, and always supervised by some expert's eyes.

The last step of the analysis involves the averaging stage. After averaging across multiple epochs and grand-averaging across participants, the ERP waveforms will be available for visual inspection and statistical evaluation. By-item analyses are not very common in the ERP literature, although they can be really interesting in some cases (DeLong et al., 2005; Molinaro, Carreiras, & Duñabeitia, 2012). We highly recommend that researchers consider this option, given its high informative value.

As already mentioned, single-subject averages are the main basis for statistical analyses. In general, two types of information can be employed to quantify some components. For some long-lasting ERP components, the averaged amplitude from a specific time interval of interest can be extracted across all electrodes. This can be done for ERP components with a duration of more than 200 ms (such as, for example, the N400; Kutas & Hillyard, 1980). A different approach is typically used for the estimation of earlier, more "peaky," and short-lasting components. In these cases the maximum value in a time interval (i.e., the amplitude of the peak expressed in microvolts) and the latency of that data point (in milliseconds) is extracted for each individual participant. This approach is critical in experiments exploring early components, since it can handle cross-individual variability in the latency of the peaks relatively well.

Recently, alternative types of analyses that aim at decomposing the EEG signal across time in different frequency components are being explored and used (e.g., Davidson & Indefrey, 2007; Molinaro, Barraza, & Carreiras, 2013; Pérez et al., 2012; for a review, see Bastiaansen & Hagoort, 2006). As indicated above, EEG represents the summation of different signal at different frequencies, and there is a family of methods (i.e., Hanning window, Multitapers, Wavelets) that can be employed to estimate the power of each specific frequency component across time. Usually, the accuracy of the time-frequency estimation depends on a large number of data points, specifically for high frequencies (such as gamma activity). Such decomposition of the EEG signal can be very useful in obtaining a detailed view of how different brain processes operate in parallel at different frequencies. At the same time, however, this type of analysis provides the researcher with another dimension or dependent factor that can increase the complexity of the data being analyzed. Hence, the researcher should carefully consider the general picture obtained from the different types of analysis, since the risk of interpreting some false positives is considerably high. In addition, the literature in this regard is not voluminous, since there are consistent scientific reports using time-frequency analyses only from the mid-90s, while ERPs have been used for almost half a century. Furthermore, the methods of analysis of the time-frequency domain are still under development. For these reasons, the researcher should strongly rely on his/her experimental hypothesis to interpret the data while keeping constantly in touch with the rich and active EEG world-wide community. Besides, the researcher should be aware of the existence of a large body of literature on methodological and practical aspects of EEG, and he/she should consider that a correct stepwise approach to the electrophysiology of the bilingual brain should start with a comprehensive reading and understanding of basic aspects of this technique that have been already compiled by experts in the field (e.g., Handy, 2009; Luck, 2005; Luck & Kappenman, 2012).

List of Keywords

Age of acquisition (AoA), Balanced bilinguals, Bilingual Interactive Activation Model (BIA), Bilingual word recognition, Code-switching cost, Cognate facilitation effect, Conceptual level, Critical Period Hypothesis, Cross-language neighbors, Cross-language repetition, Electroencephalography (EEG), Event-related potentials (ERPs), Forward translation, Function word processing, Generalized lexical decision task, Independent Component analysis (ICA), Inhibitory Control Model (IC), Inhibitory effect, Integration processes, Interference effects, Interlingual homograph, Interlingual homophones, Interlingual orthographic neighbors, Irregular words, Language nonselective activation, Language-independent lexical activation, Late bilinguals, Late positive component (LPC), Letter-length judgment task, Lexical access, Lexical competition, Lexical search, Lexical selection, Masked translation priming paradigm, Morphology, Morphosyntactic processing, Multiple priming paradigm, N400, Non-balanced bilinguals, Orthographic neighbors, P600, Phonological cognates, Phonological level, Pre-lexical, Revised Hierarchical Model (RHM), Semantic categorization task, Semantic integration, Semantic priming effects, Semantic processing, Semantic representations, Simultaneous bilinguals, Syntactic integration, Syntactic processing, Translation recognition task, Visual word recognition, Whole-word form representation, Word frequency, Word processing, World knowledge.

Review Questions

1. How do single ortho-phonological and lexico-semantic brain networks process different scripts?
2. What is the time course of lexical processing in bilinguals with varying degrees of AoA and proficiency in the L2?
3. Which is/are the electrophysiological marker/markers of bilingual lexical access?
4. What are the temporal dynamics of syntactic processing of typologically different vs. similar languages?
5. What's the oscillatory pattern associated with nonnative language processing?

Suggested Student Research Projects

1. Use EEG time-frequency information from word or sentence processing to predict whether individuals are native or nonnative speakers of a given language.
2. Create an EEG database of bilingual single-word reading using representative words from the language and a high number of participant's electroencephalographic responses.
3. Study the time course of language switching using different explicit and implicit paradigms and EEG recordings.
4. Explore the time course of bilingual syntactic processing by co-registering EEG activity and eye movements.

Related Internet Sites

Aghermann: http://johnhommer.com/academic/code/aghermann/
Basque Center on Cognition, Brain and Language: http://www.bcbl.eu/
Bilingualism Matters: http://www.bilingualism-matters.org.uk/
Brain and Language Research Institute: http://blri.weebly.com/
EEGLAB: http://sccn.ucsd.edu/eeglab/
ERP Info: http://erpinfo.org/erplab/erplab-download
Kutas Cognitive Electrophysiology Lab: http://kutaslab.ucsd.edu/
Neurocognition Laboratory: http://www.neurocoglaboratory.org/
Open EEG: http://openeeg.sourceforge.net/doc/sw/
SigViewer: http://sigviewer.sourceforge.net/
The Bilingualism Centre: http://www.bangor.ac.uk/bilingualism/

Suggested Further Reading

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT Press.

Luck, S. J., & Kappenman, E. S. (Eds.). (2012). *Oxford handbook of event-related potential components.* New York: Oxford University Press.

Van Hell, J. G., & Kroll, J. F. (2012). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism: Theoretical and applied approaches.* New York: Cambridge University Press.

Van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews, 64*(1), 104–122.

Acknowledgements The writing of this chapter was partially supported by grants PSI2012-32123 and PSI2012-32350 from the Spanish Government, ERC-AdG-295362 grant from the European Research Council, and by the AThEME project funded by the European Union (grant number 613465).

References

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic and translation priming experiments across languages. *Behavior Research Methods, 39*, 1–18.

Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meanings in two languages: An event-related brain potentials study of beginning bilinguals. *Brain and Language, 87*, 290–304.

Bastiaansen, M. C. M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. In C. Neuper & W. Klimesch (Eds.), *Event-related dynamics of brain oscillations* (pp. 179–196). Amsterdam: Elsevier.
Bentin, S., & Deutsch, A. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language, 45*, 200–224.
Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68*, 706–755.
Bongaerts, T. (1999). Second language learning data analysis. *Linguistics, 37*, 972–974.
Bowden, H. W., Sanz, C. S., Steinhauer, K., & Ullman, M. T. (2007). An ERP study of proficiency in second language. *Journal of Cognitive Neuroscience Supplement, E61*, 170.
Braunstein, V., Ischebeck, A., Brunner, C., Grabner, R. H., Stamenov, M., & Neuper, C. (2012). Investigating the influence of proficiency on semantic processing in bilinguals: An ERP and ERD/S analysis. *Acta Neurobiologiae Experimentalis, 72*(4), 421–438.
Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience, 11*(3), 261–281.
Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology. Human Perception and Performance, 25*, 137–148.
Carrasco-Ortiz, H., Midgley, K. J., & Frenck-Mestre, C. (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones. *Psychophysiology, 49*(4), 531–543.
Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: The impact of orthographic markedness. *Frontiers in Psychology, 5*, 424.
Chauncey, K., Grainger, J., & Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: A masked priming study with ERPs. *Brain and Language, 105*, 161–174.
Chauncey, K., Grainger, J., & Holcomb, P. J. (2011). The role of subjective frequency in language switching: An ERP investigation using masked priming. *Memory & Cognition, 39*, 291–303.
Chen, L., Shu, H., Liu Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition, 10*, 161–174.
Clahsen, H., & Felser, C. (2006a). How native-like is non-native language processing? *Trends in Cognitive Sciences, 10*, 564–570.
Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.
Coltheart, M., Jonasson, J. T., Davelaar, E., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. New York: Academic.
Costa, A., & Santesteban, M. (2004). Bilingual word perception and production: Two sides of the same coin? *Trends in Cognitive Sciences, 8*, 253.
Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology Section A, 38*, 367–393.
Davidson, D. J., & Indefrey, P. (2007). Inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Research, 1158*, 81–92.
De Bruijn, E., Dijkstra, A., Chwilla, D. J., & Schriefers, H. J. (2001). Language context effects on interlingual homograph recognition: Evidence from event-related potentials and response times in semantic priming. *Bilingualism: Language and Cognition, 4*, 155–168.
De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to the bilingual memory. *The Quarterly Journal of Experimental Psychology, 2*, 397–428.
DeLong, K., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21.

Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language, 64*, 344–358.

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1*, 51–66.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*, 284–301.

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011a). Two words, one meaning: Evidence of automatic co-activation of translation equivalents. *Frontiers in Psychology, 2*, 188.

Doctor, E. A., & Klein, D. (1992). Phonological processing in bilingual word recognition. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 237–252). Amsterdam: North Holland.

Duñabeitia, J. A., Dimitropoulou, M., Morris, J., & Diependaele, K. (2013). The role of form in morphological priming: Evidence from bilinguals. *Language and Cognitive Processes, 28*(7), 967–987.

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Research, 1359*, 142–154.

Duñabeitia, J. A., & Molinaro, N. (2013). The wide-open doors to lexical access. *Frontiers in Psychology, 4*, 471.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology, 57*(2), 98–107.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(6), 1340–1359.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The size of the cross-lingual masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology, 51*(2), 116–124.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 10*, 680–698.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*, 379–399.

Foucart, A., Martin, C. D., Moreno, E., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40*(5), 1461–1469.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78–84.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Science, 99*, 529–534.

Gabriele, A., Fiorentino, R., & Bañón, J. A. (2013). Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism, 3*(2), 213–232.

Geyer, A., Holcomb, P. J., Midgley, K. J., & Grainger, J. (2011). Processing words in two languages: An event-related brain potential study of proficient bilinguals. *Journal of Neurolinguistics, 24*, 338–351.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*, 1122–1139.

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology, 39A*, 295–319.

Grainger, J., & Holcomb, P. J. (2008). Neural constraints on a functional architecture for word recognition. In P. Cornelissen, P. Hansen, M. Kringelbach, & K. Pugh (Eds.), *The neural basis of reading* (pp. 3–32). Oxford: Oxford University Press.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: North Holland.

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass, 3*(1), 128–156.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science, 17*(12), 1021–1026.

Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: An electrophysiological and behavioral investigation of translation recognition. *Journal of Experimental Psychology. Learning, Memory and Cognition, 38*(5), 1165–1186.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123–141.

Hahne, A., Mueller, J. L., & Clahsen, H. (2003). Second language learners' processing of inflected words: Behavioral and ERP evidence for storage and decomposition. *Essex Research Reports in Linguistics, 45*, 1–43.

Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience, 18*(1), 121–134.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology. Learning, Memory and Cognition, 33*, 623–644.

Handy, T. C. (Ed.). (2009). *Brain signal analysis: Advances in bioelectric and biomagnetic methods*. Cambridge, MA: MIT Press.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*, 638–650.

Hipp, J. F., Hawellek, D. J., Corbetta, M., & Engel, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience, 15*(6), 884–890.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30*(1), 47–61.

Holcomb, P. J., Coffey, S. A., & Neville, H. J. (1992). Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology, 8*(2 & 3), 203–241.

Hopp, H. (2009). The syntax-discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition, 12*(4), 463–483.

Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain Research, 1371*, 100–109.

Hoshino, N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Research, 1344*, 159–172.

Isel, F. (2007). Syntactic and referential processes in second-language learners: Event-related brain potential evidence. *Neuroreport, 18*, 1885–1889.

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition, 2*, 59–75.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*, 32–51.

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*, 60–99.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of integration difficulty. *Language and Cognitive Processes, 15*, 159–201.

Kerkhofs, R., Dijkstra, A., Chwilla, D., & De Bruijn, E. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and ERP effects. *Brain Research, 1068*, 170–183.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain & Language, 109*, 68–74.

Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics, 17*(2), 215–235.

Kotz, S. A., Holcomb, P. J., & Osterhout, L. (2008). ERPs reveal comparable syntactic sentence processing in early bilinguals and monolinguals. *Acta Psychologica, 128*(3), 514–527.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Kroll, J. F., Bobb, S. C., Misra, M. M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*, 416–430.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*, 373–381.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Kutas, M., & Kluender, R. (1991). What is who violating? A reconsideration of linguistic violations in the light of event-related potentials. In H. J. Heinze, T. F. Münte, & G. R. Mangun (Eds.), *Cognitive electrophysiology: Basic and clinical applications* (pp. 183–210). Boston: Birkhauser.

Kutas, M., Moreno, E. M., & Wicha, N. Y. (2009). Code-switching and the brain. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 289–306). London: Cambridge University Press.

Lemhöfer, K., Dijkstra, A., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*, 585–611.

Lemhöfer, K. M. L., & Dijkstra, A. F. J. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language specific and generalized lexical decision. *Memory and Cognition, 32*(4), 533–550.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luck, S. J., & Kappenman, E. S. (Eds.). (2012). *Oxford handbook of event-related potential components*. New York: Oxford University Press.

Martin, C. D., Dering, B., Thomas, E. M., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the ‘active’ and the ‘non-attended’ language in early bilinguals. *NeuroImage, 47*, 326–333.

Martin, C. D., Thierry, G., & Démonet, J. F. (2010). ERP characterization of sustained attention effects in visual lexical categorization. *PloS One, 5*, e9892.

Martín, M. C., Macizo, P., & Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *British Journal of Psychology, 101*(4), 679–693.

Meuter, W., Donald, M., & Ardal, S. (1987). A comparison of first- and second-language ERPs in bilinguals. *Electroencephalography and Clinical Neurophysiology. Supplement, 40*, 412–416.

Midgley, K., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience, 23*(7), 1634–1647.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009a). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology, 46*, 551–565.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009b). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics, 22*, 281–300.

Midgley, K. J., Holcomb, P. J., van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research, 1246*, 123–135.

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47*, 908–930.

Molinaro, N., Barraza, P., & Carreiras, M. (2013). Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. *NeuroImage, 72*, 120–132.

Molinaro, N., Carreiras, C., & Duñabeitia, J. A. (2012). Semantic combinatorial processing of non-anomalous expressions. *NeuroImage, 59*(4), 3488–3501.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research, 22*(2), 205–220.

Moreno, E. M., Rodriquez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual processing. *Journal of Neurolinguistics, 21*(6), 477–508.

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152–174.

Mueller, J. L. (2006). L2 in a nut shell: The investigation of second language processing in the miniature language model. *Language Learning, 56*, 235–270.

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and non native speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience, 17*(8), 1229–1244.

Müller, O., Duñabeitia, J. A., & Carreiras, M. (2010). Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology, 47*(3), 455–466.

Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex, 2*(3), 244–258.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical-semantic processing in native and late learners of English: ERP evidence. *Journal of Cognitive Neuroscience, 24*(5), 1205–1223.

Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and the amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research, 70*, 197–205.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17*, 1212–1228.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 20*, 786–803.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience, 22*, 2728–2744.

Palmer, S. D., Van Hooff, J. C., & Havelka, J. (2010). Language representation and processing in fluent bilinguals: Electrophysiological evidence for asymmetric mapping in bilingual memory. *Neuropsychologia, 48*, 1426–1437.

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language, 68*(4), 315–332.

Pérez, A., Molinaro, N., Mancini, S., Barraza, P., & Carreiras, M. (2012). Oscillatory dynamics related to the unagreement pattern in Spanish. *Neuropsychologia, 50*(11), 2584–2597.

Petit, J. P., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2006). On the time-course of letter perception: A masked priming ERP investigation. *Psychonomic Bulletin & Review, 13*(4), 674–681.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, J. R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127–152.

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: An ERP study. *Neuropsychologia, 42*, 1636–1656.

Proverbio, A. M., Cok, B., & Zani, A. (2002). Electrophysiological measures of language processing in bilinguals. *Journal of Cognitive Neuroscience, 14*(7), 994–1017.

Rodriguez-Fornells, A., Rotte, M., Noesselt, T., Heinze, H. J., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature, 415*, 1026–1029.

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*, 2030–2048.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research, 24*, 397–430.

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing II. Segmentation, semantics and syntax in nonnative speakers. *Cognitive Brain Research, 15*, 214–227.

Schoonbaert, S., Holcomb, P. J., Grainger, J., & Hartsuiker, R. J. (2011). Testing asymmetries in noncognate translation priming: Evidence from RTs and ERPs. *Psychophysiology, 48*(1), 74–81.

Severens, E., Jansma, B. M., & Hartsuiker, R. J. (2008). Morphophonological influences on the comprehension of subject-verb agreement: An ERP study. *Brain Research, 1228*, 135–144.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research, 25*, 13–41.

Steinhauer, K., White, E., Cornell, S., Genesee, F., & White, L. (2006). The neural dynamics of second language acquisition: Evidence from Event-Related Potentials. *Journal of Cognitive Neuroscience Supplement, 1*, 99.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition, 16*, 367–382.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceedings of the National Academy of Science, 104*, 12530–12535.

Thomas, M. S. C., & Allport, D. A. (2000). Switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43*(1), 44–66.

Tokowicz, N., & MacWhinney, B. (2005). Implicit vs. explicit measures of sensitivity to violations in L2 grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27*, 173–204.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*, 780–789.

Van Hell, J. G., & Kroll, J. F. (2012). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism: Theoretical and applied approaches* (pp. 148–171). New York: Cambridge University Press.

Van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews, 64*(1), 104–122.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*(4), 380–393.
Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology. Human Perception and Performance, 28*, 616–627.
Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition, 35*(5), 938–952.
Von Studnitz, R. E., & Green, D. W. (2002). The cost of switching language in a semantic categorization task. *Bilingualism, 5*, 241–251.
Weber-Fox, C., Davis, L. J., & Cuadrado, E. (2003). Event-related brain potential markers of high-language proficiency in adults. *Brain and Language, 85*, 231–244.
Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*, 231–256.
Weber-Fox, C. M., & Neville, H. J. (2001). Sensitive periods differentiate processing of open- and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language, and Hearing Research, 44*, 1338–1353.
White, E. J., Genesee, F., Drury, J. E., & Steinhauer, K. (2007). Before and after: An ERP investigation of late second language learning in an intensive language course. *Journal of Cognitive Neuroscience, Supplement,* 290.
Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience, 30*(22), 7646–7651.
Yudes, C., Macizo, P., & Bajo, M. T. (2010). Cognate effects in bilingual language comprehension tasks. *Neuroreport, 21*, 507–512.

Chapter 12
MRI Methods in Bilingual Reading Comprehension

Angélique M. Blackburn

Abstract This chapter reviews developments in bilingual reading comprehension from anatomical *magnetic resonance imaging* (MRI), *functional magnetic resonance imaging* (fMRI), *functional connectivity analysis*, and *diffusion tensor imaging* (DTI). Networks involved in orthographic, semantic, and syntactic levels of processing are discussed, considering language-specific processing requirements and the role of age of acquisition and proficiency. Reading in the first (L1) and second language (L2) is subserved by largely overlapping networks, with additional recruitment for language-specific aspects of processing. Semantic access is largely shared, while syntactic processing in particular is affected by individual differences in language acquisition. L2 acquisition builds on the existing L1 system and, as L2 proficiency improves, processing becomes more native-like, substantiating the *convergence hypothesis*. Increased activity in a cognitive control network during effortful L2 reading provides support for an anatomically and functionally distinct task system during comprehension, as predicted by the *Bilingual Interactive Activation Plus Model* (BIA+). Converging evidence from fMRI and other techniques is considered to illustrate the power of combining methodologies.

Introduction

Before the advent of neuroimaging techniques, our knowledge of reading comprehension was restricted to what we could learn from behavioral and electrophysiological measures. Scientists designed clever behavioral manipulations to isolate a set of cognitive processes involved in reading and creating models of comprehension. While these types of designs have tremendous value and in many cases remain

A.M. Blackburn, Ph.D. (✉)
BSE 2.206, Department of Biology, Main Campus BSE 1.600, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249-1644, USA
e-mail: txd536@my.utsa.edu

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1_12

the best way to identify the timing of such processes, without a means to see inside the living brain, we were quite limited in our knowledge of where in the brain such processes occur. It was difficult to address questions regarding whether the first (L1) and second language (L2) share a location in the brain, whether separate regions of the brain are responsible for semantic retrieval and syntax, if there is a brain region that controls reading in two languages, and whether parallel processes occur in the same region or in vastly distributed regions of the brain.

In an attempt to answer these types of questions, the brains of patients with various language disorders were autopsied *postmortem* to identify which areas had undergone structural damage that might account for the unusual language behavior. Dissociation between reading disorders acquired following brain injury, such as phonological and surface dyslexia, shaped our theories of reading comprehension. Patients with phonological dyslexia could comprehend written words that they had previously encountered, but showed difficulty sounding-out novel words (Dérouesné & Beauvois, 1979). These patients appeared to rely exclusively on direct associations in memory from the word form to meaning, or a lexico-semantic route of processing. In contrast, patients with surface dyslexia could process words that sound like they are spelled, but experienced difficulty comprehending words with irregular pronunciation (e.g., the silent "t" in listen) and homophones (words that sound the same such as *pear* vs. *pair*; Marshall & Newcombe, 1973). This suggests these patients were unable to rely on a direct lexico-semantic route, but rather mapped written properties of the words (orthography) to the way the words sound (phonology; Coltheart, Masterson, Byng, Prior, & Riddoch, 1983). From these lesion deficit studies, a *Dual-Route Model* of comprehension was proposed. According to the model, words can be processed using a direct lexico-semantic route with direct links between word form and meaning, or a grapho-phonological route in which visual words are processed via a phonological association (see Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001 for a review). A great debate evolved regarding this Dual-Route theory and potential locations where visual word forms could be stored, but this debate could not be resolved by lesion deficit studies alone.

Lesion studies have provided a framework for existing theories of reading comprehension and have modeled our understanding of language systems. However, studies that rely on brain injury are limited by the unpredictable, rare, and imprecise nature of brain damage and the brain's tremendous ability to reorganize and compensate for damage. Because cognitive functions often rely on a network of areas, these types of studies cannot tell us which brain regions are involved or are sufficient for a given function, but only if the damaged area is *necessary* for a given function. The advent of relatively noninvasive neuroimaging techniques enabled us to view structural and functional changes in normal living participants. A wide variety of complementary techniques are available, including *computed tomography* (CT), *magnetic resonance imaging* (MRI), *positron-emission tomography* (PET), *transcranial magnetic stimulation* (TMS), *electroencephalography* (EEG), *event-related potentials* (ERPs), and *magnetic-encephalography* (MEG). Together, these methods help us to determine the time course (ERP/MEG) and brain areas necessary (TMS) and involved (fMRI/MEG) in a given function. Because many existing studies of neural networks involved in reading comprehension use MRI, this chapter focuses on MRI techniques, including structural MRI, functional MRI (fMRI), and diffusion tensor imaging (DTI).

Structural MRI

Structural MRI, or anatomical MRI, contrasts different types of tissue in the brain (for an in depth tutorial on the physics underlying MRI/fMRI, see Mandeville & Rosen, 2002). When participants are placed in an MRI scanner, a strong magnetic field within the scanner aligns the protons within water in the tissue. A radiofrequency pulse is then applied perpendicular to the magnetic field, which excites the atoms and knocks the protons out of alignment with the scanner's magnetic field. Over time, this energy is lost and the protons relax, that is, they return to their original alignment with the scanner's magnetic field. The time that it takes for the protons to realign with the magnetic field of the scanner is known as *T1 decay* and differs in grey and white matter. This is because white matter is highly myelinated and contains lipids, and grey matter contains more water. The relaxation period of myelin (lipids) is faster than that of water. We can identify large structures in the brain by imaging the contrast in T1 decay between different types of tissue. This is known as a *T1-weighted image*, in which lipids and white matter are typically brighter than water.

We can also obtain transverse relaxation time (*T2* and *T2**) weighted images. Following a radiofrequency pulse, the individual protons first spin together, then rapidly lose coherence as they collide with each other. As coherence is lost, the spins go out of phase and cancel each other out. This is known as *T2 decay*, which is faster for molecules with greater magnetic susceptibility. Unlike T1-weighted images, T2-weighted images are brighter for fluid than for fat. T2* decay is caused by magnetized nuclei losing magnetic coherence from colliding with each other (T2), but is also susceptible to inhomogeneities in the magnetic field. Transverse decay is used in fMRI, as will be described in the next section.

While structural MRI does not allow us to determine what areas of the brain are involved in specific processes (e.g., reading comprehension), this method has brought to light differences in the monolingual and bilingual brain (see Fig. 12.1). Structural MRI has revealed that bilinguals have a greater proportion of grey matter in the left inferior parietal lobe (IPL) than monolinguals (Mechelli et al., 2004). The IPL has been implicated in word recognition. As can be seen in Fig. 12.1, the increase in grey matter is greater in highly proficient bilinguals who acquired two languages early in life than in lower proficiency bilinguals who acquired an L2 at a later age. A similar anatomical technique known as computed tomography has been used to demonstrate differences in brain atrophy of bilingual and monolingual patients exhibiting similar symptoms of Alzheimer's disease. Despite similar cognitive performance, bilingual patients at the same disease stage exhibit a greater

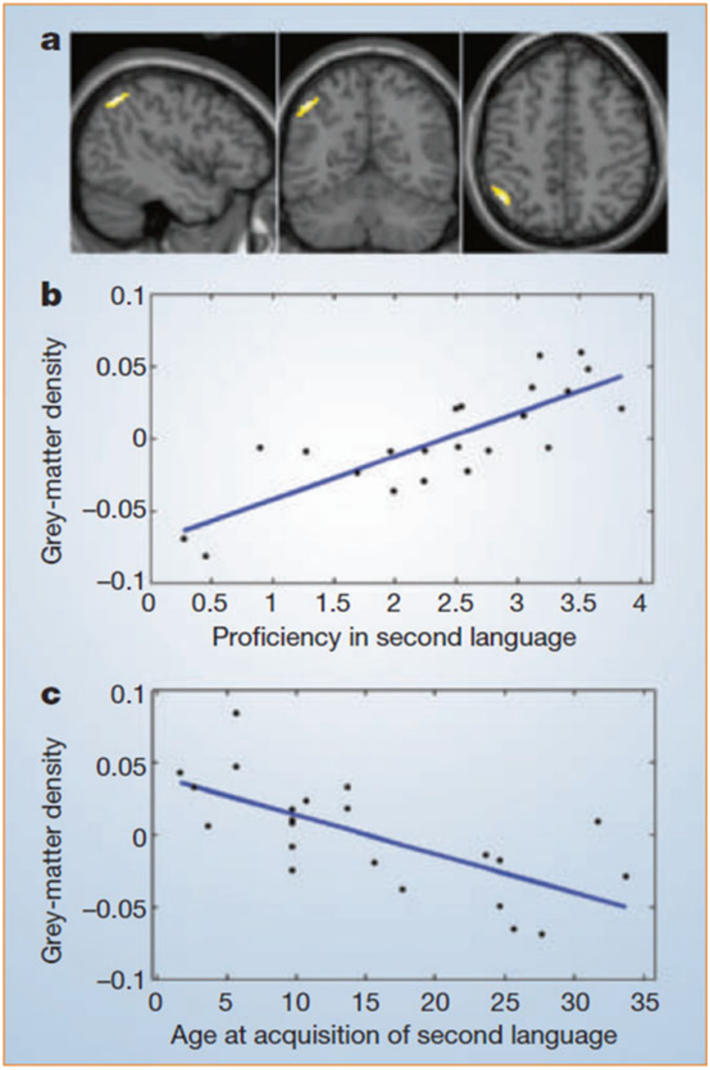


Fig. 12.1 T1-weighted MRI sagittal, coronal, and axial images highlighting the inferior parietal lobe, which has greater grey matter density in bilinguals than in monolinguals (**a**). Grey matter density in this region is positively correlated with proficiency (**b**) and negatively correlated with age of acquisition (**c**). Adapted from "Neurolinguistics: Structural Plasticity in the Bilingual Brain," by A. Mechelli, J.T. Crinion, U. Noppeney, J. O'Doherty, J. Ashburner, R.S. Frackowiak, and C.J. Price, 2004, *Nature*, *431,* pp. 757–757. Copyright 2004 by Macmillan Publishers Ltd

degree of atrophy than monolinguals, indicating that the bilinguals are better able to compensate for symptoms of the disease (Schweizer, Ware, Fischer, Craik, & Bialystok, 2012). These studies demonstrate how structural brain imaging can be used to reveal brain areas affected by bilingualism.

Structural MRI is limited in what we can learn about processing, but we often combine structural MRI with other techniques to answer questions about the function of a particular region. Functional MRI images are generally superimposed on structural images so that the areas of the brain involved in specific functions can be identified. Structural MRI is useful in guiding region of interest (ROI) analyses and can be used in conjunction with TMS and ERPs, as will be described below.

Functional MRI

Functional MRI is the most popular, and arguably the most useful, technique for investigating the brain regions involved in reading comprehension. This is because fMRI provides a spatially precise measure of changes in the blood oxygen level in the brain that occur as a result of neuronal activity. As populations of neurons fire, oxygen in the region is consumed. This oxygen is rapidly replenished via the bloodstream (beginning at about 2 s and peaking about 6 s after consumption). We can measure changes in the *blood-oxygen level dependent* (BOLD) signal to determine which areas have just experienced an increase in neuronal activity—thus, this is an indirect measure of brain activity.

As mentioned above, we collect a T2*-weighted image in BOLD fMRI, which reflects differences in magnetic susceptibility across the brain. The magnetic susceptibility of blood depends on the amount of oxygen present. In oxygenated blood, hemoglobin contains oxygen (oxyhemoglobin), which shields the magnetically susceptible iron ion. Oxyhemoglobin, like most molecules in the body, is diamagnetic, meaning that all of the electrons are spin-paired and they are largely unaffected by magnetic fields. Once oxygen is used metabolically by the neurons, the remaining deoxyhemoglobin is paramagnetic, meaning there are unpaired electrons that are susceptible to magnetic fields. When the radiofrequency pulse is administered, protons in deoxyhemoglobin rapidly lose coherence due to the magnetic field, which accelerates the transverse relaxation decay and decreases the signal. In contrast, oxyhemoglobin has less magnetic susceptibility, and less signal attenuation is observed in areas with oxygenated blood. Thus, the BOLD magnetic resonance signal, or hemodynamic response, indicates the *relative* changes in oxygenation across the brain as oxygen is consumed and replenished.

A characteristic BOLD response is observed when neurons in the same region are active (see Fig. 12.2). When a participant performs a cognitive task, neurons fire and use oxygen. We can sometimes observe a small initial dip in the signal due to the paramagnetic properties of deoxyhemoglobin. This loss of oxygen is followed by an influx of more oxygenated blood than is necessary, resulting in an overshoot in the signal. Finally, as the extra oxygenated blood is removed from the bloodstream, the signal first undershoots (as a result of temporary venous ballooning), then returns to baseline (Chen & Pike, 2009). We can model this characteristic hemodynamic response for each stimulus, assuming that overlapping BOLD responses summate linearly. We then perform statistical tests between the observed

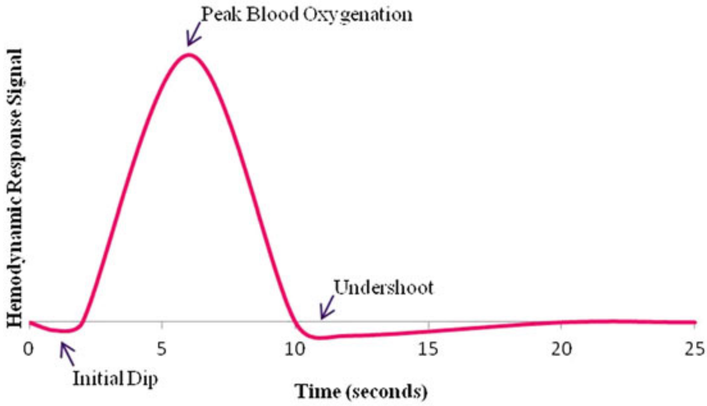


Fig. 12.2 The canonical modeled hemodynamic response function. The slow BOLD signal is characterized by an initial dip as oxygen is consumed, influx of oxygen, and an undershoot before returning to baseline

signal and the expected signal to determine which brain areas are "active" (i.e., show an increase in BOLD signal). Statistical analyses are performed in small three-dimensional sections, or voxels of the brain. We then overlay this statistical map of colored voxels with activation above a set threshold (i.e., *p*-value) onto an anatomical MRI image to visualize the regions that are "activated" for a given task. It should be noted that external contrast agents could be used to measure metabolic activity instead of hemoglobin. However, as the BOLD signal is the most commonly used in reading comprehension, we will focus on this method.

fMRI Designs

The majority of early fMRI studies of reading comprehension involved block designs. In these designs, each block of trials contains one experimental condition. Experimental blocks in which participants perform a given task (e.g., read words in Spanish) are compared to resting blocks or blocks with a different experimental condition. Block designs have high statistical power for identifying activated regions, but are not optimal for sentence reading studies in which we might want to interpolate trials of different conditions. In recent years, event-related designs have come to dominate the field. These designs have the advantage that stimuli in different conditions can be intermixed, and the responses are recorded individually over time. These designs require that the interstimulus interval (ISI) varies from trial-to-trial and that this variance is modeled in the hemodynamic response function. In fact, our understanding of the additive properties of the hemodynamic response during reading has enabled us to develop MRI sentence reading paradigms in which we can rapidly present word stimuli at a normal reading rate, every 200–300 ms

(Yarkoni, Speer, Balota, McAvoy, & Zacks, 2008). Regardless of using a blocked or event-related design, implementing an appropriate experimental design is critical. Reading comprehension studies typically have employed subtraction methodology or parametric manipulation, so we will discuss these further.

Subtraction Methodology, Parametric Contrasts, and Meta-Analyses

Any region of the brain that experiences an increase in blood-oxygen level will exhibit an increase in the hemodynamic response and be considered "active." This is problematic because much of the brain is active at any given moment. For instance, if we instruct participants to press buttons as they read words in a sentence, we might expect activation not only in regions of the brain involved in comprehension, but also in visual and motor cortices involved in sensory perception and performing the task. For this reason, we often use a subtraction technique between an experimental and baseline condition to isolate the regions involved specifically in that task. It is critical for the researchers to carefully select baseline conditions that "subtract away" any activity that is not what they intend to measure.

In early studies of reading comprehension, it was common to have participants rest or fixate on a visual marker for baseline conditions. For instance, in early fMRI studies of word reading, a baseline fixation condition was often subtracted from a condition in which words were presented on the screen (Chee, Tan, & Thiel, 1999; Cohen et al., 2000; Dehaene et al., 2001). In later attempts to control for visual processing unrelated to reading, words were contrasted with pseudowords (Fiebach, Friederici, Muller, & von Cramon, 2002) or symbol strings (Kiehl et al., 1999). Some PET studies, which are methodologically similar to fMRI, have contrasted words with letter strings (Menard, Kosslyn, Thompson, Alpert, & Rauch, 1996), lines (Bookheimer, Zeffiro, Blaxton, Gaillard, & Theodore, 1995), or with "falsefonts" (Brunswick, McCrory, Price, Frith, & Frith, 1999). Falsefonts are meaningless character combinations that are matched to the Roman alphabet in visual complexity. Words in two orthographically similar (Jamal, Piche, Napoliello, Perfetti, & Eden, 2012) and dissimilar languages (Nelson, Liu, Fiez, & Perfetti, 2009) have been contrasted to isolate regions involved in orthographical discrimination and different processing routes (e.g., direct word form-to-meaning vs. word form-phonology-meaning). Finally, fixations and sentence-like consonant strings have been subtracted from sentences, and simple sentences have been subtracted from sentences with grammatical complexity or violations to isolate regions involved in sentence-level processes (e.g., Rüschemeyer, Zysset, & Friederici, 2006). In many cases, a serial subtraction design is employed in which multiple conditions with increasing cognitive complexity are contrasted. Each experimental condition becomes the baseline for another condition with one additional cognitive process. For instance, we might have bilinguals fixate on a cross, read pseudowords, and read words. We can subtract fixation from pseudowords to isolate orthographic processing and subtract pseudowords from words to isolate semantic processing (Petersen, Fox, Snyder, & Raichle, 1990).

Subtraction methodology requires that the baseline and experimental conditions differ only in the processes under study. In reality, this is very difficult because the brain is sensitive to specific features of stimuli. To control for stimulus characteristics, researchers sometimes vary the task demands to the same stimuli. For instance, to isolate regions involved in phonological processing, participants may be instructed to make either a phonological judgment such as whether the word contains the /b/ sound or a non-phonological judgment such as whether the font is large or small. The same words can be presented in both conditions, and activation due to visual characteristics will be subtracted away. Unfortunately, even when the same stimuli are used and the task differs (e.g., press a button for English vs. Spanish words), the demands of the task may elicit activation in areas unrelated to the process we are attempting to isolate (e.g., greater activation in language control networks for the less proficient language).

To avoid confounds due to different tasks, parametric manipulations are often employed (e.g., Büchel, Holmes, Rees, & Friston, 1998). In this case, the stimuli and the task remain the same or similar, but the difficulty of performing the task increases. As task difficulty increases, processing demands and resulting BOLD signal in regions of the brain involved in the task are assumed to increase. We can use a general linear model to correlate task difficulty or other experimental parameters with the relative increase in the hemodynamic response and reveal where neural responses are explained by the experimental manipulation. For example, Cohen, Dehaene, Vinckier, Jobert, and Montavont (2008) parametrically manipulated reading difficulty by increasing the level of degradation in visually presented words. They examined the correlation between the degree of stimulus degradation and activation in specific regions of interest. As degradation level increased, they observed an increase in activation in posterior parietal cortical regions associated with visual attention. The authors suggested that the ventral reading pathway is used in cases where familiar words are easily recognized, but a dorsal pathway involving posterior parietal cortex is required to read words that are not easily recognized. A common parametric design to assess higher level processing at the sentence level is to present the same sentences with different levels of complexity or violations that disrupt a specific aspect of processing (e.g., semantics or syntax). It is also common to use the rate of stimulus presentation or response times as predictors of task difficulty in the general linear model (Binder et al., 1994; Price et al., 1992).

Finally, it is common to use the same experimental and baseline conditions in two participant groups that theoretically differ in only one dimension (e.g., bilinguals vs. monolinguals). In this way, we can determine how reading may differ between bilinguals and monolinguals, across the lifespan (children vs. adults), with differing proficiency in the L2, with age of second language acquisition, and in normal adults vs. those with reading deficits. This can be done either as a subtraction (e.g., bilingual minus monolingual) or as a parametric manipulation (e.g., using proficiency scores). It is important to select the design that is most appropriate for each study. As with any other subtraction design, comparing two groups will be affected by any differences between the groups, such as age, socioeconomic status, and working memory ability. Using a between-subjects design is especially complicated in fMRI research, as we are performing statistical tests on very small sections of the brain that do not share exact spatial locations across individuals. When it is useful to use subtraction to compare groups, such as bilinguals and monolinguals, participants should be carefully matched on potentially confounding factors, an adequate sample size should be used, and implications of using a between-subjects design should be considered when interpreting the results.

The structural brain images and overlaid statistical maps obtained for each participant are usually averaged and fit to a standard stereotactic brain space with coordinates that are comparable across studies. A number of convertible atlases can be used. Typically, brains are mapped to either Talairach space, a coordinate system based on one brain, or Montreal Neurological Institute (MNI) space, an atlas created from an average of hundreds of brains. Multiple atlases have been created; fortunately, it is usually possible to convert between these systems. Thus, it is possible to compare studies attempting to isolate the same process using slightly different tasks or comparisons. This is a useful tool because of the fMRI design limitations mentioned above and the added difficulty that many studies are limited in sample size due to the expense of running fMRI studies.

Recall that any brain area that is involved differentially in the experimental and control conditions will "light up" on the statistical map. Therefore, the regions activated often vary in studies with slight differences in the task demands, control conditions, stimuli, and subjects. Meta-analyses compare a number of studies to identify common areas involved in a cognitive process despite slight differences in experimental design. A useful database for meta-analysis, BrainMap, has been implemented in which patterns of activation can be searched by task (e.g., read, detect), modality (e.g., auditory, visual), stimulus and response types (e.g., button press), subject population (e.g., native language), context (e.g., age effects, learning, linguistic effects), experimental contrast (e.g., based on group, stimulus type), regions of interest, and other variables of interest (brainmap.org; e.g., Laird, Lancaster, & Fox, 2005; Turkeltaub et al., 2012). Meta-analyses of studies using fMRI and other methodologies have led researchers to answer a number of questions about word and sentence-level processing and identify a brain network involved in reading comprehension. These questions are discussed in the next section.

fMRI and Models of Reading Comprehension

Functional MRI is an excellent technique to answer questions regarding how information is represented in the brain and accessed during reading. We can use this method to investigate the neural overlap of the L1 and L2, whether word reading occurs via a direct (word form to semantics) or indirect route (word form to phonology to semantics), whether word form and semantic representations are stored in a central location in the brain, and whether there is a region of the brain responsible for

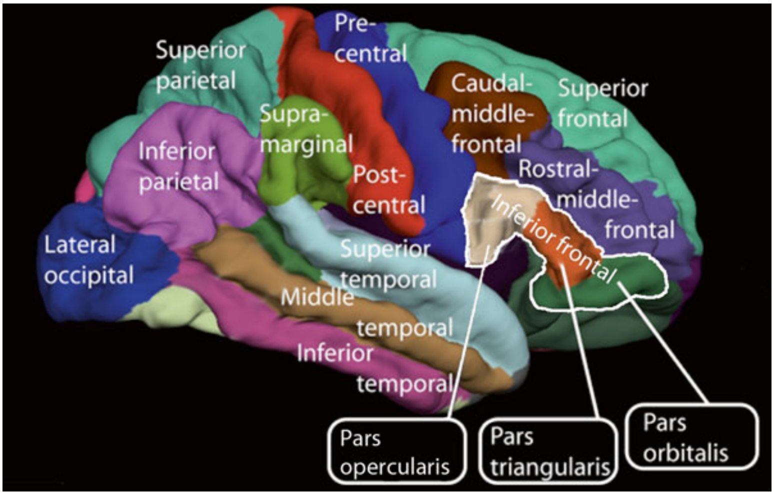


Fig. 12.3 Right lateral surface of the brain. Note that most language networks are left-lateralized but are typically highly distributed. Adapted from "Mapping the Structural Core of Human Cerebral Cortex," by P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Wedeen, and O. Sporns, 2008, *PLoS Biolology, 6,* e159. Copyright 2008 by Hagmann et al.

orthographic, word level, semantic, and syntactic processes. We can also use this method to determine if the pathways involved in reading comprehension differ across the lifespan, with changes in proficiency or performance ability, and in individuals with reading deficits. Finally, we can determine if the language network differs for bilinguals and monolinguals and if bilingual reading involves additional task demands than reading in one language alone. Here, we will address the large body of fMRI research contributing to some of these questions. It would be convenient to refer to Fig. 12.3 of the brain to understand the flow of information during comprehension.

Processing Routes for Reading

Functional MRI is perhaps best used to test existing theories that are based on lesion deficit, behavioral, and electrophysiological research. One question predating the advent of fMRI is whether there are two processing routes for word comprehension: a direct route to meaning or an indirect route through phonology. As discussed in the introduction, the Dual-Route Model originated from dissociation between two forms of dyslexia, phonological and surface dyslexia, and has implications in clinical research showing that children with dyslexia often are impaired in mapping of orthography to phonology (i.e. the core phonological deficit hypothesis of reading disability; Swan & Goswami, 1997). According to the Dual-Route Model, words can be processed using a lexico-semantic route with direct links between word form and meaning, or a grapho-phonological route in which visual words are processed via an indirect phonological association (Coltheart et al., 2001).

Consistent with the Dual-Route Model, a meta-analysis of fMRI word reading studies revealed different patterns of activation in studies contrasting tasks to isolate direct semantic retrieval (e.g., words vs. pseudowords) and indirect retrieval through phonology (e.g., reading aloud vs. silent reading; Jobard, Crivello, & Tzourio-Mazoyer, 2003). These patterns of activation implicated two functionally distinct processing routes: a dorsal path for semantic retrieval through grapho-phonological mapping and a ventral path for direct lexico-semantic retrieval (see Fig. 12.4a). Both of these left-lateralized routes originate in the visual cortex and project anteriorly. The dorsal temporo-parietal pathway, including the supramarginal gyrus in the IPL, angular gyrus, the pars opercularis of Broca's area, and superior temporal cortex, has been implicated in mapping orthography to phonology during reading and serves as the indirect phonological processing route (Fiebach et al., 2002; Jobard et al., 2003; Pugh et al., 2001). This pathway is activated when participants are required to map orthography to phonology, as during a rhyming decision on written word forms (Booth et al., 2002). On the other hand, the ventral occipito-temporal, lexico-semantic route involves the occipito-temporal junction, the left inferior temporal cortex, the posterior middle temporal gyrus (MTG), and the pars triangularis of Broca's area (Fiebach et al., 2002; Friederici & Gierhan, 2013; Jobard et al., 2003). Converging evidence from fiber tracking, a technique which will be discussed below, has shown that these two paths are not only functionally distinct, but also structurally distinct (see Fig. 12.4b, c). These processing routes have been well established using fMRI in both monolingual and bilingual readers. Next, we consider how processing may be affected by bilingualism.

L1 and L2 Processing and Word Representation Overlap

Kim, Relkin, Lee, and Hirsch (1997) conducted a landmark fMRI study to investigate the potential overlap between L1 and L2 and how second language age of acquisition (AoA) influences the degree of overlap. Age of acquisition refers to the age at which a language is learned, which may be *early*, that is before or around the age of puberty, or *late*. Largely overlapping regions of activation were observed in early learners of both languages, while late L2 learners showed more distinct regions of activation for each language. Initially, these findings were taken as evidence that late learners build representations for words in L2 around the existing L1.

However, the majority of studies show a great degree of overlap between L1 and L2, often with slightly increased activation or additional recruitment of other areas for one of the languages. One possibility for differences in neural overlap between L1 and L2 across studies is that, rather than reflecting separate areas of L1 and L2 representation, differences observed reflect specific properties of the stimulus language. Languages that share similar orthography and syntax, such as Spanish and Italian, may have a greater degree of neural overlap than languages that rely on

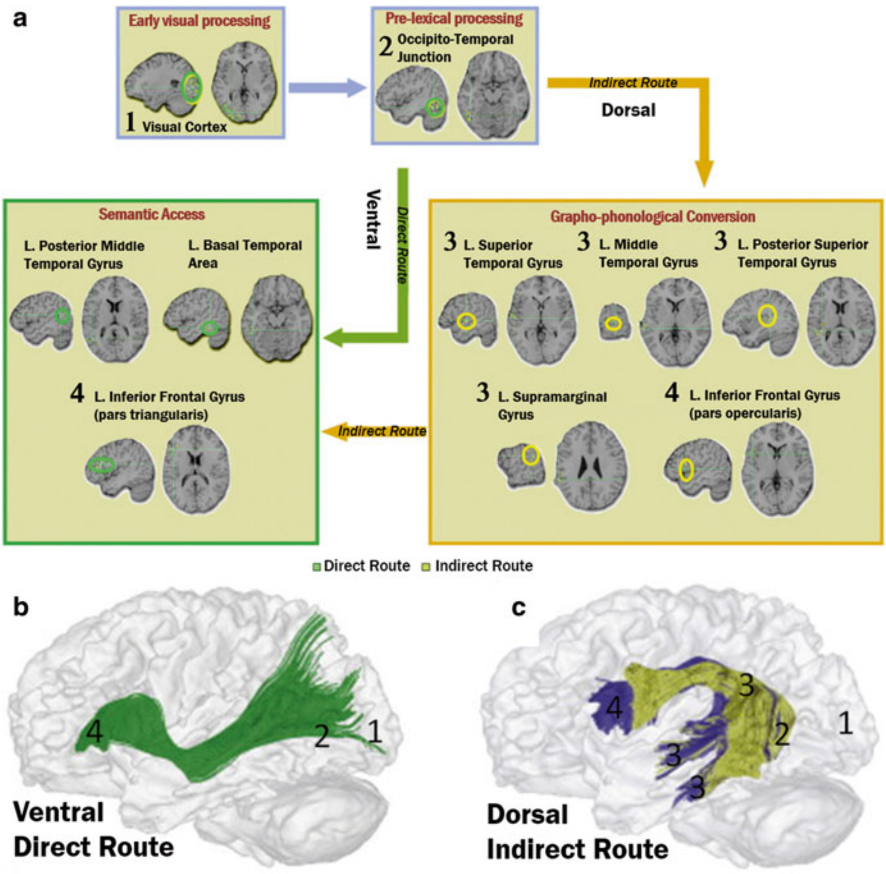


Fig. 12.4 Converging evidence from fMRI meta-analysis (**a**) and fiber tracking (**b**), (**c**) of direct and indirect reading pathways. Areas activated in fMRI studies that isolated processes involved in direct semantic access are in green; areas activated during indirect access are in yellow (**a**). The direct ventral path connects visual processing regions to the anterior part of the inferior frontal gyrus (**b**). The indirect dorsal path includes the supramarginal gyrus and terminates in the posterior inferior frontal gyrus (**c**). 1 = visual cortex, 2 = occipito-temporal junction, 3 = areas involved in grapho-phonological conversion for semantic retrieval, 4 = inferior frontal gyrus. Figure A adapted from "Evaluation of the Dual Route Theory of Reading: A Meta-Analysis of 35 Neuroimaging Studies," by G. Jobard, F. Crivello, and N. Tzourio-Mazoyer, 2003, *NeuroImage, 20,* pp. 693–712. Copyright 2003 by Elsevier. Figure B, C Adapted from "Dorsal and Ventral Pathways in Language Development." by J. Brauer, A. Anwander, D. Perani, and A.D. Friederici, 2013, *Brain and Language, 127,* pp. 289–295. Copyright 2013 by Elsevier

different syntactic strategies or orthographic processing, such as English and Chinese. Another possibility is that differences arise because bilinguals are a widely heterogeneous group with different AoAs, languages acquired, and proficiency attained. These factors were not independently considered in the landmark study by Kim et al. (1997). This led to a debate on the role of AoA and proficiency on the neural organization of an L2. Subsequent studies have shown that these factors may differentially impact organization of early and later stages of processing such as orthography and syntax. Researchers typically have pursued the question of neural organization at one or more levels of processing: orthographic (written form), phonological (sound), semantic (meaning), syntactic (grammar), and discourse level (story or text coherence across sentences).

A general left-lateralized, posterior-to-anterior reading network has been identified which involves orthographic processing in visual regions (occipital areas and the fusiform gyrus), phonological mapping in temporal and parietal regions (superior temporal gyrus [STG], parietal cortex, and inferior frontal cortex), semantic retrieval in anterior and distributed areas (inferior frontal gyrus [IFG], MTG, STG, and angular gyrus), syntactic analysis in posterior temporal and anterior areas (e.g., the arcuate fasciculus pathway connecting the posterior STG and the posterior part of Broca's area), and discourse level processing in extended bilateral networks (refer to Fig. 12.3; Brunswick, 2010; Buchweitz, Mason, Hasegawa, & Just, 2009; Friederici & Gierhan, 2013). ERP and fMRI studies in monolinguals have shown that different levels of linguistic complexity (e.g., orthography, semantics, syntax) occur over different (but sometimes parallel or cascading) time courses and differentially affect brain regions involved in reading comprehension. Many areas are involved in more than one linguistic process and may be involved in the integration of these processes (Keller, Carpenter, & Just, 2001). Of interest is whether patterns of activation for each linguistic process are modulated by language-specific properties or individual differences in the bilinguals' AoA and proficiency in each language. It is important to note that, although we focus exclusively on comprehension in this chapter, much of the comprehension networks described below are also engaged during production (Humphreys & Gennari, 2014).

Bilingual Orthographic to Phonological Processing

A substantial number of fMRI studies have investigated bilingual orthographic processing at the single-word level. Single-word comparisons are ideal for fMRI because we can experimentally isolate orthographic processes more easily than sentence-level processes, both in terms of choosing an adequate control condition and in terms of the slow hemodynamic response. Unlike sentence presentation, which is unnatural when slowed and requires complex hemodynamic response modeling to isolate rapid word-level effects, we can present single words at a sufficiently slow rate so as to capture the slow BOLD response or use a blocked design to present multiple words in one condition within a short block. At the orthographic level, we are interested in whether there are language-specific patterns of activation, whether orthographic processing is mediated by mapping to phonology, and if L1 orthographic processing strategies transfer to L2 processing. In general, we find that the orthographic demands of a language have been shown to affect processing networks and there is some degree of applying L1 orthographic strategies to L2 when possible.

One way to investigate language-specificity is to test whether the specific orthographic properties of a language are more likely to evoke indirect grapho-phonetic or direct grapheme to lexico-semantic routes. While some languages have fairly consistent grapheme to phoneme mappings (orthographical transparency), others have multiple mappings between graphemes and phonemes (opacity). For instance, English is an orthographically opaque language in which the same graphemes can be pronounced differently (e.g., *pint-mint* and *cough-through-thorough*), whereas Italian is more transparent. A common fronto-temporal network for reading in Italian and English has been identified using both PET and fMRI (Meschyan & Hernandez, 2006; Paulesu et al., 2000). In addition to a common neural system, orthographic opacity led to greater activation in areas related to semantic retrieval (e.g., the left posterior inferior temporal cortex and the anterior IFG), while orthographic transparency led to recruitment of phonological areas (i.e., the left STG). This suggests that orthographic transparency allows mapping to phonology through the STG while words in opaque languages require direct semantic retrieval (Paulesu et al., 2000). The influence of orthographic transparency on recruitment of phonological paths to meaning has been replicated in other alphabetic languages such as Spanish and English (Jamal et al., 2012).

A similar comparison has been made between two Japanese orthographies: Kana and Kanji. Kana is a transparent syllabic orthography with direct mapping to phonology and is thought to engage a phonological processing route. On the other hand, Kanji is a logographic orthography in which single complex characters map onto morphemes rather than phonemes. That is, the same Kanji graphemes can be pronounced differently in different words, and are therefore more likely to incur direct lexico-semantic access. The comparison of Kana and Kanji provides the unique opportunity to assess orthographically mediated neural differences in the same subjects without confounding proficiency across languages.

A number of PET and fMRI studies have shown that, as with alphabetic orthographies, a common neural system involving the left occipito-temporal cortex is engaged for reading both Kanji and Kana (Ino, Nakai, Azuma, Kimura, & Fukuyama, 2009; Nakamura, Dehaene, Jobert, Bihan, & Kouider, 2005; Nakamura, Dehaene, Jobert, Le Bihan, & Kouider, 2007). Areas specific to orthographic attributes were also identified. Kana typically involves visual to phonological encoding regions, including the lateral occipital and left inferior parietal areas (Nakamura et al., 2005; Nakamura, Dehaene et al., 2005; Sakurai et al., 2000). On the other hand, Kanji involves greater activity in the fusiform gyrus, a region thought to involve direct access to specific lexical forms. Kanji also has been found to involve the right hemisphere more than Kana and alphabetic languages (e.g., English). This is consistent with the interpretation that the right hemisphere is involved in global processing, as is necessary for logographic scripts more than for alphabetic or syllabic scripts (Nakamura, Oga et al., 2005).

Like Kanji, logographic languages such as Chinese recruit neural areas in addition to those engaged for alphabetic languages with direct grapho-phonetic mappings. A number of studies comparing English and Chinese have found that in addition to typical patterns of activation for reading in alphabetic languages, reading in logographic languages involves the left middle frontal gyrus (MFG), right prefrontal, and parietal areas, which may handle the visual-spatial complexity of the characters (Tan et al., 2001). In addition, reading in Chinese results in more bilateral activation than the typical left-lateralized network recruited during English reading (Tan, Spinks, Eden, Perfetti, & Siok, 2005). It has been proposed that right hemisphere recruitment in Kanji and Chinese contributes to the additional spatial representations, spatial working memory demands, and global processing of logographic vs. alphabetic languages (Tan et al., 2003). This is especially the case when the task demands require mapping to phonology, as with tasks in which participants must decide if two written words rhyme. These studies indicate that a general network is involved in orthographic and phonological processing, with additional recruitment of areas that are specific to the attributes of each language. In particular, logographic languages that cannot be subserved by phonological mapping require additional areas involved in spatial memory and global processing.

Of interest is whether learning a second orthographic system recruits new neural substrates specific to that language, or if the learned system engages existing L1 processing networks. Three studies have taken advantage of the language-specific neural substrates identified in English and Chinese to determine if learners use existing native-language networks during L2 reading or if they recruit new areas specific to the learned language. Nelson et al. (2009) showed that, unlike the left-lateralized activation found in native English speakers, Chinese learners of English show a bilateral pattern of activation in the fusiform gyrus and additional recruitment of the left MFG during both Chinese and English reading. This indicates that when possible, these learners apply strategies and networks from their native language to the learned language. On the other hand, English learners of Chinese showed the native Chinese pattern: bilateral fusiform and left MFG activation when reading Chinese. Likewise, Liu and colleagues (Liu, Dunlap, Fiez, & Perfetti, 2007) observed bilateral fusiform and left MFG activation when English speakers with minimal instruction in Chinese read Chinese words. Thus, when these learners were unable to adapt their native language networks to the spatial demands of the new language, they recruited additional regions for language-specific orthographic processing. Finally, Tan et al. (2003) assessed phonological processing of Chinese and English during a rhyming decision task. Participants saw word pairs in either language and were instructed to indicate if the words rhymed. Like in the studies by Nelson et al. and Liu et al., Chinese learners of English showed similar patterns of activation in the inferior and middle frontal cortices for English and Chinese, but a different pattern for native English speakers who did not know Chinese. The authors of all three studies concluded that language-specific attributes of orthographic and phonological processing in L1 can shape processing in L2, except when the existing networks in L1 are inadequate for the language-specific demands of the new language. The same conclusion was reached using Japanese–English bilinguals who show L2 reliance on L1 networks despite language-specific processing in native speakers of these languages (Nakada, Fujii, & Kwee, 2001).

The myriad of single-word bilingual studies have shown that while a common language system is involved in orthographic and phonological processing, language-specific attributes modulate areas involved in visual encoding, direct lexico-semantic access, and indirect access to semantics through phonology.

Bilingual Semantic Access

A few studies have investigated whether meaning is differentially accessed when reading in two different languages. A large distributed left-lateralized semantic network including IFG (especially the posterior aspect), STG, MTG, and in some cases the parietal lobe, is activated similarly for both of a bilingual's languages. To isolate semantic processes at the single-word level, studies typically contrast tasks that require meaning to be accessed (e.g., semantic decision tasks) with tasks that do not involve meaning (e.g., a font size decision task). At the sentence level, meaningful sentences are often compared to sentences composed of non-words, pseudowords, or scripts that are either not real (falsefonts) or not understood by the subjects. In some cases, sentences containing semantic violations are contrasted with plausible sentences to isolate processes involved when semantic integration is difficult.

For instance, when fluent Spanish–English bilinguals performed a semantic judgment task on written words (i.e., decide if the word represents a concrete or an abstract concept) relative to a non-semantic task (i.e., decide if the font is upper- or lower-case), a common bilateral frontal network, including left IFG, left temporal lobe, and right middle frontal gyrus, was found across both languages. Patterns of activation during the Spanish and English semantic tasks were compared to identify areas that were language specific for each individual. No consistent differences were observed during semantic retrieval in both languages. This finding provides evidence for a single conceptual representation for words in both languages (Illes et al., 1999; see also Pillai et al., 2003 for a summary of similar findings).

In a similar comparison, bilinguals were asked to determine if a target was related to a preceding prime word, a task that requires access to meaning (Mouthon, 2011). Words were presented in the L1 and L2, which varied across participants but could be French, German, or English. The bilateral temporal cortices were selected as a region of interest, due to implication of this area as a generator for the N400 ERP response and previous hypotheses that this region is involved in retrieval of concepts from memory. Briefly, the N400 is a negative-going deflection around 400 ms post-stimulus onset and is sensitive to the process of semantic retrieval. Activity in the bilateral temporal cortices was unaffected by the language of reading and proficiency in this language. However, other areas thought to be involved in early word form processes (e.g., the visual occipital regions involved in orthographic processing) were more heavily activated for the L2, especially for bilinguals with lower proficiency. Consistent with the conclusions above, the authors suggested that early word recognition processes are modulated by language but semantic processing is unaffected by language. Rather, a common or largely overlapping semantic retrieval network is engaged during reading in both languages.

In apparent contrast to these studies, different overlap between phonology and semantics was found in Spanish-English bilinguals when bilinguals were instructed to perform a semantic noun-verb association task and a phonological rhyming task (Pillai et al., 2003). In the noun-verb association task, participants were instructed to select which of two written verbs were semantically associated with a given noun.

In the rhyming task, participants decided whether a pair of written words rhymed. Comparing activation patterns in these two tasks allowed the researchers to determine if semantic and phonological processing involve the same networks and if the degree of overlap is affected by language. They found that semantic and phonological processes overlap more in L2 than L1. While this result appears contradictory to other studies showing semantic overlap between languages, the authors posit that the difference between languages may be due to difficulty in phonological processing of the L2 (see also Buchweitz et al., 2009; Rüschemeyer et al., 2006 for evidence of increased phonological processing in L2 reading).

Thus, there appear to be greatly overlapping regions involved in semantic retrieval of alphabetic languages. However, this overlap may be specific to languages with shared scripts, while languages that are dissimilar may be stored differently. It has been posited that logographic languages may facilitate semantic retrieval because the characters themselves contain semantic information that is relatively transparent to the reader (Wang, 1973). A series of studies, (Chee, Hon, Lee, & Soon, 2001; see also Chee, Tan et al., 1999; Chee et al., 2000) investigated the neural correlates of semantic processing in Chinese and English, both during single word and sentence reading. Chee and colleagues contrasted semantic judgments in which participants indicated which one of two words was closest in meaning to a third sample word (i.e., the match-to-sample task) with judgments about font size. They also contrasted sentence reading in Chinese and English with both a visual fixation and with sentences in a script unknown to the participants (Chee, Caplan, et al., 1999). They found that the same left-lateralized fronto-temporal network is involved in semantic processing for both languages at the word and sentence level and that activation is maximal over the left perisylvian cortex during single-word semantic retrieval. Although the same network was involved for both languages, they found additional activation in the bilateral inferior frontal and left prefrontal cortex for the less dominant language. This activation was ascribed to the recruitment of additional resources for processing in the language with lesser proficiency.

The conclusion that proficiency plays a role in the recruitment of neural resources has been substantiated in a number of studies contrasting semantic and non-semantic tasks. For instance, the left MFG appears to be more active during processing in highly proficient languages and is therefore thought to be involved in retrieval of learned associations between concepts in memory (Chen, Fu, Iversen, Smith, & Matthews, 2002; Ding et al., 2003). Areas involved in language and non-language cognitive control, such as the anterior cingulate (ACC), prefrontal cortex, and left IFG, have all shown greater activation for semantic processing in the less proficient language, substantiating the claim that an additional control network may be recruited to process the weaker language (see Fig. 12.5; Chee et al., 2001; Ding et al., 2003). This control network may be engaged both for controlled access to meaning and syntactic rules and/or to manage greater articulatory motor demands during reading in the less proficient language (Meschyan & Hernandez, 2006). Importantly, these control regions are activated during semantic tasks in the weaker language for early and late bilinguals, but not in late bilinguals with high proficiency levels (Ding et al., 2003; Illes et al., 1999; Indefrey, 2006; Pillai et al., 2003).

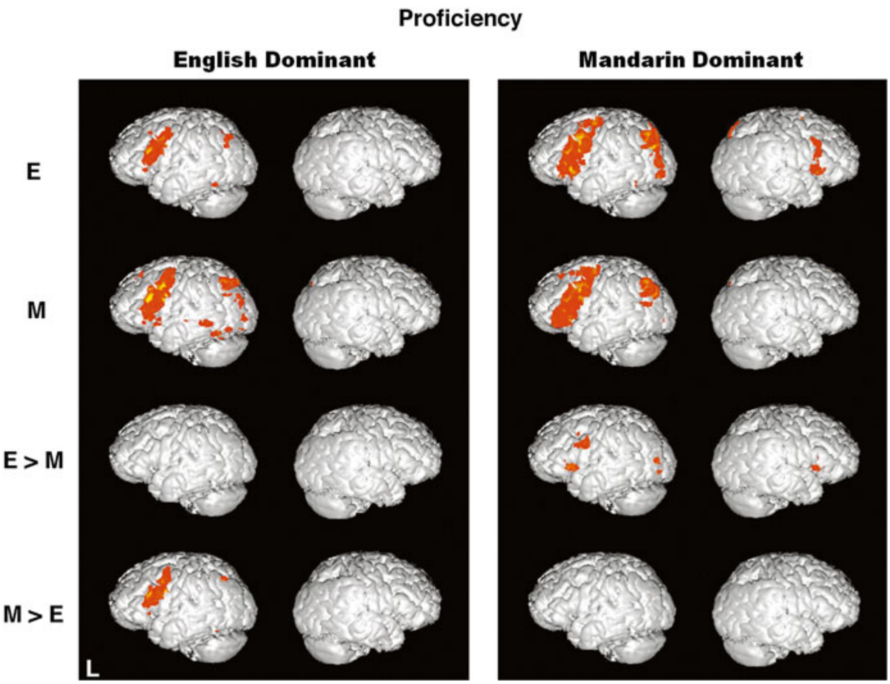


Fig. 12.5 fMRI data for native English (E) and Mandarin (M) speakers as they perform semantic judgments in E and M relative to font judgments. A largely overlapping left-lateralized semantic network is activated in both languages. Contrasting E and M for each group reveals greater activation in left prefrontal and parietal regions for the less proficient language. Adapted from "Relative Language Proficiency Modulates BOLD Signal Change when Bilinguals Perform Semantic Judgments," by M.W.L. Chee, N. Hon, H.L. Lee, Chun and S. Soon, 2001, *NeuroImage, 13,* pp. 1155–1163, Copyright 2001, by Elsevier

Thus, L2 AoA appears to play a much smaller role in semantic organization than proficiency (Abutalebi, 2008; Indefrey, 2006; Wartenburger et al., 2003).

To summarize, nearly all fMRI studies of semantic retrieval have shown largely overlapping areas for semantic access in both languages, supporting a single conceptual store for both languages. However, there are proficiency-based differences in semantic retrieval. The first language may involve direct retrieval of learned associations between concepts in memory, as evidenced by greater left MFG activity for L1 than L2. Access to meaning in the second language is more controlled, especially for bilinguals with lower proficiency. This conclusion is in line with a shared conceptual representation for words in both languages as predicted by the *Revised Hierarchical Model*, with direct access in proficient bilinguals but indirect, controlled access of L2 through the first language for low proficiency bilinguals (Kroll & Stewart, 1994). Finally, language differences found in studies that do not isolate semantic processes may reflect differences in the early processing routes to semantic access (i.e., a direct route in Chinese and an indirect grapho-phonological route in English; differences in orthographic transparency).

Bilingual Syntactic and Sentence-Level Processing

Language-Specific Syntactic Processing

A large body of literature on the neural substrates of bilingual syntactic processing exists. Studies isolating syntax often use block designs to contrast reading sentences of different levels of complexity or with syntactic violations. A general syntactic network is thought to involve pathways connecting posterior temporal and anterior areas. In particular, the arcuate fasciculus pathway connecting the posterior STG and the posterior part of Broca's area has been implicated in processing syntactic complexities.

Syntax is notoriously difficult for L2 learners to grasp, and late learners rarely reach native-like syntactic competence. Unlike semantic representations, which likely are shared for both languages, syntactic processing necessarily differs according to the rules of the language. Whether syntax in L2 builds on L1 syntax and takes advantage of networks already in place for L1 is of great interest. Importantly, both AoA and proficiency are thought to play a role in the degree of L1/L2 neural overlap for syntactic processing.

Consistent with this view, less overlap is found for the neural correlates of syntactic processing in a bilingual's two languages and the degree of overlap appears to reflect the syntactic similarity between the languages. For instance, overlap in the left inferior frontal, left precentral, and left middle temporal regions was found for syntactic processing in late German-English bilinguals who were moderately proficient in English. To circumvent the spatial resolution limitations of fMRI, Weber and Indefrey (2009) used cross-language syntactic repetition priming to identify whether the same neuronal populations are engaged for syntactic processing in both languages. Most studies comparing two languages use a voxel-based method in which they contrast activation maps in each language and determine how many voxels overlap between languages. Because each voxel contains multiple neurons, we cannot discount the possibility that different neurons within a voxel may be responsive to each language. Weber and Indefrey took advantage of the fact that neuronal responses to the same stimulus features are often temporarily reduced as the neurons adapt to repetitions of that stimulus feature. Thus, when a given syntactic structure is primed, repetition of that structure should result in reduced activation if the same neurons respond (i.e., repetition suppression). They applied this principle by priming syntactic structures in one language and determining if repetition suppression occurred when the structure was repeated in the same or a different language. They found clear within-language and between-language repetition suppression, indicating that the same neurons in the left inferior frontal, precentral, and middle temporal cortex were responding to specific syntactic properties of both languages. The repetition effect was not statistically different within and between languages. Thus, there is strong evidence of a shared syntactic processing system in German and English.

However, syntactic processing in Chinese and English is less overlapping. Reading in English is heavily governed by syntactic rules, including word order (e.g., the subject generally precedes the verb and object) and subject-verb agreement. In contrast, Chinese is less reliant on syntactic rules such as word order and lacks the requirement of subject-verb agreement. Instead, sentence meaning is largely derived from semantics and pragmatics. As such, English speakers exhibit differences in neural recruitment to perform semantic and syntactic processing tasks (Friederici, Opitz, & von Cramon, 2000). In contrast, Chinese syntactic processing appears to rely heavily on a semantic network of left inferior prefrontal and left middle and superior temporal cortices, with additional recruitment of left middle frontal cortex when syntactic plausibility judgments are made. This same semantically driven network is employed by Chinese-English bilinguals when they read in English, indicating that the existing L1 syntactic system is applied when reading in L2 (Luke, Liu, Wai, Wan, & Tan, 2002).

Thus, syntactic processing appears to be modulated by the specific properties of a language, and the degree of cross-language neural overlap depends on the degree of similarity in the syntactic rules of the languages. Because syntax is differentially represented across languages, the degree of proficiency attained in each language and AoA also play a role in the degree of neural overlap between languages. We will address these individual differences in the next section. The most common finding is that the areas of activation for syntactic processing overlap heavily in both languages, but that the patterns of activation, and therefore the mechanisms of syntactic processing, can differ for processing in L1 and L2.

Proficiency and AoA Differences in Syntactic Processing

Different linguistic processes are often found to affect the same brain region. That is, some regions may not be specific to only one task. For instance, different aspects of the left IFG have been implicated in semantic, phonological, and syntactic processing (Golestani et al., 2006). This area has been identified as a hub in which syntactic, phonological, and syntactic information are integrated (Hagoort, 2005). The pattern of activation within the IFG depends on the demands of the task at hand (Tatsuno & Sakai, 2005). Similarly, areas typically recruited for non-language tasks, such as ACC and prefrontal substrates of conflict management, are often recruited during complex processing or when a large degree of conflict is present during a language task (as with switching rapidly between languages). This has been taken as evidence that the same brain region may be recruited for different tasks, levels of difficulty, or language processes, but the mechanisms involved for each task may differ. Likewise, it stands to reason that, although largely overlapping brain areas are shared in L1 and L2 syntactic processing, the underlying neural mechanisms may differ for processing in each language.

This claim has been substantiated by three studies varying the degree of syntactic complexity of sentences in L1 and L2, as a measure of increased cognitive demand (workload). Both processing in the weaker L2 and processing complex sentences are assumed to increase cognitive demand. Hasegawa, Carpenter, and Just (2002) presented auditory sentences with easy affirmative construction and with more difficult negative construction to Japanese learners of English. They showed that the areas of activation were similar in English and Japanese, but that activation intensity within these areas was greater for L2. While complexity did not affect activation in L1, more complex sentences in L2 increased activation in the left inferior frontal, superior/middle temporal, and parietal cortices. They concluded that different activation patterns in L1 and L2 reflect the increased cognitive demand in the L2. They suggested that difficulty in processing can be compensated by recruitment of additional areas, and neural activity is sensitive to task demands.

Active and more difficult passive sentences were compared in a similar study of Japanese learners of English. These same left inferior frontal, superior/middle temporal, and parietal cortices were again activated for L1 and L2 (Yokoyama et al., 2006). Again, these regions were recruited differently for complex sentences in L1 and L2 (but note that the new patterns of complex sentence activation differed from Hasegawa et al., 2002 and were partly ascribed to language-specific differences in thematic role assignment). A third study manipulating syntactic complexity in late Korean-English bilinguals found similar results: a common L1/L2 syntactic system encompassing the left IFG, parietal cortex, and occipital lobe was differentially activated for complex sentences in each language (Suh et al., 2007). Contrary to Hasegawa et al., both of these latter studies found greater sensitivity to complexity in L1 than in L2 in the left IFG, which Suh et al. interpreted as automatic processing of simple syntactic structures in L1, but more controlled processing when difficulty increases, either as a function of language proficiency or syntactic complexity (see Rüschemeyer et al., 2006 for similar results). This would suggest that, as proficiency in L2 increases, syntactic processing should become more automatic, and more native-like activation patterns should be observed. This is known as the *convergence hypothesis*—that L2 acquisition builds on an L1 language system that already exists, and L2 therefore has convergent neural representation with L1 (Green, 2003).

According to the convergence hypothesis, as bilinguals become more proficient in the L2, syntactic processing becomes more native-like (Perani & Abutalebi, 2005). There is evidence of functional reorganization of the regions involved in syntactic processing as L2 proficiency increases. For instance, the difference in left IFG activation for L1 and L2 decreases as proficiency in L2 increases (Golestani et al., 2006; see also Lee et al., 2003). The left IFG is a highly integrative area mediating multiple levels of linguistic processing (Hagoort, 2005) and is implicated both in language control necessary to parse complex sentences and in non-language cognitive control (Koechlin & Jubault, 2006; Koechlin, Ody, & Kouneiher, 2003). This area may be differentially activated for controlled and automatic, native-like processing. Scanning participants while they performed grammaticality tasks before and after L2 learning has provided convincing evidence in favor of this hypothesis. Japanese learners of English (Sakai, Miura, Narafu, & Muraishi, 2004) and Chinese learners of Dutch (Indefrey, Hellwig, Davidson, & Gullberg, 2005) both demonstrate more native-like activation in the left IFG as grammatical competency in the learned language increases. This functional reorganization can occur as early as 6 months after learning (Indefrey et al., 2005).

However, completely native-like activation is rarely found in these studies, as native-like grammatical competency is rare in the mostly late L2 learners studied above. AoA appears to play a strong role in L2 syntactic processing. In one of the few studies contrasting the effects of L2 proficiency and AoA on syntactic processing during reading, both AoA and proficiency were found to affect the neural substrates involved in syntax. Wartenburger et al. (2003) compared early high proficiency, late high proficiency, and late low proficiency Italian-German bilinguals. Large differences in brain activation were observed in early and late high proficiency groups during a grammatical judgment task, in which participants judged sentences to be grammatically correct or incorrect. As with the studies mentioned above, greater activity was found in the IFG for the L2 than the L1 in these late learners despite having attained a high level of proficiency (although the authors mention the effect may be due to a slight proficiency difference between L1 and L2 in this group). In contrast, the early learners showed no difference between the neural substrates of L1 and L2 syntactic processing. Differences were also observed between low and high proficiency late learners over posterior areas, but these proficiency-related differences were less obviously linked to controlled vs. automatic syntactic processing than the AoA distinction. A later study of regular and irregular verb processing in Spanish confirmed that late learners exhibit increased left IFG activity as compared to early learners of matched proficiency (Hernandez, Hofmann, & Kotz, 2007). Thus, both proficiency and AoA appear to modulate the neural areas recruited for syntactic processing, with AoA playing a larger role in the degree of neural overlap between areas mediating syntactic processing in each language.

To summarize, reading appears to rely on a common neural system for both languages. In particular, a single semantic network stores conceptual representations that do not differ substantially between languages. However, specific orthographic and syntactic properties of L1 affect the organization of this system. The L1 network is highly distributed, with many areas of integration, but there is also specialization for specific linguistic processes that are differentially recruited depending on the task demands (e.g., semantic vs. syntactic judgments). When two languages are acquired from birth, there appears to be little or no difference between the reading networks. When an L2 is acquired later in life, it builds on the existing orthographic, semantic, and syntactic networks used in L1. As L2 proficiency increases, L2 processing becomes more native-like and activation patterns more closely resemble those of L1 processing. With this shift, certain aspects of syntactic processing may become more automatic than controlled. Finally, when proficiency is low or processing is difficult, additional resources from a cognitive control network can be recruited to manage the syntactic complexity in a more controlled fashion.

Limitations of fMRI

The experiments described above demonstrate the value of fMRI for understanding reading comprehension. Before designing an fMRI experiment, it is important to take into account the limitations of this modality and how these limitations might be circumvented by careful design, cautious interpretation, and seeking converging evidence from complementary techniques.

Functional MRI is notoriously temporally disadvantaged compared to other techniques such as ERPs and MEG. The BOLD signal begins 2 s and peaks around 6 s after neuronal activity (Savoy et al., 1994). The slow hemodynamic response biases fMRI capability towards sustained events, rather than the rapid transient processes involved in sentence reading. Sentences are often read at a rate of 2–3 words per second, making it difficult to have subjects read at a natural pace and isolate the response to a single word. Moreover, we know from electrophysiological studies that multiple processes in sentence and single-word comprehension begin and/or complete within the first second of reading the word. The temporal limitation of fMRI should be taken into account when attempting to determine the location of a rapid process during reading comprehension. However, in well-controlled experiments (e.g., using the same preceding sentence context in multiple conditions) with appropriate modeling of the hemodynamic response, it is possible to attribute differences in activation patterns between conditions to a given factor. It is important to note that each fMRI experiment is only as good as its design and careful consideration should be made during the design phase to create manipulations that isolate the function of interest.

Furthermore, it is important to consider how the statistical measures employed may bias the results (e.g., see Cox & Savoy, 2003 for a comparison of multivariate vs. univariate statistics in detecting distributed patterns of fMRI activity). Because statistical analysis involves comparison of conditions in small voxels, it has been argued that localized activity is typically visualized while more distributed networks may be overlooked. Distributed networks are better studied when informed by our knowledge of the pathways through the brain. Recent functional connectivity and diffusion tensor imaging (DTI) methodologies have made tremendous strides in mapping the connections between co-activated brain regions.

Functional Connectivity MRI

Functional connectivity refers to the degree to which brain regions are co-activated, or activated in response to the same task or stimulus condition. Functional connectivity MRI is a type of fMRI analysis that represents statistical relationships in the BOLD signal of distributed brain regions. Regions with correlated activity or spectral coherence may be functionally interdependent. Functional connectivity is a useful way to reveal experience-related changes in neural networks (e.g., acquiring an L2 or enhanced proficiency). Of particular interest to researchers of bilingual reading comprehension, functional connectivity can be measured during task performance or when participants are in a resting state. Measuring connectivity at rest eliminates confounds presented by performance differences, and therefore task-demand load, between two groups of different proficiency levels.

Recall that L1 and L2 appear to be subserved by the same neural structures but that the mechanism of processing may differ between the two languages. A functional connectivity study was conducted which demonstrates this point (Dodel et al., 2005). While the same brain areas are engaged for L1 and L2 syntactic tasks, these areas are functionally more connected during L2 syntactic tasks than during the same tasks in L1. Functional connectivity studies are useful in corroborating suggested differences in processing mechanisms involving the same regions. They are also useful in building models of connected network activity and potential pathways of comprehension. Recent advances allow us to estimate the connectivity during discrete stages of processing by modeling the hemodynamic response at every stage of processing and determining which regions of the brain show correlated fluctuations across trials (Rissman, Gazzaley, & D'Esposito, 2004).

Importantly, functional connectivity is distinct from anatomical connectivity. Brain regions may be anatomically connected but respond differently to a given stimulus condition. Conversely, regions that are anatomically distinct may be co-activated. To measure anatomical connectivity, other measures such as DTI are employed. We can construct anatomical pathways of processing by combining evidence from DTI and functional connectivity studies for a comprehensive picture of connectivity during processing (see Fig. 12.6).

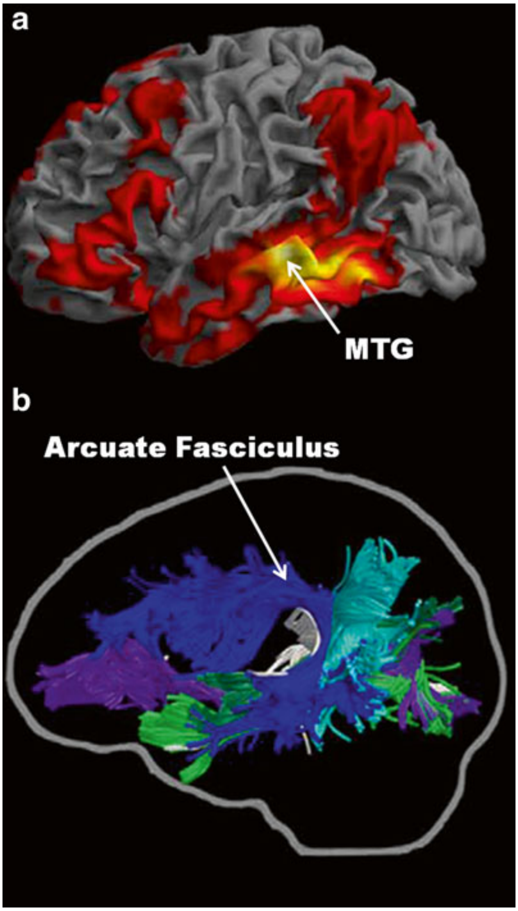


Fig. 12.6 Together, functional connectivity and DTI have illuminated language comprehension networks. Here we show resting state functional connectivity (**a**) and structural connectivity (**b**) profiles of the MTG. Both functional connectivity and DTI reveal that the MTG is a highly connected region in the language comprehension network. The arcuate fasciculus, a dorsal path involved in syntactic processing, is tracked in blue. Adapted "The Neural Architecture of the Language Comprehension Network: Converging Evidence From Lesion and Connectivity Analyses" by A.U. Turken, A.U. and N.F. Dronkers, *Frontiers. in Systems. Neuroscience, 5,* pp. 1–20. Copyright 2011 by Turken and Dronkers

DTI

Functional MRI measures changes in blood flow during task performance. Using fMRI, we can tell which regions are involved in performing a task and which areas work together. But we cannot determine if regions are anatomically connected. DTI takes into account the diffusion of water as it moves through the white matter tracts in the brain and allows us to reconstruct anatomical pathways through the brain.

In a homogenous medium, water diffusion is described by Brownian motion, meaning that the direction of diffusion is random and isotropic, that is, it occurs equally in all directions. Most long axons that connect regions of the brain are wrapped in myelin, a fatty substance that speeds signal propagation. Importantly, myelin constrains diffusion of water such that the principle direction of diffusion will be along the interior of the axon. We can measure the degree of anisotropy, that is, how elliptical the overall path of diffusion is as water moves in different directions. To measure diffusion with MRI, first a gradient pulse is applied such that protons are knocked out of alignment. An equal pulse in the opposite direction is then applied to realign the protons. Protons which have not moved (i.e., diffused) will cause no net difference in the signal observed before and after the equal and opposite pulses, but any movement will be detected as a change in the signal. We then create a tensor, or a model of the anisotropic diffusion along the axon.

DTI is useful in a number of ways, but most relevant here is tractography, also known as *fiber tracking*. In tractography, we essentially connect voxels with the same principle diffusion direction to track fibers through the brain and obtain a picture of white matter tracts connecting brain regions. The posterior-to-anterior dorsal and ventral pathways involved in comprehension have been mapped using fMRI results to inform fiber tracking (Ben-Shachar, Dougherty, & Wandell, 2007; Brauer, Anwander, Perani, & Friederici, 2013; Friederici, 2009; Turken & Dronkers, 2011). A dorsal pathway from the sensory areas (e.g., auditory or visual cortex), through the temporal cortex, to the premotor cortices is involved in mapping orthography/auditory speech perception onto phonology. The arcuate fasciculus is another dorsal pathway involved in complex syntactic processing, which connects the MTG and the posterior aspect of Broca's area in the IFG (see Fig. 12.6 for functional and structural connectivity between these regions). A ventral pathway from sensory areas to temporal and frontal areas is involved in direct semantic retrieval (Brauer et al., 2013; Friederici & Gierhan, 2013). DTI studies have contributed greatly to our understanding of the language pathways involved in comprehension. However, only combination with fMRI and other methodologies has revealed the flow of processing and differences in *how* bilinguals comprehend during reading.

Converging Evidence

Together with behavioral and electrophysiological methods, we can test predictions of bilingual reading comprehension models. A predominant theory of reading comprehension is the Bilingual Interactive Activation Plus Model (BIA+) proposed by Dijkstra and van Heuven (Dijkstra & van Heuven, 2002). While other theories of reading comprehension exist, the BIA+ will be used to demonstrate how neuroimaging data can provide converging evidence with other methodologies to substantiate and lead to adjustments in such models. This model has been selected in particular to highlight the role of cognitive control and task demands during bilingual reading comprehension, as prefaced in the discussion of proficiency.

The BIA+ model has at its core a language processing system which is subdivided into a word-level identification system and a sentence processing system. The model allows for the Dual-Route Theory of processing, in which orthographic representations simultaneously activate phonological and semantic information. As discussed above, tremendous support in favor of Dual-Route processing has been found using fMRI. Core assumptions of the BIA+ model are that the lexicon is integrated and access is language nonselective. This means that orthographic, phonological, and semantic features that are shared across languages have a central representation and words in both languages are activated when we read. As we have seen from fMRI studies, there is support for an integrated lexicon in terms of a large amount of overlap in the neural representations of both of a bilingual's languages. Behavioral and ERP studies using cross-language orthographic neighbors and priming have shown that word recognition in one language is influenced by the other language (Thierry & Wu, 2004, 2007; Van Heuven, Dijkstra, & Grainger, 1998). Together findings from these three different methodologies suggest that both languages are stored in an integrated lexicon and are activated in parallel, as predicted by the BIA+ model.

As a result of parallel access to overlapping representations in both languages, a fair degree of interference from one language may be anticipated while reading in the other. Thus, bilingual reading comprehension may require cognitive control to suppress irrelevant words and syntactic representations in the other language. The prevailing theory is that when the L2 is weaker than L1, words and syntactic rules in L1 may be more quickly activated than those in L2 and may interfere with L2 processing. Processing in the second language is more effortful than processing in the first language, and therefore additional resources are recruited for controlled L2 processing. Cognitive control brain areas, including a task/decision system, may prevent or inhibit interference from the other language, especially when a bilingual is reading in their weaker language.

Dijkstra and van Heuven's model has been modified to accommodate evidence from recent neuroimaging studies of a functionally and anatomically distinct task/decision system. The BIA+ now specifies a task/decision system which maintains the goals at hand and mediates decision making relevant to the task (e.g., decide if the words rhyme or have related meanings). Importantly, the original BIA model did not contain a separate task system, but converging evidence of such a system from behavioral studies, ERPs, and neuroimaging led to this modification (Dijkstra & van Heuven, 1998; Van Heuven & Dijkstra, 2010).

Complementary results from fMRI and TMS suggest that modifying the task demands during reading alters the control networks involved and necessarily engaged, respectively (Moss, Schunn, Schneider, McNamara, & VanLehn, 2010; Nakamura et al., 2006). In addition, ERPs recorded while participants read code-switched words embedded in sentences revealed that access to meaning is independent of identifying the language (i.e., the reading task; Blackburn, 2013). Instead of affecting an ERP index of semantic retrieval (the N400), code switches elicit a positivity which resembles a task-reconfiguration ERP component, and may be evidence of a distinct task system (although the authors suggest other possible accounts; see for example, Blackburn, 2013; Moreno, Federmeier, & Kutas, 2002; Moreno, Rodríguez-Fornells, & Laine, 2008; Van Der Meij, Cuetos, Carreiras, & Barber, 2011). Critically, ERP indices of task-reconfiguration are localized to the dorsolateral prefrontal cortex (Jamadar, Provost, Fulham, Michie, & Karayanidis, 2009), an area which is activated both during cognitive control tasks and during effortful reading (Miller & Cohen, 2001; Rodriguez-Fornells, Rotte, Heinze, Noesselt, & Münte, 2002). Therefore, the prefrontal cortex appears to be part of a task/decision system which aids in maintaining the language of reading and task-relevant goals. This system may mediate effortful processing both during complex L1 processing and processing in the weaker language.

Functional MRI has been instrumental in the ongoing mission to identify the anatomical substrates of this task system and ascertain which aspects are language specific and nonspecific. The left IFG, ACC, and prefrontal cortex are typically more active for L2 than L1 processing. The prefrontal cortex and ACC are implicated in both language and non-language tasks that require cognitive control, or the ability to manage conflicting information (Abutalebi & Green, 2007, 2008; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Miller & Cohen, 2001; Rodriguez-Fornells et al., 2002). These regions are part of a general cognitive control network, which also includes the caudate nucleus and left IPL. This network can be recruited during effortful language tasks, and recruitment depends on the nature of the task (Abutalebi, 2008; Abutalebi et al., 2007; Abutalebi & Green, 2007). The left IFG appears to be involved in cognitive control as well as having specific roles in language. This region is subdivided into areas specialized for syntactic and semantic interference during processing (Glaser, Martin, van Dyke, Hamilton, & Tan, 2013) and is thought to play a role in integrating linguistic processes and maintaining goals for the task at hand (Hagoort, 2005; Miller & Cohen, 2001). In fact, the left inferior frontal cortex is more active for bilinguals than monolinguals during reading tasks and may be the seat of ambiguity resolution between activated representations in a bilingual's two languages during reading (Kovelman, Baker, & Petitto, 2008).

Thus, there is some degree of evidence from behavioral, ERP, fMRI, and TMS studies in favor of a functionally and anatomically distinct task system which is affected by cross-language or syntactic conflict and degree of control needed during reading. As such, this task/control system has been implemented in the later version of the BIA model, as well as other prevalent models (Hagoort, 2005; Van Heuven & Dijkstra, 2010). This example illustrates the value of converging evidence from different modalities to shape bilingual reading models.

Simultaneous Recording and Novel Methods

The temporal precision of ERPs complements the spatial resolution of fMRI. We can employ these methods concurrently or use data from one modality to constrain analysis in the other. One way to combine techniques is to isolate areas of activity using fMRI and apply this knowledge to ERP analysis. Unlike the slow hemodynamic response measured by fMRI, ERPs allow us to measure processes occurring within milliseconds after stimulus presentation. Unfortunately, due to the many convolutions of the brain and properties of signal propagation, brain potentials recorded at the scalp may originate from any of a set of sources. This is known as the *inverse problem*. But if we know the source in the brain of the ERP signal, we can determine where on the scalp this activity would be recorded (i.e., the *forward solution*). Activated regions of interest identified using fMRI can be used to constrain the inverse problem and effectively increase the spatial precision of ERPs.

A frequently used technique to estimate the distribution of the ERP current and identify the most likely neural sources of activity recorded at the scalp is *low resolution electromagnetic tomography* (LORETA; Pascual-Marqui, Michel, & Lehmann, 1994). We can use localization techniques like LORETA to estimate where activity from a known source will be observed at the scalp and determine if this scalp-observed activity is correlated with or converges with fMRI activation identified for the same process (e.g., Meyer, Obleser, Kiebel, & Friederici, 2012). For instance, ERP studies have effectively isolated the time course of semantic and syntactic processing. We know from fMRI studies that semantic access is tied to activity in the angular gyrus and the posterior temporal lobe more generally. The inferior frontal lobe is activated for many processes and is fairly consistently activated during syntactic processing (Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004). Converging evidence from fMRI activity patterns and LORETA estimates during known time intervals of semantic and syntactic processing implicate the posterior middle temporal lobe and left angular gyrus in semantic retrieval during the first 500–600 ms of processing and the inferior frontal cortex and anterior temporal lobe in syntactic processing around 500–1000 ms after stimulus onset (Osterhout et al., 2004).

A particularly promising methodology entails combining techniques in the same study. A novel method of combining fMRI, TMS, and EEG has been used in visual attention studies (see Fig. 12.7; Zanto, Rubens, Thangavel, & Gazzaley, 2011). First, fMRI is used to identify regions in each individual that may be involved in a given task. Next, TMS is used to render these areas ineffective. TMS is a way to selectively stimulate neuronal populations to the point that they are unable to respond naturally for a short period of time, similar to a transient lesion. Note that while the lesion studies discussed at the onset of this chapter only reveal regions that are *necessary* but not the regions *involved* in a function, fMRI highlights all regions that are *involved* but cannot easily distinguish those that are *necessary* for a function. TMS and fMRI are therefore complementary techniques. Before neurons recover from TMS, we can determine if ERP responses linked to rapid processing

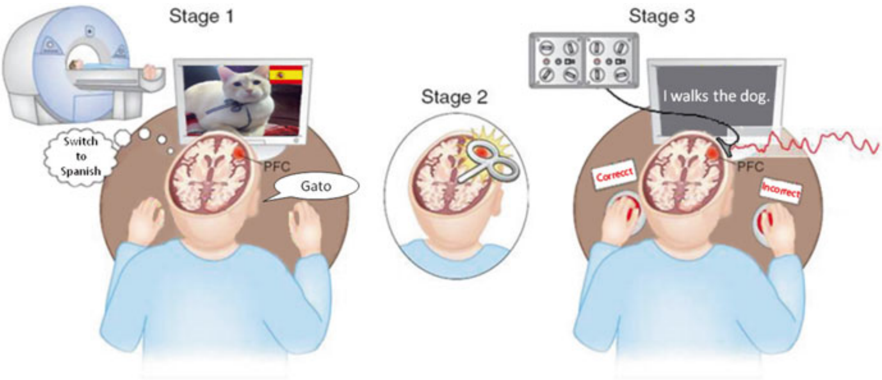


Fig. 12.7 Reading experiment modeled after the converging techniques employed by Zanto et al. (2011). Functional magnetic resonance images are obtained to localize region of interest (ROI) activity while participants perform a task. TMS application is guided by the ROIs obtained using functional MRI to inhibit the region of interest. ERPs are then recorded to determine if the regions stimulated by TMS are necessary for the new task. Adapted from "The visual Attention Network Untangled," by S. Nieuwenhuis and T.H. Donner, 2011, *Nature Neuroscience, 14,* pp. 542–543. Copyright 2011 by Macmillan Publishers Ltd

are modulated by rendering the stimulated area ineffective (see also Nieuwenhuis & Donner, 2011).

For instance, we could combine fMRI, TMS, and ERPs to test if the same cognitive control areas engaged during language production are involved in complex syntactic processing, and how this may be affected by proficiency. First, fMRI would be used to identify regions involved in controlled language production. For example, we can subtract a low-control condition, in which participants name pictures in only one language, from a high-control condition, in which participants are cued to switch rapidly between languages while naming (Fig. 12.7, Stage 1). We might expect to find activation in the prefrontal cortex during controlled language production. We can use the fMRI coordinates obtained for the region of interest in each participant to guide application of TMS (Fig. 12.7, Stage 2). Immediately after TMS, we can record ERPs while participants read syntactic violations in sentences with varying syntactic complexity (Fig. 12.7, Stage 3). If the stimulated region is necessary for syntactic processing, the ERP component typically elicited to syntactic violations will be smaller immediately after TMS than before TMS. This result would imply that the same stimulated region is engaged for reading complex sentences and switching between languages during picture naming. Further, we might predict that low proficiency bilinguals would be affected for simple and complex sentences, but high proficiency bilinguals may only show effects on complex sentences. This would support the claim that controlled processing is always used by low proficiency bilinguals but automatic processing is used for simple syntactic constructions in high proficiency bilinguals. Combining techniques within a given study opens doors to the kinds of questions we can ask.

Each of the available neuro-cognitive techniques provides a sliver of information regarding either where or when various aspects of processing occur. Linking findings from fMRI studies with ERP and behavioral studies lends spatial and temporal resolution necessary to create models of comprehension at all levels of processing. For this reason, it is important to combine results from various techniques and obtain converging evidence regarding the underlying processes in reading comprehension.

Summary and Conclusions

The advent of MRI techniques has enabled us to answer questions about how reading networks are organized in the bilingual brain. Structural MRI provides a static image of the brain and has revealed anatomical changes that are a result of bilingualism. In functional MRI, we obtain a statistical map of activity in the brain. This method has been used to identify the processing routes for reading, determine whether cognitive comprehension processes are language specific, and identify the degree of neural overlap in L1 and L2 reading networks. We have also used fMRI to investigate the role of proficiency and AoA on the organization of the second language around the existing first language. Functional connectivity analysis identifies regions that are co-activated and confirms that, although the same brain regions are often engaged for L1 and L2 processing, the mechanism of processing may differ in the stronger and weaker language. Finally, DTI is useful for mapping the pathways connecting regions revealed by fMRI to mediate processing.

Together, these techniques have been used to test models of reading comprehension. Overall, reading in L1 and L2 is subserved by a largely overlapping network, with additional recruitment for language-specific aspects of processing. Late L2 learners tend to apply language-specific strategies from their L1 to their L2, when possible. When this is not possible, network activity is fairly rapidly modified to accommodate the demands of the new language. Thus, L2 acquisition builds on the existing L1 system and as L2 proficiency improves, processing becomes more native-like. This is known as the convergence hypothesis.

Functional MRI studies have elucidated the networks involved in each processing stage and degree of overlap between L1 and L2 in each of these processing networks. Reading relies on a number of left-lateralized dorsal and ventral pathways that project anteriorly from the visual cortex to a distributed fronto-temporal network. The orthographic transparency of a language plays a role in the processing route engaged. Transparent languages that are easily mapped to phonology can engage an indirect word form to meaning pathway through phonological processing in the STG. Opaque and logographic languages are more likely to be processed via a direct route from word form to lexico-semantic access. Semantic access to a shared conceptual representation in both languages is mediated by largely overlapping brain regions, including the IFG, MTG, and STG. Access to meaning is more controlled for the weaker L2 than the L1, but AoA plays a smaller role in the neural organization of semantic networks. On the other hand, syntactic processing is less consistent across languages and is modulated in particular by AoA. The degree of cross-language neural overlap in syntactic processing also depends on the degree of similarity between syntactic rules in the language and proficiency. Areas of activation for syntactic processing overlap heavily in both languages, but the patterns of activation, and therefore the mechanisms of syntactic processing, can differ for processing in L1 and L2. Enhanced intensity in the left IFG, ACC, and prefrontal cortex for processing in L2 than in L1 implies that reading may involve an anatomically and functionally distinct task-specific control system. When processing is effortful or involves conflict between two languages, as with processing in the L2, this general cognitive control system is recruited in concert with the language system. The IFG in particular appears to play a role in managing conflict during reading in one language that results from interference from the other language. These findings are consistent with the BIA+ model.

While fMRI has been instrumental in mapping bilingual reading pathways, any new design should take into account the limitations in spatial resolution and the need for a well-designed contrast or parametric manipulation. Patterns of activation vary across studies attempting to isolate the same process, which are often due to slight differences in task demands or experimental design. Meta-analyses are a useful way to highlight consistencies and determine factors responsible for inconsistencies across studies. Combining fMRI with temporally precise measures such as ERPs combats the temporal and spatial limitations of each, respectively. TMS is another complementary technique that can determine the necessity of regions identified by fMRI to be involved in a process. The most complete models of reading comprehension will take into account converging evidence from different neuro-cognitive modalities.

List of Keywords

Age of acquisition (AoA), Arcuate fasciculus pathway, Automatic syntactic processing, Bilingual Interactive Activation Plus Model (BIA+), Blocked design, Blood-oxygen level dependent (BOLD), Broca's area, Cognitive control tasks, Computed tomography (CT), Controlled processing, Convergence Hypothesis, Diffusion tensor imaging (DTI), Direct route, Dominant language, Dorsal pathway, Dual-Route Model, Dyslexia, Event-related potentials (ERPs), Event-related design, Fiber Tracking, Font size decision task, Functional Magnetic Resonance Imaging (fMRI), Fusiform gyrus, Grammatical judgment task, Grapho-phonological route, Hemodynamic response, Homophones, Identification system, Indirect route, Inferior frontal cortex, Inverse problem, Lexico-semantic access, Lexico-semantic route, Low resolution electromagnetic tomography (LORETA), Magnetic Resonance Imaging (MRI), Magnetic-encephalography (MEG), Meta-Analyses, N400, Neuroimaging Techniques, Orthographic opacity, Orthographic transparency, Phonological rhyming task, Positron-emission tomography (PET), Posterior-to-anterior reading network, Pragmatics, Region of interest (ROI), Repetition effect, Repetition priming, Repetition suppression, Revised Hierarchical Model, Rhyming decision, Semantic decision tasks, Semantic-noun-verb association task, Sentence-level processing, Syntactic processing, Syntactic rules, Syntactic system, T1 decay, T1-weighted image, T2 decay, T2-weighted image, Talairach space, Task-reconfiguration ERP component, Task/decision system, Match-to-sample task, Tractography, Transcranial magnetic stimulation (TMS), Ventral occipito-temporal lexico-semantic route, Ventral reading pathway, Voxel-based method, Word order.

Review Questions

1. What method(s) would be most appropriate to identify brain areas *involved* in a given cognitive process? How could we determine if each of the involved areas is *necessary* for the task?
2. Explain the difference between structural and functional connectivity. What methods are typically used to assess connectivity? Describe how these methods can be used to obtain converging evidence in support of a connected brain network. Find an example either in this chapter or in a recent article in which methods are combined to obtain converging evidence of a brain pathway.
3. What are the different types of fMRI experimental designs? Which do you think is most appropriate to use? Think of at least one example where each design would be desirable and explain why. What are some of the limitations of these designs and how might you take these into account when designing a study?
4. Explain the convergence hypothesis. What predictions does the convergence hypothesis make regarding neural organization as proficiency in the second language improves?

Suggested Student Research Projects

1. An excellent website detailing how to construct a canonical hemodynamic response curve and how to implement the most efficient fMRI design for a study can be found at: http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency#The_BOLD_impulse_response_.28IR.29. Go to this website and read about how the interstimulus interval (ISI) affects the hemodynamic response curve. Sketch a canonical hemodynamic response curve or create one in a spreadsheet such as Excel or LibreOffice (libreoffice.org). On the *x*-axis, label the time in seconds for the initial dip, peak response, and return to baseline. Now create a predicted response curve when stimuli are presented in rapid succession. What factors may affect the BOLD response curve? (Hint: is it the same in everyone or in every brain region?)

2. There are a number of ways to design a reading comprehension fMRI study. As we have seen, activity patterns differ slightly between studies due to differences in the stimuli, control conditions, task demands, and participants. Consistencies across studies often highlight the process we are attempting to capture. Meta-analyses are good ways to compile consistencies and determine potential factors responsible for experimental differences. Conducting a meta-analysis using BrainMap is a good place to start a new study and to get an idea of experimental parameter options. Conveniently, this database contains multiple types of imaging so we can search for converging evidence across imaging modalities or constrain the search to fMRI. To get started, go to brainmap.org and enter the BrainMap under the "Tools" tab (see Fig. 12.8).

   To find bilingual sentence reading data, select the conditions "visual" as the stimulus modality, "read" as the instructions. Over 100 reading experiments appear, with the number of subjects and the imaging modality easily referenced. We can constrain the search by adding the experimental parameter "linguistic

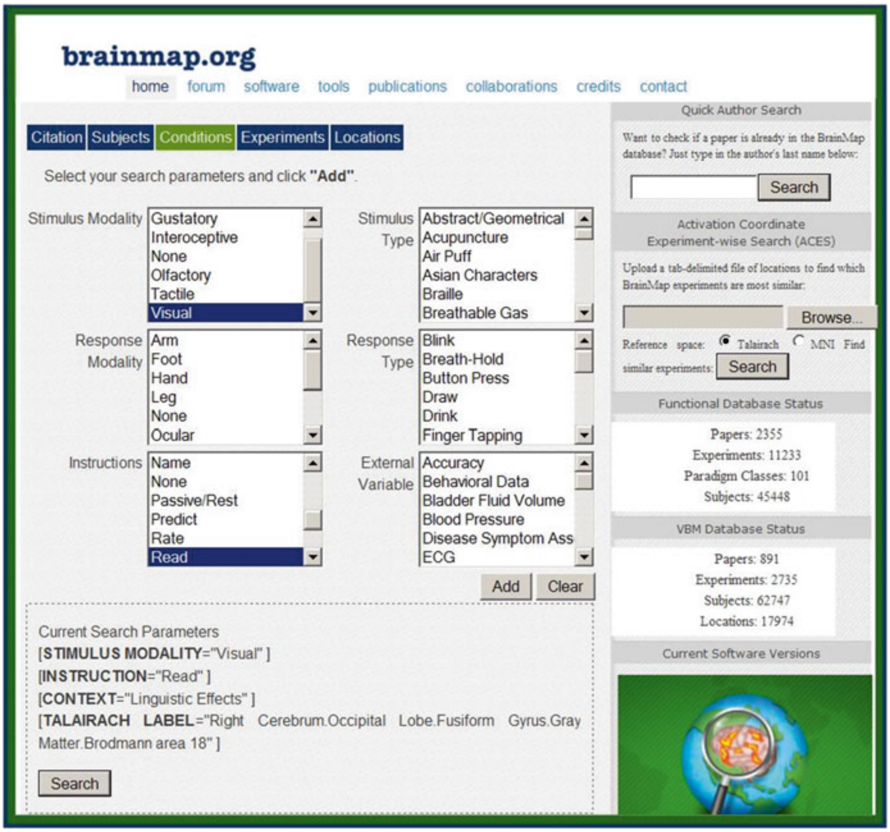


Fig. 12.8 The BrainMap database is a useful meta-analysis tool, available at www.brainmap.org. Parameters are selected here to constrain the search to sentence reading studies

effects" as the context. To save options before moving between screens, it is important to select the add button. This search results in 13 studies. We could also constrain the analysis to only stimuli of words, falsefonts, or pseudowords, or to subjects with a particular native language. We can further constrain the dataset to include only those studies in which activation in the right fusiform gyrus was observed, narrowing results to only one study on word form processing in Chinese (Fu, Chen, Smith, Iversen, & Matthews, 2002).

3. As you can see, this is an efficient way to determine both which brain areas are consistently obtained in studies of a given design and what types of studies evoke activity in a given region of interest. Another way to use this database is to determine in what types of tasks brain regions appear to be functionally connected. For instance, we can select Broca's area (BA 44/45) and Wernicke's area (BA 22), regions classically thought to mediate language production and comprehension, respectively, to determine the types of studies in which both regions are co-activated. Based on the classical descriptions of these areas, we might expect to find studies on comprehension and production. As expected, one of the few results refers to the common language network for comprehension and production (Papathanassiou et al., 2000). To conduct a classic fMRI study, it is important to consider whether an event-related or blocked design is best suited to your research question. Most studies of sentence processing will be best addressed with an event-related design. Next, a well-designed baseline subtraction condition or parametric manipulation is necessary to isolate the process of interest. An important consideration will need to be made regarding the software (which includes hemodynamic response modeling) and coordinate space to use. More studies are moving to MNI space, as this space is representative of more than one brain. Finally, if comparing groups, it will be important, as with any psycholinguistic study, to carefully consider sample size and on which factors the groups should be matched. As a final precaution, subjects should be able to hold still in tight spaces for extended periods to avoid artifacts caused by movement in the scanner. This may make certain populations (e.g., young children learning to read or claustrophobic individuals) difficult to study.

   Keeping these factors in mind, design a simple fMRI study to test a hypothesis you have about sentence reading. State the conditions you will be comparing, who your participants will be, and what kind of fMRI design you will use. If you chose to use a subtraction method, does your baseline adequately isolate the process of interest?

Related Internet Sites

Brainmap: http://www.brainmap.org/
fMRI Statistical Modeling: http://mindhive.mit.edu/node/55
How fMRI works: http://www.csulb.edu/~cwallis/482/fmri/fmri.html
MRC Cognition and Brain Sciences Unit: http://imaging.mrc-cbu.cam.ac.uk/

Suggested Further Reading

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128*(3), 466–478.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *NeuroImage, 20*(2), 693–712.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat Neurosci, 14*(5), 656–661.

Acknowledgement The author would like to thank Christopher Hwozdek for helping to adapt figures and Charlie Giattino, Charlotte Blackburn, and the editors for their helpful comments.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128*(3), 466–478.

Abutalebi, J., Brambati, S. M., Annoni, J.-M., Moro, A., Cappa, S. F., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *The Journal of Neuroscience, 27*(50), 13762–13769.

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*, 242–275.

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes, 23*(4), 557–582.

Ben-Shachar, M., Dougherty, R. F., & Wandell, B. A. (2007). White matter pathways in reading. *Current Opinion in Neurobiology, 17*, 258–270.

Binder, J. R., Rao, S. M., Hammeke, T. A., Frost, J. A., Bandettini, P. A., & Hyde, J. S. (1994). Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. *Cognitive Brain Research, 2*, 31–38.

Blackburn, A. M. (2013). *A study of the relationship between code switching and the bilingual advantage: Evidence that language use modulates neural indices of language processing and cognitive control.* (Unpublished doctoral dissertation). University of Texas at San Antonio, San Antonio.

Bookheimer, S., Zeffiro, T. A., Blaxton, T., Gaillard, W., & Theodore, W. (1995). Regional cerebral blood flow during object naming and word reading. *Human Brain Mapping, 3*, 93–106.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *NeuroImage, 16*, 7–22.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179–181.

Brauer, J., Anwander, A., Perani, D., & Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain and Language, 127*(2), 289–295.

Brunswick, N. (2010). The functional neuroanatomy of reading. In P. L. Cornelissen, P. C. Hansen, M. L. Kringelbach, & K. R. Pugh (Eds.), *The neural basis of language*. New York: Oxford University Press.

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz. *Brain, 122*, 1901–1917.

Büchel, C., Holmes, A. P., Rees, G., & Friston, K. J. (1998). Characterizing stimulus–response functions using nonlinear regressors in parametric fMRI experiments. *NeuroImage, 8*, 140–148.
Buchweitz, A., Mason, R. A., Hasegawa, M., & Just, M. A. (2009). Japanese and English sentence reading comprehension and writing systems: An fMRI study of first and second language effects on brain activation. *Bilingualism: Language and Cognition, 12*(2), 141–151.
Chee, M. W. L., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W. L., Thiel, T., et al. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron, 23*(1), 127–137.
Chee, M. W. L., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage, 13*(6), 1155–1163.
Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience, 19*(8), 3050–3056.
Chee, M. W. L., Weekes, B., Lee, K. M., Soon, C. S., Schreiber, A., Hoon, J. J., et al. (2000). Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: Evidence from fMRI. *NeuroImage, 12*(4), 392–403.
Chen, Y., Fu, S., Iversen, S. D., Smith, S. M., & Matthews, P. M. (2002). Testing for dual brain processing routes in reading: A direct contrast of Chinese character and Pinyin reading using fMRI. *Journal of Cognitive Neuroscience, 14*, 1088–1098.
Chen, J. J., & Pike, G. B. (2009). Origins of the BOLD post-stimulus undershoot. *NeuroImage, 46*(3), 559–568. doi:10.1016/j.neuroimage.2009.03.015.
Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split brain patients. *Brain, 123*, 291–307.
Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage, 40*(1), 353–366.
Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *The Quarterly Journal of Experimental Psychology Section A, 35*(3), 469–495. doi:10.1080/14640748308402483.
Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. doi:10.1037/0033-295X.108.1.204.
Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage, 19*(2), 261–270.
Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature, 4*, 752–758.
Dérouesné, J., & Beauvois, M. F. (1979). Phonological processing in reading: Data from alexia. *Journal of Neurology, Neurosurgery & Psychiatry, 42*, 1125–1132.
Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist Connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Lawrence Erlbaum.
Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197.
Ding, G., Perry, C., Peng, D., Ma, L., Li, D., Xu, S., et al. (2003). Neural mechanisms underlying semantic and orthographic processing in Chinese-English bilinguals. *NeuroReport, 14*(12), 1557–1562.
Dodel, S., Golestani, N., Pallier, C., Elkouby, V., Le Bihan, D., & Poline, J. B. (2005). Condition-dependent functional connectivity: Syntax networks in bilinguals. *Philosophical Transactions of the Royal Society of London, 360*(1457), 921–935.

Fiebach, C. J., Friederici, A. D., Muller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience, 14*, 11–23.

Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences, 13*(4), 175–181.

Friederici, A. D., & Gierhan, S. M. E. (2013). The language network. *Current Opinion in Neurobiology, 23*, 250–254.

Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex, 10*, 698–705.

Fu, S., Chen, Y., Smith, S. M., Iversen, S. D., & Matthews, P. M. (2002). Effects of word form on brain processing of written Chinese. *NeuroImage, 17*(3), 1538–1548.

Glaser, Y. G., Martin, R. C., van Dyke, J. A., Hamilton, A. C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain and Language, 126*, 314–326.

Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., & Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia, 44*(7), 1029–1040.

Green, D. W. (2003). Neural basis of lexicon and grammar in L2 acquisition: The convergence hypothesis. In R. van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The lexicon-syntax interface in second language acquisition* (pp. 197–218). Amsterdam: John Benjamins.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9*, 416–423.

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage, 15*(3), 647–660.

Hernandez, A. E., Hofmann, J., & Kotz, S. A. (2007). Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *NeuroImage, 36*(3), 912–923.

Humphreys, G., & Gennari, S. P. (2014). Competitive mechanisms in sentence processing: Common and distinct production and reading comprehension networks linked to the prefrontal cortex. *NeuroImage, 84*, 354–366.

Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D. E., Glover, G. H., Poldrack, R., et al. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language, 70*(3), 347–363.

Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning, 56*, 279–304.

Indefrey, P., Hellwig, F., Davidson, D., & Gullberg, M. (2005). Native-like hemodynamic responses during sentence comprehension after six months of learning a new language. *Poster presented at the 11th Annual Meeting of the Organization for Human Brain Mapping*. Toronto, Canada.

Ino, T., Nakai, R., Azuma, T., Kimura, T., & Fukuyama, H. (2009). Recognition and reading aloud of kana and kanji word: An fMRI study. *Brain Research Bulletin, 78*(4–5), 232–239.

Jamadar, S., Provost, A., Fulham, W. R., Michie, P. T., & Karayanidis, F. (2009). Multiple sources underlie ERP indices of task-switching. In W. Christensen, E. Schier, & J. Sutton (Eds.), *ASCS09: Proceedings of the 9th Conference of the Australasian Society for Cognitive Science* (pp. 154–161). Sydney: Macquarie Centre for Cognitive Science.

Jamal, N. I., Piche, A. W., Napoliello, E. M., Perfetti, C. A., & Eden, G. F. (2012). Neural basis of single-word reading in Spanish–English bilinguals. *Human Brain Mapping, 33*, 235–245.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *NeuroImage, 20*(2), 693–712.

Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The neural bases of sentence comprehension: An fMRI examination of syntactic and lexical processing. *Cerebral Cortex, 11*, 223–237.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping, 7*, 225–233.

Kim, K. H. S., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*(6638), 171–174. doi:10.1038/40623.

Koechlin, E., & Jubault, T. (2006). Broca's Area and the hierarchical organization of human behavior. *Neuron, 50*(6), 963–974.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*, 1181–1185.

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible "neural signature" of bilingualism. *Journal of Cognitive Neuroscience, 20*(1), 153–169.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetic connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). BrainMap: The social evolution of a functional neuroimaging database. *Neuroinformatics, 3*, 65–78.

Lee, H.-S., Fujii, T., Okuda, J., Tsukiura, T., Umetsu, A., Suzuki, M., et al. (2003). Changes in brain activation patterns associated with learning of Korean words by Japanese: An fMRI study. *NeuroImage, 20*, 1–11.

Liu, Y., Dunlap, S., Fiez, J., & Perfetti, C. (2007). Evidence for neural accommodation to a writing system following learning. *Human Brain Mapping, 28*(11), 1223–1234. doi:10.1002/hbm.20356.

Luke, K.-K., Liu, H.-L., Wai, Y.-Y., Wan, Y.-L., & Tan, L. H. (2002). Functional anatomy of syntactic and semantic processing in language comprehension. *Human Brain Mapping, 16*, 133–145.

Mandeville, J. B., & Rosen, B. R. (2002). Functional MRI. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain mapping: The methods* (pp. 315–349). San Diego: Academic Press.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2*(3), 175–199.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., et al. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature, 431*(7010), 757–757.

Menard, M. T., Kosslyn, S. M., Thompson, W. L., Alpert, N. M., & Rauch, S. L. (1996). Encoding words and pictures: A positron emission tomography study. *Neuropsychologia, 34*, 185–194.

Meschyan, G., & Hernandez, A. (2006). Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation. *NeuroImage, 29*, 1135–1140.

Meyer, L., Obleser, J., Kiebel, S. J., & Friederici, A. D. (2012). Spatiotemporal dynamics of argument retrieval and reordering: An fMRI and EEG study on sentence processing. *Frontiers in Psychology, 3*, Article 523.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Moreno, E., Federmeier, K., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language, 80*, 188–207.

Moreno, E., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics, 21*(6), 477–508.

Moss, J., Schunn, C. D., Schneider, W., McNamara, D. S., & VanLehn, K. (2010). An fMRI study of strategic reading comprehension. *Paper Presented at the Proceedings of the Thirty-Second Conference of the Cognitive Science Society*, Portland, OR.

Mouthon, M. S. (2011). *Neuroimagery of language control mechanism and semantic processing in the bilingual brain with a special emphasis on the effects of second language proficiency.* (Doctorat en Neurosciences), Université de Genève, Genève. Retrieved from http://archive-ouverte.unige.ch/unige:18233 (Neur. 77)

Nakada, T., Fujii, Y., & Kwee, I. L. (2001). Brain strategies for reading in the second language are determined by the first language. *Neuroscience Research, 40*, 351–358.

Nakamura, K., Dehaene, S., Jobert, A., Bihan, D. L., & Kouider, S. (2005). Subliminal convergence of Kanji and Kana words: Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. *Journal of Cognitive Neuroscience, 17*(6), 954–968.

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., & Kouider, S. (2007). Task-specific change of unconscious neural priming in the cerebral language network. *Proceedings of the National Academy of Sciences, 104*(49), 19643–19648.

Nakamura, K., Hara, N., Kouider, S., Takayama, Y., Hanajima, R., Sakai, K. L., et al. (2006). Task-guided selection of the dual neural pathways for reading. *Neuron, 52*, 557–564.

Nakamura, K., Oga, T., Okada, T., Sadato, N., Takayama, Y., Wydell, T., et al. (2005). Hemispheric asymmetry emerges at distinct parts of the occipitotemporal cortex for objects, logograms and phonograms: A functional MRI study. *NeuroImage, 28*(3), 521–528.

Nelson, J., Liu, Y., Fiez, J., & Perfetti, C. A. (2009). Assimilation and accommodation patterns in ventral occipitotemporal cortex in learning a second writing system. *Human Brain Mapping, 30*(3), 810–820.

Nieuwenhuis, S., & Donner, T. H. (2011). The visual attention network untangled. *Nature Neuroscience, 14*(5), 542–543.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. J. Clifton (Eds.), *The on-line study of sentence comprehension: Eye tracking, ERP, and beyond*. New York: Psychology Press.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: A contribution to the definition of language epicenters with PET. *NeuroImage, 11*(4), 347–357.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology, 18*, 49–65.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*(1), 91–96.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology, 15*(2), 202–206.

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extra striate and frontal cortical areas by visual words and word like stimuli. *Science, 249*, 1041–1044.

Pillai, J. J., Araque, J. M., Allison, J. D., Sethuraman, S., Loring, D. W., Thiruvaiyaru, D., et al. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: Preliminary findings. *NeuroImage, 19*(3), 565–576.

Price, C. J., Wise, R. J. S., Ramsay, S., Friston, K. J., Howard, D., Patterson, K., et al. (1992). Regional response differences within the human auditory cortex when listening to words. *Neuroscience Letters, 146*, 179–182.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders, 34*(6), 479–492. doi:10.1016/s0021-9924(01)00060-0.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage, 23*, 752–763.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Noesselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature, 415*, 1026–1029.

Rüschemeyer, S.-A., Zysset, S., & Friederici, A. D. (2006). Native and non-native reading of sentences: An fMRI experiment. *NeuroImage, 31*(1), 354–365.

Sakai, K. L., Miura, K., Narafu, N., & Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex, 14*(11), 1233–1239.

Sakurai, Y., Momose, T., Iwata, M., Sudo, Y., Ohtomo, K., & Kanazawa, I. (2000). Different cortical activity in reading of Kanji words, Kana words and Kana nonwords. *Cognitive Brain Research, 9*(1), 111–115.

Savoy, R. L., O'Craven, K. M., Weisskoff, R. M., Davis, T. L., Baker, J., & Rosen, B. (1994). Exploring the temporal boundaries of fMRI: Measuring responses to very brief visual stimuli. *Society for Neuroscience Abstract, 20*, 1264.

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex, 48*(8), 991–996.
Suh, S., Yoon, H. W., Lee, S., Chung, J.-Y., Cho, Z.-H., & Park, H. (2007). Effects of syntactic complexity in L1 and L2: An fMRI study of Korean–English bilinguals. *Brain Research, 1136*, 178–189.
Swan, D., & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language, 56*(3), 334–353.
Tan, L. H., Liu, H.-L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J.-H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage, 13*(5), 836–846.
Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences, 102*(24), 8781–8785.
Tan, L. H., Spinks, J. A., Feng, C. M., Siok, W. T., Perfetti, C. A., Xiong, J., et al. (2003). Neural systems of second language reading are shaped by native language. *Human Brain Mapping, 18*(3), 158–166.
Tatsuno, Y., & Sakai, K. L. (2005). Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *Journal of Neuroscience, 25*(7), 1637–1644.
Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport, 15*, 1555–1558.
Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*, 12530–12535.
Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. T. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping, 33*, 1–13.
Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers Systems Neuroscience, 5*, 1. doi:10.3389/fnsys.2011.00001.
Van Der Meij, M., Cuetos, F., Carreiras, M., & Barber, H. A. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology, 48*, 44–54.
Van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews, 64*, 104–122.
Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483.
Wang, W. S.-Y. (1973). The Chinese language. *Scientific American, 228*, 50–60.
Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron, 37*(1), 159–170.
Weber, K., & Indefrey, P. (2009). Syntactic priming in German-English bilinguals during sentence comprehension. *NeuroImage, 46*(4), 1164–1172.
Yarkoni, T., Speer, N. K., Balota, D. A., McAvoy, M. P., & Zacks, J. M. (2008). Pictures of a thousand words: Investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *NeuroImage, 42*(2), 973–987.
Yokoyama, S., Okamoto, H., Miyamoto, T., Yoshimoto, K., Kim, J., Iwata, K., et al. (2006). Cortical activation in the processing of passive sentences in L1 and L2: An fMRI study. *NeuroImage, 30*(2), 570–579.
Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience, 14*(5), 656–661.

Author Index

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1

C

I

J

N

Subject Index

R.R. Heredia et al. (eds.), *Methods in Bilingual Reading Comprehension Research*, The Bilingual Mind and Brain Book Series 1,
DOI 10.1007/978-1-4939-2993-1

J

K

L

O

Q

R

S

X