languages

Special Issue Reprint

Multilingualism

Consequences for the Brain and Mind

Edited by
John W. Schwieter

www.mdpi.com/journal/languages

MDPI

Multilingualism: Consequences for the Brain and Mind

Multilingualism: Consequences for the Brain and Mind

Editor

John W. Schwieter

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editor
John W. Schwieter
Language Acquisition,
Multilingualism, and
Cognition Laboratory
Wilfrid Laurier University
Waterloo
Canada

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Languages* (ISSN 2226-471X) (available at: www.mdpi.com/journal/languages/special_issues/multilingualism_brain_mind).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-8039-5 (Hbk)
ISBN 978-3-0365-8038-8 (PDF)

Contents

About the Editor

John W. Schwieter

John W. Schwieter is Full Professor of Spanish, Linguistics, and Psychology, and Director of the Language Acquisition, Multilingualism, and Cognition Laboratory and Bilingualism Matters at Wilfrid Laurier University. He is also Adjunct Professor of Linguistics at McMaster University. Schwieter's research interests include cognitive and neuroscientific approaches to multilingualism and language acquisition; translation and cognition; and second language teaching and learning. His work has appeared in journals such as: *Bilingualism: Language and Cognition; Biological Psychology; Brain Structure and Function; Cerebral Cortex; Cognitive Neurodynamics; Frontiers in Psychology; Human Brain Mapping; International Journal of Psychophysiology; Language Learning; Linguistic Approaches to Bilingualism; Neuropsychologia; Psychophysiology*; among others. Some of his recent books include *The Cambridge Handbook of Working Memory and Language* (CUP, 2022) and *The Cognitive Neuroscience of Bilingualism* (CUP, 2024).

Preface to "Multilingualism: Consequences for the Brain and Mind"

For multilinguals, acquiring and processing language is similar to other cognitive skills: they are grounded in mechanisms of sensory processing and motor control (Paradis, 2019). Recent clinical and experimental research on multilingualism have introduced innovative neuroimaging measures and psychological methods that have significantly shed light on what we know (and do not know) about how multiple languages are processed, represented, and controlled in the mind/brain (Schwieter, 2019).

Since the 1990s and 2000s, a plethora of behavioral and neurological research has demonstrated that for multilinguals, all languages are active to some degree in the mind, even when only using one. Furthermore, the need for the mind to manage the ongoing competition that arises from this parallel activation has been shown to affect cognition (e.g., executive functioning) (Giovannoli et al., 2020), modify the structure and functioning of the brain (e.g., changes in the areas where language control and executive control overlap) (Costa and Sebastián-Gallés, 2014), and slow the onset or progression of cognitive and neural decline (Bialystok, 2017).

The goal of "Multilingualism: Consequences for brain and mind" is to bring together state-of-the art papers that examine the cognitive and neurological consequences of multilingualism through an exploration of how two or more languages are processed, represented, and/or controlled in one brain/mind. The included peer-reviewed papers are either theoretically or empirically oriented and present new findings, frameworks, and/or methodologies on how multilingualism affects the brain and mind.

References

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. Psychological Bulletin, 143, 233-262.

Costa, A. & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? Nature Reviews Neuroscience, 15, 336-345.

Giovannoli, J., Martella, D., Federico, F., Pirchio, S. & Casagrande, M. (2020). The impact of bilingualism on executive functions in children and adolescents: A systematic review based on the PRISMA method. Frontiers in Psychology, 11(574789).

Paradis, M. (2019). Special foreword. In J. W. Schwieter (Ed.), The handbook of the neuroscience of multilingualism (pp. xxxiii-xxxvii). Malden, MA/Oxford, England: Wiley-Blackwell.

Schwieter, J. W. (Ed.). (2019). The handbook of the neuroscience of multilingualism. Malden, MA/Oxford, England: Wiley-Blackwell.

John W. Schwieter
Editor

languages

Review

Regulation and Control: What Bimodal Bilingualism Reveals about Learning and Juggling Two Languages

Anne Therese Frederiksen * **and Judith F. Kroll**

Bilingualism, Mind and Brain Lab, University of California Irvine, Irvine, CA 92697, USA
* Correspondence: a.t.frederiksen@gmail.com

Abstract: In individuals who know more than one language, the languages are always active to some degree. This has consequences for language processing, but bilinguals rarely make mistakes in language selection. A prevailing explanation is that bilingualism is supported by strong cognitive control abilities, developed through long-term practice with managing multiple languages and spilling over into more general executive functions. However, not all bilinguals are the same, and not all contexts for bilingualism provide the same support for control and regulation abilities. This paper reviews research on hearing sign–speech bimodal bilinguals who have a unique ability to use and comprehend their two languages at the same time. We discuss the role of this research in re-examining the role of cognitive control in bilingual language regulation, focusing on how results from bimodal bilingualism research relate to recent findings emphasizing the correlation of control abilities with a bilingual's contexts of language use. Most bimodal bilingualism research has involved individuals in highly English-dominant language contexts. We offer a critical examination of how existing bimodal bilingualism findings have been interpreted, discuss the value of broadening the scope of this research and identify long-standing questions about bilingualism and L2 learning which might benefit from this perspective.

Keywords: bimodal bilingualism; language regulation; cognitive control; bilingualism; interactional context; variation in language environments

Citation: Frederiksen, Anne Therese, and Judith F. Kroll. 2022. Regulation and Control: What Bimodal Bilingualism Reveals about Learning and Juggling Two Languages. *Languages* 7: 214. https://doi.org/10.3390/languages7030214

Academic Editor: John W. Schwieter

Received: 25 April 2022
Accepted: 28 July 2022
Published: 11 August 2022

1. Introduction

Although bilingualism takes many different forms, there is evidence that knowing and using more than one language leads to both behavioral and neurocognitive changes or accommodations (Bialystok 2017; Pliatsikas 2020). Bilinguals and multilinguals experience language co-activation because all of their languages come online to some degree when even one language alone is processed (Kroll et al. 2008). The co-activation of the bilingual's two languages occurs at all levels of language processing (see Kroll et al. 2015 for a review). In this context, bilinguals must be able to regulate the intended language so that the unintended language does not intrude. How speakers acquire and use mechanisms of control to enable proficient bilingual performance is a topic that has been at the center of current research on bilingual language processing (e.g., Declerck and Koch 2022). The same mechanisms of regulation and control have been shown to be important during language learning.

In this paper, we examine the way that bilinguals who use markedly different languages, one a spoken language and the other a signed language, regulate the use of their two languages to enable skilled language use. We first review what is known about the ways that unimodal bilinguals who speak two languages engage these processes, and we then compare the evidence on unimodal bilinguals to bimodal bilinguals. We note at the outset that bimodal bilingualism is not the only form of bilingualism in which the two languages differ in significant ways. Many studies of unimodal bilinguals have examined language pairings with distinct lexical, grammatical, and phonological features. What is striking about the research on cross-language activation and its consequences for regulation and control is that the findings are largely similar across many different pairs of languages. There are modulations that result from language distance, but the overarching picture is one of greater similarity than the difference (Degani et al. 2018; see also Barac and Bialystok 2012). Critically, the role of contextual factors has been shown to be as or more important than language distance (e.g., Beatty-Martínez et al. 2020 found effects of contextual factors in Spanish–English bilinguals that are comparable to those found in Chinese–English bilinguals by Zhang et al. 2015, 2021). The environment places differential demands on speakers as a function of the opportunities to use the two languages across different contexts and distinct social networks (e.g., Green and Abutalebi 2013; Gullifer and Titone 2020). The form of bilingualism may itself vary with these demands or be independent of them.

In what follows, we review the research that has laid a foundation for understanding the regulation and control of a bilingual's two languages, and we then examine bimodal bilingualism in more detail. We view bimodal bilingualism as a phenomenon of interest in and of itself but also as a tool to investigate the constraints and plasticity associated with the mappings between language, cognition, and their neural underpinnings.

2. What Is Bimodal Bilingualism?

Bimodal bilinguals can be Deaf individuals who know a natural signed language and the spoken and/or written form of a spoken language (depending on their language experience and when they became deaf), or hearing individuals who know a spoken language and a signed language (see Berent 2004 for an overview of the forms that bimodal bilingualism can take).[1] Here, we restrict our focus to hearing bimodal bilinguals (see Morford et al. 2011, 2014, 2019 for studies of Deaf sign-print bilinguals). Crucially, bimodal bilinguals' languages are in two different modalities.[2] This offers an opportunity to explore the role of overlapping forms on different aspects of bilingual linguistic and cognitive behavior (e.g., Morford and Kroll 2021). For example, what does language co-activation look like when there is no shared phonology? Does the way that bilinguals regulate their languages differ, quantitatively or qualitatively, when the two languages do not rely on the same articulatory and perceptual systems, and what are the consequences of this for non-linguistic cognitive processing?

In many ways, bimodal bilinguals are similar to unimodal bilinguals. Children of Deaf Adults (CODAs), that is, hearing individuals who grew up in deaf families, in particular, have much in common with those unimodal bilinguals who use their two languages in different contexts and cultures (e.g., Heritage speakers of a home language who speak the language of the community at school and work but the use the home language with family).[3] Natural sign languages are phonologically, lexically, and grammatically different from the spoken languages surrounding them (Sandler and Lillo-Martin 2006). That means that bimodal bilinguals learn two distinct languages, and many also come to know two different cultures (Singleton and Tittle 2000). Similarly, although the two languages are in different modalities, they appear to be supported largely by the same neural systems, much as is observed for unimodal bilinguals. This is the case when comparing native sign language and spoken language in different individuals (see MacSweeney et al. 2008; and Emmorey 2002, for reviews) and in bimodal bilinguals (Emmorey et al. 2014; see Emmorey et al. 2016, for an overview).

Despite many similarities between unimodal and bimodal bilinguals, a crucial difference is that bimodal bilinguals can use and comprehend both of their languages at the same time because they can sign while they speak or speak while they sign. It is important to note that this does not mean that bimodal bilinguals regularly converse fluently and grammatically in both languages at the same time; rather, bimodal bilinguals often accompany some spoken words in an utterance with signs, or vice versa. Nevertheless, this ability means that bimodal bilinguals are not subject to the constraints that unimodal bilinguals face of being physiologically unable to produce their two languages at the same time because their languages are perceived and produced using different production and perceptual systems. In contrast, spoken languages are primarily produced orally and perceived aurally, whereas signed languages are produced manually and with the body and perceived visually.[4] While language processing in both modalities largely makes use of the same neural networks, some differences have been reported in patterns of brain activity responsible for modality-specific language components, e.g., listening vs. watching, as well as differences in motor regions moving hands vs. vocal tract (Corina et al. 2013).

Using both languages at the same time is not something that is merely technically possible. Unlike unimodal bilinguals who know two spoken languages and who often switch languages sequentially, a phenomenon known as code-switching, bimodal bilinguals tend to code-blend, that is, to use two languages at the same time. Such language blending is a normal and natural process for bimodal bilinguals (Emmorey et al. 2008a), and especially CODAs (Bishop 2010; Bishop and Hicks 2005; Zeshan and Panda 2018; Kuntze 2000). [5] Code-blending is not restricted to lexical items (e.g., signing *HOME* in ASL while saying *home* in English). A study by Pyers and Emmorey (2008) showed that bimodal bilinguals use facial expressions that convey grammatical information in ASL while speaking in English. In a communicative task in English, bimodal bilinguals produced more ASL-like facial expressions (i.e., eyebrow raises to indicate if-statements and furrowed brows to indicate wh-questions) than did non-signers, despite the fact that their interlocutor did not know ASL. Language blending is not unique to proficient bimodal bilinguals. There is a growing body of evidence from studies of adult sign language learners, suggesting that signs or sign language-based distinctions begin appearing in the co-speech gestures that accompany the spoken language from very early stages of language acquisition (Casey et al. 2012; Gu et al. 2019; Weisberg et al. 2020; Frederiksen 2021).

Bimodal bilinguals' ability to use signed and spoken language at the same time has consequences on how they manage their languages, which in turn affects how bilingualism modulates cognitive control abilities in this population. Understanding how two languages are managed under these conditions is not only informative in its own right but also offers a way of establishing how different aspects of the bilingual experience are related to cognitive abilities and processes.

Language Co-Activation in Bimodal Bilinguals

While bimodal bilingualism presents some unique opportunities for using the two languages together, it also shares a fundamental feature of unimodal bilingualism: When one language is presented alone, the other language is activated to some degree. An extensive body of research on unimodal bilingualism has established the parallel activation of the two languages when unimodal bilinguals listen to speech, when they read, and when they plan speech in each of the two languages. Reviews of the evidence on the nonselective activation of bilinguals' two languages are widely available (e.g., Dijkstra and Van Heuven 2018; Kroll et al. 2006, 2015; Kroll and Navarro-Torres 2018). An important observation in this research on unimodal bilingualism is that bilingualism with spoken languages does not depend on shared properties of the written or spoken form of the language; even languages that do not share the same written script or phonological form give rise to cross-language activation (e.g., Hoshino and Kroll 2008; Thierry and Wu 2007; but see also Costa et al. 2017 for an account that assumes mapping between L1 and L2 during learning instead of cross-language activation during processing). These findings are important because they suggest that we might also see cross-language activation in bimodal bilinguals and that bimodal bilinguals, like unimodal bilinguals, may need to engage regulatory mechanisms that draw on domain-general cognition.

Several studies have shown that although bimodal bilinguals' languages make use of two different modalities, these bilinguals experience the same kind of language co-activation as unimodal bilinguals do (Shook and Marian 2012; Giezen et al. 2015; Giezen and Emmorey 2016; Villameriel et al. 2016; Williams and Newman 2016). Morford et al. (2014) investigated if hearing English speakers who were proficient signers of American

Sign Language (ASL) would activate ASL while performing a lexical decision task in written English. The English-ASL bilinguals were presented with English word pairs that were semantically related (heart-brain) or unrelated (baby-lion), and their task was to indicate with a button press whether the words were related in meaning or not. Unbeknownst to the participants, half of the words in each semantic condition had ASL translation equivalents that were phonologically related; that is, they shared some combination of hand shape, movement, or location in signing space (i.e., shared hand shape and movement in the semantically unrelated, but phonologically related ASL word pair MOVIE-PAPER in Figure 1).

Figure 1. ASL signs for MOVIE (**left**) and PAPER (**right**). Reprinted with permission from Morford et al. (2011). Copyright 2011 Elsevier.

Despite there being no phonological overlap between ASL and English because the languages are produced in different modalities, Morford and colleagues found that the English-ASL bilinguals co-activated ASL while performing the task in English. Specifically, participants were slower to indicate that semantically unrelated English words were not similar in meaning when the ASL translation equivalents were related in phonology, compared to when semantically unrelated words were also phonologically unrelated in ASL.

These findings highlight that language co-activation is not restricted to language pairs with shared form elements and suggest that it is instead a feature of the bilingual language system. Bimodal bilinguals, however, are uninhibited by the factors that restrict unimodal bilinguals to using one language at a time, and code-blending may thus reflect how the bilingual language system handles both languages being active in the absence of any articulatory constraints. The next sections present evidence on the bilingual adaptations resulting from the need to regulate language co-activation and discuss how the fact that bimodal bilinguals may have a much-reduced need to regulate this offers a way to investigate the role of language regulation in the development of bilingual cognitive abilities.

3. Control and Regulation in Bilinguals

The traditional approach to bilingualism research, perhaps especially in studies taking a cognitive perspective on bilingualism, has aimed to identify group differences in bilinguals compared to monolinguals across different cognitive dimensions. Understanding that the bilingual is not two monolinguals in one (Grosjean 1989), research has sought to characterize the nature of the difference between monolinguals and bilinguals, not only in terms of language use but also in the cognitive and neural consequences that result from the use of more than one language. An obvious difference between monolinguals and bilinguals is that bilinguals have two languages and must always select which one to use, which creates different demands on language monitoring in these two groups of language users, especially because of the demands in the bilinguals on the mechanisms responsible for language selection. In this way, cross-language activation, and the need to regulate it, results in different every-day cognitive demands for bilinguals compared to monolinguals and different bilingual adaptations to meet these demands (e.g., Bialystok 2017). While there is considerable ongoing controversy about the scope of the consequences and adaptations produced by bilingual experience (see, for example, Antoniou 2019; van den Noort et al. 2019; and Grundy 2020), the fact that bilinguals' two languages are always active is uncontroversial. Yet, the nature of cross-language activation and its regulation is also complex as research uncovers more details about how variation in the bilingual experience modulates aspects of bilingual language and cognition. Although variation in factors such as the age of acquisition and proficiency level has long been of interest, other types of variation have received limited attention until recently. If we consider how people around the world come to learn more than one language, it becomes clear that there are many different circumstances that lead to bilingualism ((Baum and Titone 2014; Fricke et al. 2019), just as there are different environments in which bilinguals use their languages when they have already been acquired (Luk and Bialystok 2013; Surrain and Luk 2019)). Bimodal vs. unimodal bilingualism represents an obvious instantiation of variability in bilingual experiences. As discussed, bimodal bilingualism is in some respects drastically different from unimodal bilingualism while also sharing many characteristics with it. This offers a unique opportunity to explore how different variable aspects of bilingualism influence bilingual cognitive adaptations. In recent years, a growing interest in differences among bilingual populations has begun to uncover significant variation in what language co-activation looks like based on individual circumstances and in how bilinguals' language regulation needs and abilities change in response (Green and Abutalebi 2013; Giezen et al. 2015; Zirnstein et al. 2018; Treffers-Daller 2019; Beatty-Martínez et al. 2020). From this perspective, the properties of bimodal bilingualism represent variation not only in the modality difference compared to unimodal bilingualism but also among bimodal bilinguals in different circumstances.

In what follows, we review foundational evidence about language regulation and control from studies of unimodal bilingualism and then discuss how accounting for variation in bilingual experiences, including in modality, reveals systematicity in previous findings.

3.1. Cognitive Control Adaptations

Cognitive control refers to domain-general abilities that help with regulating mental activity in order to resolve information conflicts (Braver 2012; Miyake and Friedman 2012). While cognitive control is a domain-general ability, it has important uses in language processing, for example, in resolving competing interpretations of sentence or utterance meaning (Novick et al. 2005; Nozari et al. 2016; Hsu et al. 2017). Similarly, dynamic engagement of cognitive control has been shown to facilitate efficient language processing (Hsu and Novick 2016; Hsu et al. 2021; Ovans et al. 2022).

Many studies have shown that bilingualism can lead to behavioral changes, not only in language use but also in the ways that cognitive and neural resources are recruited to enable fluent language use. Proficient bilinguals often outperform monolinguals on a variety of tasks measuring executive functions, including cognitive control (Bialystok et al. 2008; Bialystok and Craik 2022). This is assumed to be due to bilinguals' long-term experience with managing two languages. Specifically, the extensive experience in engaging in mental behaviors involved in controlling one's languages is assumed to confer benefits to other non-verbal domains (Bialystok 2009). Bilingualism also leads to changes to neural patterns in the brain, as well as changes to anatomical structure, including increased grey matter and cortical thickness in the temporo-parietal cortex, as well as enhanced white matter integrity, e.g., in the inferior fronto-occipital fasciculus (Li et al. 2014; Abutalebi and Green 2016), some of which have been linked to observed behavioral changes to cognitive control (Della

Rosa et al. 2013). Such neural and behavioral changes can have long-term consequences, as evidenced, for example, by research showing that bilingualism acts as a protective factor against dementia, such that bilinguals experience a later onset of symptoms relative to monolinguals (Bialystok et al. 2007).

Studies in the past decade have begun to identify the ways in which variability in bilinguals' language and learning experiences may affect cognitive processes. According to the adaptive control hypothesis (Green and Abutalebi 2013), the interactional contexts in which bilinguals live and use their two languages are predicted to differentially shape control processes. For example, this hypothesis holds that bilinguals who generally use their languages separately in different environments are expected to adapt their control processes to their surroundings differently than bilinguals who generally use both languages in many contexts but with different speakers, and both of these bilingual groups are expected to differ from those regularly code-switch, that is, who use both languages in many contexts and with the same speakers (Green and Abutalebi 2013).

Investigating bimodal bilinguals has been particularly fruitful for disentangling the possible contributing factors underlying changes to cognitive control in bilinguals. Bimodal bilinguals' ability to use both languages at the same time suggests that they have a reduced need to fully inhibit one language while they are using the other. If processes such as cognitive control develop in response to the need to manage one's languages, then bimodal bilinguals may not necessarily show the same kinds of changes to cognitive control abilities that unimodal bilinguals do. Indeed, Emmorey et al. (2008b) found that adult unimodal but not bimodal bilinguals outperformed monolinguals on the speed of decision on a set of tasks measuring cognitive control. They argue that this provides evidence that the enhanced cognitive control often observed in bilinguals can be traced to the experience of managing two languages in the same modality rather than simply managing the representation of two languages. Similarly, Olulade et al. (2016) measured differences in grey matter volume in frontal brain areas associated with executive functions and found that while unimodal bilinguals had larger grey matter volume compared to monolinguals, bimodal bilinguals did not (but see Zou et al. 2012b). What these studies suggest is that differences in cognitive control do not necessarily happen as a result of bilingualism, per se. In this way, research on bimodal bilinguals offers the chance to establish causal links between bilingual experiences and outcomes. We further note that the interest in bimodal bilingualism has the potential to uniquely contribute to the paradigm shift in the broader field of bilingualism over the past decade, away from the traditional, relatively narrow interest in how bilingualism, broadly construed, differs from monolingualism and towards an interest also in individual differences between bilinguals as well as differences between bilingual groups whose language experiences differ in ways that go beyond differences in proficiency.

Cognitive Control and Language Learning

In proficient bilinguals, cognitive control abilities and the ability to manage multiple languages are assumed to develop alongside language proficiency. Results from studies with babies and toddlers have suggested that bilingualism may lead to different cognitive development with respect to memory generalization and improved cognitive control abilities even before the age at which most children master basic two-word utterances (Brito and Barr 2012) or even start to produce words (Kovács and Mehler 2009; Ferjan Ramírez et al. 2017). More uncertain is what role cognitive control plays for developing bilinguals, that is, learners of a second language (L2) after childhood when the first language (L1) has been solidly established, with robust links between world and language, such as between concepts and lexical items, and what kind of relationship exists between cognitive control, language regulation and L2 proficiency and use at different stages of acquisition.

For decades, second language learning research has investigated what enables adult learners to become proficient in an L2, investigating both external and internal factors (e.g., learning environment vs. motivation, Luque and Morgan-Short 2021). An existing assumption is that the skills and processes that enable the proficient bilingual to function in their two languages are also related to achieving proficiency in a second language in the first place (Michael and Gollan 2005). One possibility is that individual differences in cognitive control abilities in monolinguals predict their success with L2 learning. In order to assess whether increased cognitive control helps initial language learning, studies have compared monolinguals and bilinguals on tasks related to novel word learning (Cenoz 2003; Kaushanskaya and Marian 2009). A study by Bartolotti and Marian (2012) that specifically investigated word retrieval in a newly-learned language found that resolution of between-language competition occurred earlier in bilinguals than in monolinguals. One interpretation of these results is that individuals with better cognitive control abilities make more efficient L2 learners. Another possibility is that cognitive control changes as L2 proficiency and/or the need to regulate activation of the L1 vs. L2 increases. Luque and Morgan-Short (2021) examined the relationship between L2 proficiency and different cognitive measures, including cognitive control. They administered a range of cognitive and L2 proficiency tasks to test the correlation between cognitive control and L2 proficiency. Reactive and proactive cognitive control was measured using the AX Continuous Performance Task (AX-CPT), in which participants' reaction times are measured to five-letter strings in which the first letter is the cue and the final letter is the probe (see Braver 2012 for a discussion of how this task provides a framework for analyzing cognitive control). The study found a negative correlation between L2 proficiency and overall processing time and reactive control, with the more proficient L2 users reacting faster overall as well as on trials measuring reactive control. Across studies, however, results have been mixed, with a relationship between cognitive control and second language learning occurring only sometimes and varying as a function of proficiency and task (Luk et al. 2011; Bak et al. 2014; Misra et al. 2012; Vega-Mendoza et al. 2015; Xie 2018).

In this context, studying bimodal bilingualism might offer the means to disentangle the relationship between L2 learning and cognitive control. As discussed above, findings from adults who learned a sign language and a spoken language in childhood have suggested that their cognitive control abilities may not show the same enhancements as in unimodal bilinguals, and this difference is explained by the fact that bimodal bilinguals are not required to regulate their two languages as carefully as unimodal bilinguals. If proficient bilinguals do not require the kind of cognitive control abilities that unimodal bilinguals do to function in their two languages, then learning a sign language as a second language may similarly result in little or no changes to cognitive control, nor should individual differences in cognitive control predict success in learning a signed L2. To date, no studies that we are aware of have examined the effect of cognitive control on the acquisition of a signed language, nor whether knowing a second sign language facilitates learning a new spoken language, as in the study by Bartolotti and Marian (2012). Interestingly, however, the results from two studies examining the learning of a sign language as a second language in adulthood after a childhood of only spoken language exposure suggest that cognitive functions, in fact, may change in adult sign language learners. In a longitudinal study of ASL interpreting students, Macnamara and Conway (2014) found improvement in emerging bimodal bilinguals' cognitive abilities, suggesting that, under the right circumstances, bimodal bilingualism can, in fact, have similar effects as unimodal bilingualism, although it remains unclear whether the level of improvement observed in the study would be enough to differentiate the bimodal bilinguals from monolinguals, as well as how much the cognitive changes should be attributed to specifically to interpreter training as opposed L2 learning in and of itself. Similarly, changes to cognition may appear in bimodal bilinguals who learned sign language after childhood. Proficient bilinguals who have learned both speaking and signing in childhood do not show increased grey matter volume in brain areas such as the dorsolateral prefrontal and parietal cortices that are implicated in executive functions such as attention shifting, inhibition, and conflict detection and resolution (Derrfuss et al. 2004; Cole and Schneider 2007) and that are frequently seen to have greater volume in proficient unimodal bilinguals (Olulade et al. 2016). However, Chinese speakers who learned Chinese Sign Language as an L2 in adulthood showed structural differences from monolinguals in specific brain areas implicated in language switching and control (Zou et al. 2012b, see also Li et al. 2015).

In sum, many questions remain for future research about the relationship between cognitive control and L2 learning for bimodal bilinguals and how this may differ from proficient bimodal bilinguals. For proficient bilinguals, there is behavioral and neurocognitive evidence to support the conclusion that the changes often observed in unimodal bilinguals' cognitive control abilities result from the experience of needing to regulate two languages in the same modality, and as bimodal bilinguals need less regulation, they may not show similar cognitive enhancement. However, the evidence from signed L2 learning, as well as the research focused on bilingual variation in unimodal bilinguals, makes it increasingly clear that differing patterns of bilingual language use may contribute to the variation observed in bilingual cognition, including cognitive control patterns (Treffers-Daller 2019; Green and Abutalebi 2013). The next section discusses questions related specifically to language regulation in bilinguals, how cognitive control is modulated by variation in how much language regulation a bilingual engages in as a function of language environment and interactional contexts (Zirnstein et al. 2018; Beatty-Martínez et al. 2020), and how research on bimodal bilinguals may shed light on these issues.

3.2. Language Regulation

Language regulation, a term first used by Zirnstein et al. (2018), is closely related to but not fully overlapping with cognitive control. The term describes bilinguals' use of different executive functions to shift the activation state of the L1 and L2, such that the demands of the environment can be met (Kroll et al. 2022). Bilinguals must learn flexibility in how they inhibit and disinhibit their two languages as a function of a range of factors, including immersion, language status, individual-level proficiency, conversational partner, and other types of contexts. Language regulation is also important for learning an additional language in the first place. Findings from studies of adult language learners have suggested that successfully learning a second language partially depends on the ability to regulate the first (or native) language. While language regulation likely involves cognitive control, recent findings suggest that the two processes do not fully overlap (Kang et al. 2020; Zirnstein et al. 2018). This section discusses evidence of language regulation stemming from language switching as well as language learning and immersion and also discusses the relationship between cognitive control, language regulation, and a bilingual's interactional contexts, emphasizing what research on bimodal bilingualism contributes to current understandings of these phenomena.

3.2.1. Language Switching as Evidence of Language Control and Regulation

Much of what is known about language regulation comes from studies of language switching. The critical insights from this work are that although bilinguals can communicate in both their languages and many proficient bilinguals code-switch frequently, it can be costly to switch between languages. Especially, the type of cued or forced switching employed in typical experimental paradigms results in high switching costs (see Zhu et al. 2022). A foundational study by Meuter and Allport (1999) used a cued language switching paradigm in a number naming task and showed that naming was overall faster in L1 as compared to L2. Crucially, on trials where bilinguals had to switch language (switch trials), response latencies were longer than on trials where the same language was used in the preceding trial (non-switch trials) compared within the same language. This difference is termed the switching cost. Meuter and Allport moreover found that switch costs were asymmetrical between the L1 and L2, with the L2 to L1 switch producing longer response latencies than L1 to L2 switches, and, importantly, L1 naming becoming slower than L2 naming on switch trials, suggesting that it is more effortful to regulate the L1 compared to the L2. Switch costs have been found both in production (e.g., Costa and Santesteban 2004; Meuter and Allport 1999) and comprehension (e.g., Grainger and Beauvillain 1987; Thomas and Allport 2000). Based on these insights, the field has long assumed that language switching is costly and that the cost associated with language switching results from the effort required to regulate the relative activation of the two languages, specifically as related to overcoming the inhibition from the preceding trial (e.g., Kroll et al. 2008; Philipp and Koch 2009).

Cued language switching has also been investigated for bimodal bilinguals. As research has shown that both languages are similarly activated in bimodal bilinguals as in unimodal bilinguals (e.g., Shook and Marian 2012), bimodal bilinguals should be similar to unimodal bilinguals in needing to rely on language regulation in order to switch between languages. Dias et al. (2017) investigated language switch costs in bimodal Spanish-Spanish Sign Language (LSE) bilinguals and found that they performed similarly to unimodal bilinguals in exhibiting greater switch costs when switching from the weaker (L2) to the stronger (L1) language. However, in a study comparing language switch costs in German–English-German Sign Language (DGS) trilinguals, Kaufmann et al. (2018) found that although there was a cost to switching between German and DGS, the costs were higher in unimodal German–English contexts than bimodal German-DGS contexts in intermediate sign language learners. One suggested explanation is that the difference is due to the fact that it is the lexical item that has to be inhibited in the unimodal context, but in the bimodal context, only the articulators need to be inhibited—and that lexical inhibition is costlier.[6] Casey et al. (2012) similarly propose that language control could be uniquely challenged when the two languages are in different modalities. This raises the question of whether language regulation is, in fact, the same unified process in the contexts of unimodal vs. bimodal bilinguals.

3.2.2. The Locus of Language Switch Costs

Emmorey et al. (2008a) conducted one of the first studies to highlight the unique capacity of bimodal bilinguals to use both of their languages simultaneously. This work demonstrated that bimodal bilinguals (CODAs) tend to code-blend rather than code-switch the way unimodals do. Accordingly, research on bimodal bilinguals has provided a way to tease apart which aspect of switching is responsible for the cognitive cost. Several processes are involved in switching from one language to the other. Simplistically, for an individual to switch between languages, they first must recognize that a switch has to be executed, then they must inhibit ("turn off") one language and activate or increase activation ("turn on") of the other language. While there is broad agreement that language switching is costly, it is less clear which components of the switch are responsible for the cost. Some theories have emphasized activation as being costly, while others have emphasized inhibition (see Declerck and Philipp 2015, for review). In unimodal bilinguals, the switching on and off happens at the same time; that is, in order to produce a language switch, the unimodal bilingual must simultaneously turn off one language and turn on the other. In bimodal bilinguals, these two processes can be separated by taking advantage of bimodal bilinguals' ability to use both languages at the same time. Specifically, studies of switch costs in bimodal bilinguals include dual response trials, that is, trials in which items are named in both languages at the same time by signing and speaking (see Kaufmann et al. 2018; and Emmorey et al. 2020). When bimodal bilinguals name in a single language after naming in both languages simultaneously, this isolates the process of turning a language off. Conversely, naming in both languages simultaneously after naming in only one of the languages isolates the process of turning a language on. In this way, preceding and following dual response trials with single language trials allows for an investigation of whether costs for turning a language on (going from single language trial to dual language trial) differ from the costs for turning a language off (going from dual language response to single language).

A study by Emmorey et al. (2020) compared response latencies in a task in which bimodal bilinguals named pictures in English, ASL, and both languages simultaneously. Their results showed no cost to bimodal bilinguals for switching a language on, only for switching one-off. These findings about behavior have been corroborated by neuro-imaging research. Recording neural activity using magnetoencephalography (MEG) while bimodal bilinguals performed a picture naming task, Blanco-Elorrieta et al. (2018) showed that the brain regions involved in cognitive control are not involved when bimodal bilinguals engage in a language, only when they inhibit one. Studying bimodal bilinguals thus offers a way to test hypotheses about links between different bilingual language behaviors and cognitive processes, which can be assumed but not tested from correlational studies of unimodal bilinguals.

3.2.3. What Code-Blending Reveals about the Bilingual Language System

The studies discussed above—showing that inhibition and not activation of a language is what is costly—help explain why bimodal bilinguals code-blend when intuition suggests that using two languages at the same time should be more effortful than using one. If adding a language is not costly, however, using two languages at the same time may, in fact, not be more effortful than using only one. Emmorey et al. (2008a) demonstrated that bimodal bilinguals frequently express the same lexical meaning in their two languages simultaneously instead of suppressing one of the languages. Along the lines of the discussion in the previous section, this suggests that it may be less costly to select two lexical representations compared to selecting one and suppressing the other (p. 57). Emmorey et al. (2012) provided empirical evidence for this suggestion by investigating the processing of code-blends (see also Giezen and Emmorey 2016; Emmorey et al. 2020). They showed that production and comprehension of simultaneous ASL and English (i.e., signing CUP while saying 'cup') do not appear to be more difficult than using one language at a time. Specifically, in production, ASL signs were produced with the same speed in code-blends compared to ASL alone. English words were slower in code-blends, but this was attributed to the desire to synchronize lexical onsets in the two languages—with ASL being overall slower (even in monolinguals)—and thus not attributable to processing costs. In comprehension, code-blends facilitated access to both languages. The results suggest that the dual language mode is not more costly than the single language mode, a remarkable finding given studies of non-language dual response tasks show worse performance in the dual compared to a single task mode (see Pashler 1994, for review). While simultaneously naming in two languages at the same time is less effortful than naming in one language while turning the other language off, the relative cost of inhibiting one language compared to dual lexical production may depend on whether it is the dominant or non-dominant language that is being inhibited (Blanco-Elorrieta et al. 2018). Thus, it is critical that future work investigate bimodal bilingual language behaviors in groups with varying language experiences and proficiency. Further, recent work on unimodal bilinguals has suggested that freely mixing languages can, in fact, be less effortful than staying in a single language (Zhu et al. 2022). Future research should investigate whether bimodal bilingual code-blending can be considered a form of free language mixing, absent any articulatory constraints, or whether the behaviors of bimodal and unimodal bilinguals are the result of qualitatively different systems for language regulation.

3.2.4. Language Regulation and L2 on L1 Effects

The traditional assumption has been that in sequential L2 learning, and especially L2 learning in adulthood after the L1 is well-established, the L1 has a privileged status. As discussed above, this assumption influenced the understanding of cognitive control in bilinguals and, for a long time, also meant that language regulation for adult L2 learners was understood to consist mostly or solely of regulating the L1, and especially limiting transfer from L1 to L2. More recently, studies have shown a preponderance of bidirectional language influences, and specifically that learning a second language has effects on the first language as well. This suggests that bidirectional language effects start emerging from the very beginning of second language learning, especially at the lexical level (see reviews in Kroll et al. 2021, 2022; see also Bice and Kroll 2015; Bogulski et al. 2019; Brice et al. 2021). What such findings show is that bilingualism, even at the earliest stages, changes the language system away from the monolingual system. In the new bilingual system, language regulation becomes a crucial component of successful communication, and it is possible that language regulation is central not only to regulating what has been acquired in the L2 but to the process of learning an L2 in the first place. Acquiring L2 proficiency may thus depend in part on how open the existing native language system is to influences from the L2, possibly recapitulating the openness of the language system in infants exposed to more than one language (Ferjan Ramírez et al. 2017; Petitto et al. 2012).

Given that proficient bimodal bilinguals experience cross-language activation similar to that of unimodal bilinguals, we might expect L2-on-L1 effects to be relatively independent of language modality, and, in fact, studies have observed L2 on L1 effects in L2 sign language learners. Casey et al. (2012) conducted a large-scale survey of novice ASL students and showed that over 60% of the learners self-reported sometimes using ASL signs when speaking. In the same study, the results of an experiment investigating narrative retellings, novice English L1-ASL L2 learners showed effects of learning to sign when they use English. Specifically, in face-to-face communication, they use co-speech gestures more (see also Weisberg et al. 2020; Emmorey et al. 2008a) than developing unimodal bilinguals, and as many as 25% of the learners produced one or more ASL signs alongside speech when retelling a story in English. This suggests that bimodal bilinguals experience lexical intrusions from ASL despite speaking to a non-signer and that this likely happens from the earliest stages of acquisition. Frederiksen (2021) similarly found that intermediate ASL learners accompanied English verbs of object placement (*put*, *place*) with co-speech gestures that were different from those used by non-signers. Specifically, the signers used handshapes that reflected the shape and orientation properties of the object being placed, and they acted so in a manner that is consistent with ASL but not with English. These findings suggest that not only the life-long exposure to and use of a sign language but also short-term exposure can affect the L1 by increasing and qualitatively changing gesture use as well as resulting in simultaneous use of the two languages. While it remains to be seen how patterns such as these evolve as proficiency increases, there is evidence that the functional brain networks supporting spoken L1 production look different in proficient L2 signers compared to monolinguals, suggesting that learning a signed L2 can have profound effects on language processing in the brain (Zou et al. 2012a) Nevertheless it is not yet clear how these patterns might inform the understanding of how efficient language regulation is acquired and how the ability to tolerate changes to the L1 affect the language learning process across modalities.

The relationship between L2 learning and L1 processing can perhaps be seen most clearly in studies investigating what happens to L1 access after using the L2 for shorter or longer periods of time. The asymmetrical switching costs discussed above showed that the L1 could be inhibited immediately after L2 use (e.g., Meuter and Allport 1999). Misra et al. (2012) found evidence that translation equivalents were inhibited in L1 after L2 use and that this inhibition lasted for longer than just a few trials (it lasted across at least two blocks, see also Van Assche et al. 2013). On a scale of months, a study by Linck et al. (2009) found that learners immersed in an L2 environment had temporarily reduced access to L1 while immersed. Moreover, the immersed learners showed an unexpected insensitivity to L1 lexical interference in a translation recognition task. This effect was found both during immersion and after returning to the L1 environment. The authors argue that these findings support an important role in the inhibition of the L1 in successful L2 acquisition. Other studies have reported similar findings (e.g., Baus et al. 2013; Brice et al. 2021; and see Kroll et al. 2018, for a review of the findings on language immersion). To date, inhibition effects have been investigated in bimodal bilinguals only on the shortest timescale, that is, in the context of cued language switching, where studies have found asymmetrical switch costs, suggesting that the spoken language is more strongly inhibited than the signed language. To date, no studies have examined the long-term effects or the effects of immersion on language access in bimodal bilinguals. As such, the answers are not clear to questions such as whether immersion in ASL reduces bimodal bilinguals' access to English or whether their ability to use both languages simultaneously creates an opportunity for maintaining the relative activation levels of the two languages, regardless of language environment.

4. Language Environment and the Relationship between Language Regulation, Cognitive Control, and Language Processing

In the past decade, research has begun to explore the effects on the cognition of different bilingual experiences. Evidence is emerging to suggest that fine-grained variation in the contexts of language use affects the relationship between different linguistic and cognitive processes. Different language environments, both L1 and L2, pose different demands on language users, and recent years have seen a variety of metrics proposed to capture variation in bilingual's language use and interactional contexts. Green and Abutalebi (2013) distinguish between bilinguals based on whether they use their languages mostly in separated situations (single-language contexts), consistently with different interlocutors (dual-language contexts), or whether they mix languages with the same interlocutor (dense code-switching contexts). Similarly, Beatty-Martínez and Titone (2021) propose different bilingual phenotypes based on patterns of language use and regulation in competitive vs. cooperative contexts. The entropy measure developed by Gullifer et al. (2018) considers the diversity of individuals' social and language networks, with lower language entropy scores reflecting less diversity and thus more predictable contexts for the use of each language, and higher language entropy scores indicating the use multiple languages in a variety of contexts. The growing understanding of the importance of accounting for bilinguals' contexts of language use in shaping the relationship between language and cognition provides a context for a better understanding of the theoretical importance of the existing evidence from studies of bimodal bilinguals and the questions we still need to ask. At the same time, this understanding highlights the unique contribution that studies of bimodal bilingualism can make to the field of bilingualism at large.

Studies have shown that various types of contexts affect bilingual language regulation, which in turn shapes control processes and their interaction with language processing (e.g., Pot et al. 2019; Gullifer and Titone 2020; Ooi et al. 2018). Gullifer et al. (2018) conducted a neuroimaging study of highly proficient French–English bilinguals with similar results. This study found that the bilinguals with lower entropy scores, indicating more predictability in when to use French vs. English, relied more on reactive cognitive control compared to bilinguals with high entropy scores, and low entropy bilinguals also showed less connectivity between different brain areas that have been found in previous research to be implicated in monitoring and language switching. Beatty-Martínez et al. (2020) investigated the relationship between lexical access and cognitive control in bilinguals who were alike in language proficiency but differed in their interactional contexts and the patterns of language regulation necessitated by these contexts. This study found that among Spanish–English bilinguals in compartmentalized language contexts where they generally only use one language per context, greater reliance on reactive cognitive control in the AX-CPT task correlated with better picture naming accuracy in L1 and L2. Among Spanish–English bilinguals who are immersed in the L2 (English) in environments where only some community members use both languages, there was overall greater reliance on proactive cognitive control processes and a positive correlation between more proactive control and higher picture naming accuracy in Spanish. For bilinguals in integrated language contexts, who can freely mix their two languages (code-switch) because they are in communities with speakers who are proficient in both languages and where code-switching is culturally acceptable, the pattern was in between those of the two other groups (Beatty-Martínez et al. 2020).[7]

The results of such studies suggest that it is long-term experience with a specific type of language regulation that shapes cognitive control abilities and their relationship to language processing. However, as is clear from studies of language immersion, bilinguals' regulation needs can change when the language environment changes. The studies discussed above show that L1 access can change as a result of the type of language regulation required in contexts of immersion, and an important question is how such experiences affect cognitive control. Zhang et al. (2015) conducted a study on the effect of language switching training and showed that Chinese–English bilinguals living in China improved their performance on proactive control in the AX-CPT task after several days of being trained in switching between English and Chinese. Using the same paradigm with Chinese–English bilinguals in the U.S., however, revealed no proactive cognitive control improvements (Zhang et al. 2021), but crucially, the English-immersed bilinguals had much higher initial proactive control scores, suggesting that the experience of immersion and the resulting increased demands on language regulation had already led to the cognitive control benefits that the bilinguals in China obtained after language switching training. While more research is needed, the results of this study suggest that bilinguals who have a high need to inhibit either language in unpredictable contexts may be expected to most consistently show increased cognitive control abilities. Conversely, bilinguals with lower inhibitory needs or in highly predictable environments may be expected to experience fewer changes to their cognitive control abilities. From this perspective, bilinguals with experiences that look drastically different on the surface may, in fact, show similar adaptations. For example, bilinguals who do not need to fully inhibit one language, such as bimodal bilinguals, may look like those who can switch languages relatively freely, such as bilinguals in integrated or dense code-switching contexts.

From the perspective of understanding control and regulation abilities as a function of the demands of the language environment, the reason why bimodal bilinguals do not show cognitive advantages may thus not necessarily be due to the lack of perceptual or production overlap between their languages. Instead, it may result in part from bimodal bilinguals rarely needing to decide which language to use in any context. In an English-speaking context, a bimodal bilingual might supply signs for lexical items while speaking without disrupting communication because such signs are likely to be perceived as co-speech gestures by a non-signing interlocutor. In fact, Pyers and Emmorey (2008) showed that CODAs do not fully suppress ASL even when their productions (i.e., grammatical facial expressions such as furrowed eyebrows) could be construed by the interlocutor as conveying negative emotions or attitudes in the context of speaking in English.

As discussed previously, the experience of learning a signed second language may lead to changes in cognitive abilities in some cases (Macnamara and Conway 2014). It is also clear that cognitive control plays a similar role for bimodal and unimodal bilinguals during language processing. Giezen et al. (2015) investigated whether bimodal bilinguals showed a correlation between non-linguistic conflict resolution (measured by the Stroop effect) and efficiency of managing cross-language activation, similar to unimodal bilinguals (Blumenfeld and Marian 2013). Giezen and colleagues found that bimodal bilinguals who had smaller Stroop effects (showing better inhibitory control) also experienced less competition from ASL in recognizing English words. This result suggests that although bimodal bilinguals look like monolinguals in their cognitive control abilities, they may *use* their abilities as unimodal bilinguals do. This raises the possibility that the regulation and control patterns of some bimodal bilinguals may look more like unimodal bilinguals than others. As the studies of unimodal bilinguals by Zhang and colleagues showed, bilinguals' language environment affects their cognitive control abilities, but what this relationship looks like for bimodal bilinguals is as of yet unknown. The majority of the evidence from bimodal bilinguals comes from CODAs and proficient L2 learners in mostly English-speaking contexts, but experiences may vary between bimodal bilinguals. Some bimodal bilinguals may be switching regularly between ASL and English contexts. Others may be using ASL in the Deaf community, where it may not be culturally appropriate to voice in English. Such contexts may involve regularly switching languages and an increased need to inhibit the non-selected language and may consequently result in different cognitive control adaptations. It is also important to acknowledge that many highly proficient bimodal bilinguals are professional interpreters. Studies of this group of language users, therefore, tend to include interpreters whose experiences with language activation and inhibition are likely to differ from non-interpreters (but see Christoffels et al. 2006 for evidence that interpreting experience in unimodal bilinguals does not alter primary language processes above and beyond their high proficiency). Whether the inclusion or exclusion of professional interpreters, as well as considerations of situational changes in the individual, such as whether an interpreter is at work or not, would affect the basic cognitive and linguistic patterns identified in previous studies is an open question. Another open question concerns whether there are differences in terms of the presence and strength of language co-activation between CODAs vs. bimodal bilinguals who learned the sign language in a classroom as adults vs. those who learned in contexts of language immersion.

Dynamic Recruitment of Cognitive Control during Language Processing

Recent studies have investigated dynamicity in the relationship between language processing and cognitive control. By using eye-tracking, Hsu and Novick (2016) investigated how efficiently monolingual English speakers were able to revise incorrectly parsed sentences as a function of their level of cognitive control engagement. High vs. low engagement contexts were created by interleaving congruent and incongruent Stroop and language processing tasks. When the sentence processing task took place during high cognitive control engagement (following an incongruent Stroop trial), the initial incorrect parse was corrected faster than in low cognitive control engagement contexts. This result suggests that monolinguals' language processing benefits from dynamically engaging cognitive control (see also Hsu et al. 2021; Thothathiri et al. 2018). However, similar studies with bilinguals have not revealed the same benefit of dynamic cognitive control engagement for language processing. Specifically, bilinguals appear to take advantage of their experience with language regulation to activate a high degree of cognitive control, irrespective of the experimental manipulation. Crucially, the effect of cognitive control recruitment in bilinguals appears to suffice to overcome any potential disadvantages associated with language co-activation, such as the well-documented delays in lexical processing observed in bilinguals compared to monolinguals (see Bialystok 2009). What has not yet been discovered, however, is how a bilingual's language environment and their experience with language regulation may affect their recruitment of cognitive control for the purposes of language processing. Bimodal bilingualism does not confer the boost to cognitive control abilities that unimodal bilingualism does, and we might therefore expect to see bimodal bilinguals' processing benefit from the dynamic engagement of cognitive control. At the same time, work by Giezen and Emmorey (2016) provides evidence that the bimodal bilinguals' ability to use both their languages at the same time does not appear to mitigate the typically observed disadvantage in lexical access in bilinguals compared to monolinguals (Bialystok et al. 2008; Gollan et al. 2002; but see Blanco-Elorrieta and Caramazza (2021) for an account that assumes that language processing requires the same level of effort for monolinguals and bilinguals). This could disadvantage bimodal bilinguals' language processing compared to monolinguals', while at the same time, bimodal bilinguals are not expected to show the same ability to recruit cognitive control abilities to support language processing as unimodal bilinguals are. Such issues are among the many unanswered questions about bimodal bilingualism, which offer the potential to better understand the interactions between language experiences, language regulation, cognitive control, and language processing.

5. Conclusions

The preceding sections have presented an overview of what the study of bimodal bilingualism contributes to our understanding of bilingualism more generally and in particular to understanding the interactions between language regulation, cognitive control, and language processing.

Bimodal bilingualism can be a tool for investigating questions about language and cognition that are less approachable when studying monolinguals or unimodal bilinguals. There is by now a substantial body of research that has investigated language behaviors, language switching, and cognitive control in proficient bimodal bilinguals. Among other things, this research has revealed which aspects of bilingualism are subject to modality effects and how the language system handles two languages when they can be used at the same time. Crucially, this work has also provided the means to establish clearer links between bilingual experiences and adaptive outcomes. Yet, for all that the research on bimodal bilingualism has contributed in the last decades, many questions still remain unanswered.

In this paper, we have highlighted the value of investigating bilingual variation. It will be an important endeavor for future research on bimodal bilingualism to grapple with the questions of variation and how the language environment affects language use, regulation, and cognitive adaptations in individuals who sign and speak. Another avenue of research that should be pursued relates to sign language learning. While most of the current research has focused on proficient bimodal bilinguals, there are important implications to understanding learning in bimodal bilinguals. Does knowing a signed second language make subsequent spoken language learning easier? Are unimodal bilinguals advantaged in learning a sign language? To what extent does the ability to regulate two spoken languages affect the nature of language regulation in an additional signed language? We noted, for example, that language immersion in a spoken L2 results in long-lasting inhibition of the L1. An important question for understanding the role of L1 regulation for L2 learning is whether immersion in a sign language produces similar effects and whether outcomes are similar in learners with and without prior spoken L2 proficiency. Questions such as these hold great promise for revealing the interactions between bilingualism, cognition, and language learning ability.

In sum, bilingualism provides a lens through which to investigate plasticity in the interactions between language and cognition that may be obscured in monolinguals. Research from the last decade has shown that variation in bilingual language experiences impacts minds and brains, and it has become clear that an approach that embraces variation in language proficiency, language environment, different types of contexts, as well as in language modality is necessary to reveal the plasticity of which the human brain is capable.

Author Contributions: Conceptualization, A.T.F. and J.F.K.; resources, A.T.F. and J.F.K.; writing—original draft preparation, A.T.F. and J.F.K.; writing—review and editing, A.T.F. and J.F.K.; funding acquisition, A.T.F. and J.F.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation, grant number SMA-2005246 and OISE-1545900.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Note that bimodal bilingualism is different from signed unimodal bilingualism, where individuals know two languages from the signed modality (although not much is known about bilingualism in the visual modality, but see Zeshan and Panda 2018; Koulidobrova 2019. See also Chen Pichler et al. 2019 for an overview of work on signed unimodal bilingualism).
2. We focus here on bimodal bilingualism in adults, but research has also been performed on child bimodal bilinguals (see, for example, Baker and Van den Bogaerde 2008; Van den Bogaerde and Baker 2005; Lillo-Martin et al. 2014).
3. A potentially important difference between CODAs and most heritage speakers of a spoken home language is that sign languages do not have a written form. We will not focus on the implications of this difference for literacy development but note that there are parallels with the phenomenon of diglossia, e.g., as seen in languages such as Arabic.
4. Deaf–blind individuals can perceive a variant of sign languages in the tactile modality
5. The natural language blending discussed here stands in contrast to SimCom (simultaneous communication), which is a form of communication that did not evolve naturally. Empirical evidence for the difference between natural and non-natural language blending comes from a study by Emmorey et al. (2005). In this study, CODAs were asked to narrate stories in SimCom or "coda talk", and the former was found to cause disfluencies, whereas the latter did not.
6. Kaufmann et al. (2018) note that they cannot rule out that the greater switching cost in unimodal contexts might be partially attributable to the participants' greater proficiency in English compared to DGS.

7 For bimodal bilinguals, the interactional context includes features that may shape the cognitive consequences of their bilingualism. The signed languages must be maintained, but in most environments, they may be unlikely to encounter signers, given the small number of signers in the larger environment.

References

Abutalebi, Jubin, and David W. Green. 2016. Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition* 19: 689–98. [CrossRef]

Antoniou, Mark. 2019. The Advantages of Bilingualism Debate. *Annual Review of Linguistics* 5: 395–415. [CrossRef]

Bak, Thomas H., Jack J. Nissan, Michael M. Allerhand, and Ian J. Deary. 2014. Does bilingualism influence cognitive aging? *Annals of Neurology* 75: 959–63. [CrossRef] [PubMed]

Baker, Anne, and Beppie Van den Bogaerde. 2008. Code-Mixing in Signs and Words in Input to and Output from Children. In *Sign Bilingualism: Language Development, Interaction, and Maintenance in Sign Language Contact Situations*. Edited by Carolina Plaza-Pust and Esperanza Morales-López. Amsterdam and Philadelphia: John Benjamins Pub Co., pp. 1–27.

Barac, Raluca, and Ellen Bialystok. 2012. Bilingual Effects on Cognitive and Linguistic Development: Role of Language, Cultural Background, and Education. *Child Development* 83: 413–22. [CrossRef]

Bartolotti, James, and Viorica Marian. 2012. Language Learning and Control in Monolinguals and Bilinguals. *Cognitive Science* 36: 1129–47. [CrossRef]

Baum, Shari, and Debra Titone. 2014. Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics* 35: 857–94. [CrossRef]

Baus, Cristina, Albert Costa, and Manuel Carreiras. 2013. On the effects of second language immersion on first language production. *Acta Psychologica* 142: 402–9. [CrossRef]

Beatty-Martínez, Anne L., and Debra A. Titone. 2021. The Quest for Signals in Noise: Leveraging Experiential Variation to Identify Bilingual Phenotypes. *Languages* 6: 168. [CrossRef]

Beatty-Martínez, Anne L., Christian A. Navarro-Torres, Paola E. Dussias, Maria T. Bajo, Rosa E. Guzzardo Tamargo, and Judith F. Kroll. 2020. Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 1022–47. [CrossRef]

Berent, Gerald P. 2004. Sign Language—Spoken Language Bilingualism: Code Mixing and Mode Mixing by ASL-English Bilinguals. In *The Handbook of Bilingualism*. Edited by Tej K. Bhatia and William C. Ritchie. Hoboken: Blackwell Publishing Ltd., pp. 312–35. [CrossRef]

Bialystok, Ellen. 2009. Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition* 12: 3–11. [CrossRef]

Bialystok, Ellen. 2017. The Bilingual Adaptation: How Minds Accommodate Experience. *Psychological Bulletin* 143: 233–62. [CrossRef]

Bialystok, Ellen, and Fergus I. M. Craik. 2022. How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review*. [CrossRef]

Bialystok, Ellen, Fergus I. Craik, and Gigi Luk. 2008. Cognitive Control and Lexical Access in Younger and Older Bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 34: 859–73. [CrossRef] [PubMed]

Bialystok, Ellen, Fergus I. Craik, and Morris Freedman. 2007. Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45: 459–64. [CrossRef] [PubMed]

Bice, Kinsey, and Judith F. Kroll. 2015. Native language change during early stages of second language learning. *NeuroReport* 26: 966–71. [CrossRef] [PubMed]

Bishop, Michele. 2010. Happen Can't Hear: An Analysis of Code-Blends in Hearing, Native Signers of American Sign Language. *Sign Language Studies* 11: 205–40. [CrossRef]

Bishop, Michele, and Sherry Hicks. 2005. Orange Eyes: Bimodal Bilingualism in Hearing Adults from Deaf Families. *Sign Language Studies* 5: 188–230. [CrossRef]

Blanco-Elorrieta, Esti, and Alfonso Caramazza. 2021. A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition* 213: 104625. [CrossRef]

Blanco-Elorrieta, Esti, Karen Emmorey, and Liina Pylkkänen. 2018. Language switching decomposed through MEG and evidence from bimodal bilinguals. *Proceedings of the National Academy of Sciences* 115: 9708–13. [CrossRef]

Blumenfeld, Henrike K., and Viorica Marian. 2013. Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology (Hove, England)* 25: 547–67. [CrossRef]

Bogulski, Cari A., Kinsey Bice, and Judith F. Kroll. 2019. Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition* 22: 1052–67. [CrossRef] [PubMed]

Braver, Todd S. 2012. The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences* 16: 106–13. [CrossRef] [PubMed]

Brice, Henry, Stephen J. Frost, Atira S. Bick, Peter J. Molfese, Jay G. Rueckl, Kenneth R. Pugh, and Ram Frost. 2021. Tracking second language immersion across time: Evidence from a bi-directional longitudinal cross-linguistic fMRI study. *Neuropsychologia* 154: 107796. [CrossRef]

Brito, Natalie, and Rachel Barr. 2012. Influence of bilingualism on memory generalization during infancy: Bilingualism and memory generalization. *Developmental Science* 15: 812–16. [CrossRef] [PubMed]

Casey, Shannon, Karen Emmorey, and Heather Larrabee. 2012. The effects of learning American Sign Language on co-speech gesture. *Bilingualism: Language and Cognition* 15: 677–86. [CrossRef]

Cenoz, Jasone. 2003. The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism* 7: 71–87. [CrossRef]

Christoffels, Ingrid K., Annette M. B. De Groot, and Judith F. Kroll. 2006. Memory and language skill in simultaneous interpreting: The role of expertise and language proficiency. *Journal of Memory and Language* 54: 324–45. [CrossRef]

Cole, Michael W, and Walter Schneider. 2007. The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage* 37: 343–60. [CrossRef]

Corina, David, Laurel Lawyer, and Deborah Cates. 2013. Cross-Linguistic Differences in the Neural Representation of Human Language: Evidence from Users of Signed Languages. *Frontiers in Psychology* 3: 587. [CrossRef]

Costa, Albert, and Mikel Santesteban. 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. [CrossRef]

Costa, Albert, Mario Pannunzi, Gustavo Deco, and Martin J. Pickering. 2017. Do Bilinguals Automatically Activate Their Native Language When They Are Not Using It? *Cognitive Science* 41: 1629–44. [CrossRef] [PubMed]

Declerck, Mathieu, and Andrea M. Philipp. 2015. A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review* 22: 1630–45. [CrossRef]

Declerck, Mathieu, and Iring Koch. 2022. The concept of inhibition in bilingual control. *Psychological Review*. [CrossRef]

Degani, Tamar, Anat Prior, and Walaa Hajajra. 2018. Cross-language semantic influences in different script bilinguals. *Bilingualism: Language and Cognition* 21: 782–804. [CrossRef]

Della Rosa, Pasquale A., Gerda Videsott, Virginia Borsa, Matteo Canini, Brendan Weekes, Rita Franceschini, and Jubin Abutalebi. 2013. A neural interactive location for multilingual talent. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 49: 605–8. [CrossRef] [PubMed]

Derrfuss, Jan, Marcel Brass, and D. Yves von Cramon. 2004. Cognitive control in the posterior frontolateral cortex: Evidence from common activations in task coordination, interference control, and working memory. *Neuroimage* 23: 604–12. [CrossRef] [PubMed]

Dias, Patricia, Saul Villameriel, Marcel R. Giezen, Brendan Costello, and Manuel Carreiras. 2017. Language switching across modalities: Evidence from bimodal bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43: 1828–34. [CrossRef] [PubMed]

Dijkstra, Ton, and Walter J. B. Van Heuven. 2018. Visual word recognition in multilinguals. In *The Oxford Handbook of Psycholinguistics*, 2nd ed. Edited by S. Rueshemeyer and M. G. Gaskell. Oxford: Oxford University Press, pp. 118–43.

Emmorey, Karen. 2002. *Language, Cognition, and the Brain: Insights from Sign Language Research*. Mahwah: Lawrence Erlbaum Associates Publishers, pp. xvii, 383.

Emmorey, Karen, Chuchu Li, Jennifer Petrich, and Tamar H. Gollan. 2020. Turning languages on and off: Switching into and out of code-blends reveals the nature of bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 443–54. [CrossRef]

Emmorey, Karen, Gigi Luk, Jennie E. Pyers, and Ellen Bialystok. 2008a. The Source of Enhanced Cognitive Control in Bilinguals. *Psychological Science* 19: 1201–6. [CrossRef]

Emmorey, Karen, Helsa B. Borinstein, and Robin Thompson. 2005. Bimodal Bilingualism: Code-blending between Spoken English and American Sign Language. In *ISB4: Proceedings of the 4th International Symposium on Bilingualism*. Edited by James Cohen, Kara T. McAlister, Kellie Rolstad and Jeff MacSwann. Somerville: Cascadilla Press, pp. 663–73.

Emmorey, Karen, Helsa B. Borinstein, Robin Thompson, and Tamar H. Gollan. 2008b. Bimodal bilingualism. *Bilingualism: Language and Cognition* 11: 43–61. [CrossRef] [PubMed]

Emmorey, Karen, Jennifer Petrich, and Tamar H. Gollan. 2012. Bilingual processing of ASL-English code-blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language* 67: 199–210. [CrossRef] [PubMed]

Emmorey, Karen, Marcel R. Giezen, and Tamar H. Gollan. 2016. Psycholinguistic, cognitive, and neural implications of bimodal bilingualism. *Bilingualism (Cambridge, England)* 19: 223–42. [CrossRef] [PubMed]

Emmorey, Karen, Stephen McCullough, Sonya Mehta, and Thomas J. Grabowski. 2014. How sensory-motor systems impact the neural organization for language: Direct contrasts between spoken and signal language. *Frontiers in Psychology* 5: 484. [CrossRef] [PubMed]

Ferjan Ramírez, Naja, Rey R. Ramírez, Maggie Clarke, Sammu Taulu, and Patricia. K. Kuhl. 2017. Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science* 20: e12427. [CrossRef] [PubMed]

Frederiksen, Anne Therese. 2021. Emerging ASL Distinctions in Sign-Speech Bilinguals' Signs and Co-speech Gestures in Placement Descriptions. *Frontiers in Psychology* 12: 686485. [CrossRef] [PubMed]

Fricke, Melinda, Megan Zirnstein, Christian Navarro-Torres, and Judith F. Kroll. 2019. Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition* 22: 1–8. [CrossRef]

Giezen, Marcel R., and Karen Emmorey. 2016. Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. *Bilingualism: Language and Cognition* 19: 264–76. [CrossRef]

Giezen, Marcel R., Henrike K. Blumenfeld, Anthony Shook, Viorica Marian, and Karen Emmorey. 2015. Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition* 141: 9–25. [CrossRef]

Gollan, Tamar, Rosa Montoya, and Grace Werner. 2002. Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology* 16: 562–76. [CrossRef]
Grainger, Jonathan, and Celine Beauvillain. 1987. Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology—QUART J EXP PSYCH A-HUM EXP P* 39: 295–319. [CrossRef]
Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef] [PubMed]
Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36: 3–15. [CrossRef]
Grundy, John G. 2020. The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science* 4: 177–99. [CrossRef]
Gu, Yan, Yequi Zheng, and Marc Swerts. 2019. Having a different pointing of view about the future: The effect of signs on co-speech gestures about time in Mandarin–CSL bimodal bilinguals. *Bilingualism: Language and Cognition* 22: 836–47. [CrossRef]
Gullifer, Jason W., and Debra Titone. 2020. Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition* 23: 283–94. [CrossRef]
Gullifer, Jason W., Xiaoqian J. Chai, Veronica Whitford, Irina Pivneva, Shari Baum, Denise Klein, and Debra Titone. 2018. Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia* 117: 123–34. [CrossRef]
Hoshino, Noriko, and Judith F. Kroll. 2008. Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition* 106: 501–11. [CrossRef]
Hsu, Nina S., and Jared M. Novick. 2016. Dynamic Engagement of Cognitive Control Modulates Recovery From Misinterpretation During Real-Time Language Processing. *Psychological Science* 27: 572–82. [CrossRef]
Hsu, Nina S., Stefanie E. Kuchinsky, and Jared M. Novick. 2021. Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience* 36: 211–39. [CrossRef]
Hsu, Nina S., Susanne M. Jaeggi, and Jared M. Novick. 2017. A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language* 166: 63–77. [CrossRef]
Kang, Chunyan, Fengyang Ma, Shuhua Li, Judith F. Kroll, and Taomei Guo. 2020. Domain-general inhibition ability predicts the intensity of inhibition on non-target language in bilingual word production: An erp study. *Bilingualism: Language and Cognition* 23: 1056–69. [CrossRef]
Kaufmann, Emily, Irene Mittelberg, Iring Koch, and Andrea M. Philipp. 2018. Modality effects in language switching: Evidence for a bimodal advantage. *Bilingualism: Language and Cognition* 21: 243–50. [CrossRef]
Kaushanskaya, Margarita, and Viorica Marian. 2009. The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review* 16: 705–10. [CrossRef]
Koulidobrova, Elena. 2019. Argument omission in SignL2 acquisition by deaf learners. In *Three Streams of Generative Language Acquisition Research: Selected Papers from the 7th Meeting of Generative Approaches to Language Acquisition*. Edited by Tania Ionin and Matthew Rispoli. Amsterdam and Philadelphia: John Benjamins Publishing Company, vol. 63, pp. 497–519. Available online: https://uci.primo.exlibrisgroup.com (accessed on 27 July 2022).
Kovács, Á. Melinda, and Jacques Mehler. 2009. Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 6556–60. [CrossRef]
Kroll, Judith F., and Christian Navarro-Torres. 2018. Bilingualism. In *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. Edited by John Wixted. Hoboken: Wiley/Blackwell, pp. 245–74.
Kroll, Judith F., Andrea Takahesu Tabori, and Christian Navarro-Torres. 2021. Capturing the variation in language experience to understand language processing and learning. *Language, Interaction and Acquisition* 12: 82–109. [CrossRef]
Kroll, Judith F., Kinsey Bice, Mona R. Botezatu, and Megan Zirnstein. 2022. On The Dynamics of Lexical Access In Two or More Languages. In *The Oxford Handbook of the Mental Lexicon*. Edited by Anna Papafragou, John C. Trueswell and Lila R. Gleitman. Oxford: Oxford University Press, pp. 582–97. [CrossRef]
Kroll, Judith F., Paola E. Dussias, and Maria T. Bajo. 2018. Language Use Across International Contexts: Shaping the Minds of L2 Speakers. *Annual Review of Applied Linguistics* 38: 60–79. [CrossRef]
Kroll, Judith F., Paola E. Dussias, Kinsey Bice, and Lauren Perrotti. 2015. Bilingualism, Mind, and Brain. *Annual Review of Linguistics* 1: 377–94. [CrossRef]
Kroll, Judith F., Susan Bobb, and Zofia Wodniecka. 2006. Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition* 9: 119–35. [CrossRef]
Kroll, Judith F., Susan C. Bobb, Maya Misra, and Taomei Guo. 2008. Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica* 128: 416–30. [CrossRef]
Kuntze, Marlon. 2000. Codeswitching in ASL and Written English Language Contact. In *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*. Edited by Karen Emmorey and Harlan L. Lane. Hove: Psychology Press, pp. 287–302. [CrossRef]
Li, Le, Jubin Abutalebi, Lijuan Zou, Xin Yan, Lanfang Liu, Xiaoxia Feng, Ruiming Wang, Taomei Guo, and Guosheng Ding. 2015. Bilingualism alters brain functional connectivity between "control" regions and "language" regions: Evidence from bimodal bilinguals. *Neuropsychologia* 71: 236–47. [CrossRef]

Li, Ping, Jennifer Legault, and Kaitlin A. Litcofsky. 2014. Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 58: 301–24. [CrossRef] [PubMed]

Lillo-Martin, Diane, Ronice M. de Quadros, Deborah Chen Pichler, and Zoe Fieldsteel. 2014. Language Choice in Bimodal Bilingual Development. *Frontiers in Psychology* 5. [CrossRef]

Linck, Jared A., Judith F. Kroll, and Gretchen Sunderman. 2009. Losing Access to the Native Language While Immersed in a Second Language: Evidence for the Role of Inhibition in Second-Language Learning. *Psychological Science* 20: 1507–15. [CrossRef] [PubMed]

Luk, Gigi, and Ellen Bialystok. 2013. Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology* 25: 605–21. [CrossRef] [PubMed]

Luk, Gigi, Ellen Bialystok, Fergus I. M. Craik, and Cheryl L. Grady. 2011. Lifelong Bilingualism Maintains White Matter Integrity in Older Adults. *The Journal of Neuroscience* 31: 16808–13. [CrossRef] [PubMed]

Luque, Alicia, and Kara Morgan-Short. 2021. The relationship between cognitive control and second language proficiency. *Journal of Neurolinguistics* 57: 100956. [CrossRef]

Macnamara, Brooke N., and Andrew R. A. Conway. 2014. Novel evidence in support of the bilingual advantage: Influences of task demands and experience on cognitive control and working memory. *Psychonomic Bulletin & Review* 21: 520–25. [CrossRef]

MacSweeney, Mairéad, Cheryl M. Capek, Ruth Campbell, and Bencie Woll. 2008. The signing brain: The neurobiology of sign language. *Trends in Cognitive Sciences* 12: 432–40. [CrossRef]

Meuter, Renata F. I., and Alan Allport. 1999. Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language* 40: 25–40. [CrossRef]

Michael, Erica B., and Tamar H. Gollan. 2005. Being and Becoming Bilingual: Individual Differences and Consequences for Language Production. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, pp. 389–407.

Misra, Maya, Taomei Guo, Susan C. Bobb, and Judith F. Kroll. 2012. When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language* 67: 224–37. [CrossRef]

Miyake, Akira, and Naomi P. Friedman. 2012. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science* 21: 8–14. [CrossRef] [PubMed]

Morford, Jill P., and Judith F. Kroll. 2021. Bilingualism in Deaf and Hearing Signers. In *Bilingualism Across the Lifespan*, 1st ed. Edited by Wendy S. Francis. New York: Routledge, pp. 165–81. [CrossRef]

Morford, Jill P., Corinne Occhino, Megan Zirnstein, Judith F. Kroll, E. Wilkinson, and Pilar Piñar. 2019. What is the Source of Bilingual Cross-Language Activation in Deaf Bilinguals? *The Journal of Deaf Studies and Deaf Education* 24: 356–65. [CrossRef] [PubMed]

Morford, Jill P., Erin Wilkinson, Agnes Villwock, Pilar Piñar, and Judith F. Kroll. 2011. When deaf signers read English: Do written words activate their sign translations? *Cognition* 118: 286–92. [CrossRef] [PubMed]

Morford, Jill. P., Judith F. Kroll, Pilar Piñar, and Erin Wilkinson. 2014. Bilingual word recognition in deaf and hearing signers: Effects of proficiency and language dominance on cross-language activation. *Second Language Research* 30: 251–71. [CrossRef] [PubMed]

Novick, Jared M., John C. Trueswell, and Sharon L. Thompson-Schill. 2005. Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience* 5: 263–81. [CrossRef]

Nozari, Nazbanou, Daniel Mirman, and Sharon L. Thompson-Schill. 2016. The ventrolateral prefrontal cortex facilitates processing of sentential context to locate referents. *Brain and Language* 157–58: 1–13. [CrossRef]

Olulade, Olumide A., Nasheed I. Jamal, David S. Koo, Charles A. Perfetti, Carol LaSasso, and Guinevere F. Eden. 2016. Neuroanatomical Evidence in Support of the Bilingual Advantage Theory. *Cerebral Cortex* 26: 3196–204. [CrossRef]

Ooi, Seok H., Winston Goh, Antonella Sorace, and Thomas Bak. 2018. From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition* 21: 1–13. [CrossRef]

Ovans, Zoe, Nina S. Hsu, Donald Bell-Souder, Phillip Gilley, Jared M. Novick, and Albert E. Kim. 2022. Cognitive control states influence real-time sentence processing as reflected in the P600 ERP. *Language, Cognition and Neuroscience 1–9*. [CrossRef]

Pashler, Harold. 1994. Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin* 116: 220–44. [CrossRef]

Petitto, Laura Ann, Melody S. Berens, Iulia Kovelman, Matthew H. Dubins, Kaja Jasinska, and Mark Shalinsky. 2012. The "Perceptual Wedge" hypothesis as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language* 121: 130–43. [CrossRef] [PubMed]

Philipp, Andrea M., and Iring Koch. 2009. Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35: 1187–95. [CrossRef] [PubMed]

Pichler, Deborah C., Elena Koulidobrova, and Jeffrey L. Palmer. 2019. Modality-(in) dependent second language learning. *Senri Ethnological Studies* 101: 165–86.

Pliatsikas, Christos. 2020. Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition* 23: 459–71. [CrossRef]

Pot, Anna, Joanna Porkert, and Merel Keijzer. 2019. The Bidirectional in Bilingual: Cognitive, Social and Linguistic Effects of and on Third-Age Language Learning. *Behavioral Sciences* 9: 98. [CrossRef] [PubMed]

Pyers, Jennie E., and Karen Emmorey. 2008. The Face of Bimodal Bilingualism: Grammatical Markers in American Sign Language Are Produced When Bilinguals Speak to English Monolinguals. *Psychological Science* 19: 531–35. [CrossRef]

Sandler, Wendy, and Diane Lillo-Martin. 2006. *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Shook, Anthony, and Viorica Marian. 2012. Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition* 124: 314–24. [CrossRef]

Singleton, Jenny L., and Matthew D. Tittle. 2000. Deaf Parents and Their Hearing Children. *Journal of Deaf Studies and Deaf Education* 5: 221–36. [CrossRef]

Surrain, Sarah, and Gigi Luk. 2019. Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition* 22: 401–15. [CrossRef]

Thierry, Guillaume, and Yan Jing Wu. 2007. Brain potentials reveal unconscious translation during foreign-language comprehension. *The Proceedings of the National Academy of Sciences* 104: 12530–35. [CrossRef]

Thomas, Michael S. C., and Alan Allport. 2000. Language switching costs in bilingual visual word recognition. *Journal of Memory and Language* 43: 44–66. [CrossRef]

Thothathiri, Malathi, Christine T. Asaro, Nina S. Hsu, and Jared M. Novick. 2018. Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition* 178: 162–77. [CrossRef] [PubMed]

Treffers-Daller, Jeanine. 2019. What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics* 5: 375–93. [CrossRef]

Van Assche, Eva, Wouter Duyck, and Tamar H. Gollan. 2013. Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 39: 1781–92. [CrossRef]

Van den Bogaerde, Beppie, and Anne Baker. 2005. Code Mixing in Mother–Child Interaction in Deaf Families. *Sign Language & Linguistics* 8: 153–76.

van den Noort, Maurits, Esli Struys, Peggy Bosch, Lars Jaswetz, Benoit Perriard, Sujung Yeo, Pia Barisch, Katrien Vermeire, Sook-Hyun Lee, and Sabina Lim. 2019. Does the Bilingual Advantage in Cognitive Control Exist and If So, What Are Its Modulating Factors? A Systematic Review. *Behavioral Sciences* 9: 27. [CrossRef]

Vega-Mendoza, Mariana, Holly West, Antonella Sorace, and Thomas H. Bak. 2015. The impact of late, non-balanced bilingualism on cognitive performance. *Cognition* 137: 40–46. [CrossRef]

Villameriel, Saul Patricia Dias, Brendan Costello, and Manuel Carreiras. 2016. Cross-language and cross-modal activation in hearing bimodal bilinguals. *Journal of Memory and Language* 87: 59–70. [CrossRef]

Weisberg, Jill, Shannon Casey, Zed Sevcikova Sehyr, and Karen Emmorey. 2020. Second language acquisition of American Sign Language influences co-speech gesture production. *Bilingualism: Language and Cognition* 23: 473–82. [CrossRef]

Williams, Joshua, and Sharlene Newman. 2016. Interlanguage dynamics and lexical networks in nonnative L2 signers of ASL: Cross-modal rhyme priming. *Bilingualism: Language and Cognition* 19: 453–70. [CrossRef]

Xie, Zhilong. 2018. The Influence of Second Language (L2) Proficiency on Cognitive Control Among Young Adult Unbalanced Chinese-English Bilinguals. *Frontiers in Psychology* 9: 412. [CrossRef]

Zeshan, Ulrike, and Sibaji Panda. 2018. Sign-speaking: The structure of simultaneous bimodal utterances. *Applied Linguistics Review* 9: 1–34. [CrossRef]

Zhang, Haoyun, Chunyan Kang, Yanjing Wu, Fengyang Ma, and Taomei Guo. 2015. Improving proactive control with training on language switching in bilinguals. *NeuroReport* 26: 354–59. [CrossRef] [PubMed]

Zhang, Haoyun, Michele T. Diaz, Taomei Guo, and Judith F. Kroll. 2021. Language immersion and language training: Two paths to enhanced language regulation and cognitive control. *Brain and Language* 223: 105043. [CrossRef] [PubMed]

Zhu, Judy D., Esti Blanco-Elorrieta, Yanan Sun, Anita Szakay, and Paul F. Sowman. 2022. Natural vs. forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage* 247: 118797. [CrossRef] [PubMed]

Zirnstein, Megan, Janet G. van Hell, and Judith F. Kroll. 2018. Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition* 176: 87–106. [CrossRef] [PubMed]

Zou, Lijuan, Guosheng Ding, Jubin Abutalebi, Hua Shu, and Danling Peng. 2012a. Structural plasticity of the left caudate in bimodal bilinguals. *Cortex* 48: 1197–206. [CrossRef]

Zou, Lijuan, Jubin Abutalebi, Benjamin Zinszer, Xin Yan, Hua Shu, Danling Peng, and Guosheng Ding. 2012b. Second language experience modulates functional brain network for the native language production in bimodal bilinguals. *NeuroImage* 62: 1367–75. [CrossRef]

languages

MDPI

Concept Paper

On Path Diagrams and the Neurophenomenal Field in Bilinguals

David William Green

Department of Experimental Psychology, Faculty of Brain Sciences, University College London, London WC1E 6BT, UK; d.w.green@ucl.ac.uk

Abstract: Conversation is a major site for our use of language. Each conversation elicits a distinct subjective experience: a specific and dynamic phenomenal field, and it is this field that controls our communicative actions. We cannot hope to understand the neural bases of conversation without relating these to the phenomenal field. We need a neurophenomenology of the bilingual speaker. I propose and illustrate an approach involving path diagrams together with retrospective experience sampling to capture the richness of the phenomenal field as a speaker talks through an issue of concern, and relate this process to large-scale attentional networks. The proposal offers a general approach to developing a neurophenomenology of the bilingual speaker and listener.

Keywords: path diagrams; neurophenomenology; neurophenomenal field; language attrition; mental simulation; counterfactual thinking; attentional networks; retrospective experience sampling

1. Introduction: Rationale and Setting the Scene

An important goal of research in bilingualism is to understand the linguistic and cognitive processes involved in language use, and that entails understanding such processes in a key site of language use: conversation. We talk to others to share our thoughts and feelings about different topics and so express our lived, subjective experience about them. This subjective or phenomenal experience is important if we are to make sense of the neural bases of language use, as revealed by analyses of the neural networks involved. How can we tell just by observing the dynamics of neural activity as a person looks at a film or talks about an episode in their lives what the neural flows of activity actually signify, without knowing what they are actually thinking or feeling? Linking the phenomenal and the neural domains is an exercise in neurophenomenology, a termed coined by Varela (1996), and is an exercise vital to further our understanding of the nature of human consciousness. Its purpose is to connect the study of the structure of human experience, as in the work of the phenomenologists such as Merleau-Ponty (1964), with research that examines the dynamics of brain activity. This exercise is vital too for the narrower project of understanding the neural basis of language use in bilingual speakers.

Consider a small thought experiment. Suppose, hypothetically, we record neural activity as a bilingual speaker narrates an experience in her first language, and then ask that she retell it in her second language. How might we explain a dramatic increase in the activation of networks mediating affect in the latter case when we know from self-reports that words or phrases in a non-native language may lack affective charge and lead to a misconstrual of the significance of the utterance for the addressee (Pavlenko 2017), and that current experimental research suggests that talking in a second language may increase emotional distance from an event (e.g., individuals may be more utilitarian in their decisions (Costa et al. 2014))? Well, for some bilingual speakers, the disconnect between their lived experience and the felt capacity of the language they are currently using to express it induces a real sense of frustration, and even anger (e.g., Wierzbicka 2004). If, after the recording, we actually ask the speaker what they were thinking or feeling at different

Citation: Green, David William. 2022. On Path Diagrams and the Neurophenomenal Field in Bilinguals. *Languages* 7: 260. https://doi.org/10.3390/languages7040260

Academic Editor: John W. Schwieter

Received: 18 May 2022
Accepted: 6 October 2022
Published: 12 October 2022

moments in time, we are in a position to develop an explanation. Particularly poignant examples of disconnection arise when children become attriters in the language of their parents. I make use of one reported example of such attrition to illustrate how we might bring subjective experience into our theorizing and experimental research, and so afford a deeper understanding of the neural bases of language use during conversation between bilingual speakers. In the next section I overview the approach before then detailing the case study.

Our experience of the world is integrated, multi-modal and situated from a first-person point of view: a centered spatial structure or phenomenal field. Its point of origin, at least for sighted individuals, is experienced behind the eyes. It is sustained as we act in the world and engage with others. It allows us to envisage other points of view. For example, it allows us to envisage the hidden parts of objects so that we can act appropriately (e.g., I see just one side of a pot, not its hidden side, but infer it and so grasp it appropriately). It also allows us to envisage the points of view of other agents in our world, and so is a grounding for intersubjectivity and our ability to attribute mental states to others (e.g., Wiliford et al. 2018). The perspectival properties of the phenomenal field are therefore essential to our ability to act in the world. My interest is with its corollary—its psychological characterization. It is distinct, as personal experience attests—we can look out on the world as we walk, for example, and be somewhere else in our thoughts.

Psychologically, for the purposes of this paper, the focus of the phenomenal field is a current issue of concern (a particular topic) with a periphery of factors that bear on it. The core idea is that its more detailed structure can be expressed as a path diagram connecting together factors that we deem relevant to the issue of concern. Such diagrams have been used to understand how individuals perceive a wide range of issues (e.g., the risk and prevention of coronary heart disease (Green and McManus 1995); the causes of unemployment and employment (Green et al. 1998); the causes of overfishing (Nikolic and Lagnado 2015)) and have been used in formal analyses identifying causal factors in numerous phenomena (Pearl 2000; Pearl and Mackenzie 2018). Here I propose such diagrams as a means to capture the psychological relations involved in the phenomenal field as a precursor to mapping these onto the neural substrates to capture the neurophenomenal field.

The paths in the diagram reflect our understanding of which factor can affect which other factors based on our interactions with our physical, cultural and interpersonal worlds. Such understanding is captured in the notion of a mental causal model (e.g., Craik 1943; Green 1997, 2001; Johnson-Laird 1983, chp. 15; Pearl 2000; Pearl and Mackenzie 2018) that may be intersubjectively shared within a community. Our models are generative and we use them to explain, predict and control experience. For instance, in the case of the interpersonal world, we can ask why did someone speak as they did? Further, we can answer it: Why did John argue with Mary? Because she voted for someone he disliked. One or other party can intervene and defuse the argument by agreeing to disagree. Such questions, answers and interventions are possible because our models express relations amongst relevant factors, including their causal relationships. We also hold models of ourselves expressed in the narratives we use as we interact with others to make sense of our past and give meaning to our future (Mead 1959; and see also Andersen et al. 2020, for the important narratives in the context of social action).

Whilst our models of our physical, cultural and interpersonal world underlie our perceptions and attributions, our phenomenal experience at some moment in time arises from our current concerns: the cognitive–affective nexus. This is the psychological focus of the phenomenal field. Talking about it (or reflecting on it) creates further structure in terms of the paths linking different factors to the cognitive–affective nexus. The paths themselves are an outcome of our own personal memories, and the possible futures we envisage are intimately related to whom we consider ourselves to be. As we will see in the illustration, "loving daughter" is a salient theme of the personal narrative, and is an important driver of transformation in the phenomenal field. The approach here makes evident that our phenomenal worlds mediated by language are permeated by imaginary and not simply by actual worlds.

In the illustration, I envisage that after talking about the specific concern to a friend, the participant constructs the path diagrams and also describes her thoughts and feelings at different points in time during the conversation. It is then the job of the researcher to code utterances in terms of the network of paths they reflect, and their experiential qualities, and analyze the neural dynamics coincident with the coded speech acts. I will say a little more about how we might construe that complexity in Section 2.3, but first I explore the illustrative case.

2. An Illustrative Case

Consider a familiar story: the life worlds of families migrating to a country whose main language (or lingua franca) they do not speak—Cantonese speakers, for instance, who migrate to America. A local community, whose language the parents speak, may support them and they thrive with hard work. Unsurprisingly, parents want their children to enjoy a better life with more material advantages than themselves. They muster an education for them so they become proficient in the dominant language (English, in the illustration). Some children as they grow older perhaps continue to live within the local community and function as language brokers with respect to the wider society. Others, with progressive mastery of English and the opportunities arising from their education, move away, identify themselves as members of American culture (assume an American identity) rather than as members of their parents' diaspora, and become attriters in their first language. The salient issue here is the practical and affective communicative rupture experienced by a hypothetical young woman in the circumstances described by Liao (2021): in conversation with her parents, she is no longer fluent in Cantonese, and they speak no English.

2.1. The Path Diagrams

Imagine if you will the rich data recorded as our participant describes her thoughts and feelings in conversation with a friend. She first talks about how she has been feeling.

She expresses awkwardness at having to use a translation app as she tries to find a Cantonese word or expression, regret at the loss of Cantonese, a sense of loss of her childhood, despite the sudden recall of specific childhood memories or a felt remembrance of childhood (see Sheldon et al. 2019, for this distinction). Then in a second part of the conversation, she talks about what she has decided to do as a result of a felt recognition that parents do not live forever. Her interventions are to relearn Cantonese and to envisage a new sense of self: a Cantonese–American identity. She anticipates a transformation in her lived experience. Awkwardness in conversation now becomes part of a joint project with the parents to relearn Cantonese—an exercise, at least initially, in translanguaging perhaps (Wei 2018). She sees it is possible to laugh about the process with her parents as she reaches for her translation app once more (see Appendix A) indicates the kind of change required to accommodate such a joint endeavor in an existing hierarchical control model of speech production and comprehension. The decision to form a new Cantonese–American identity becomes the intimation of a new self-narrative grounded in shared memories elicited through conversation and parental recollections—a richer intersubjectivity.

After the first part we ask her to draw a path diagram of the factors as she sees them that contribute to her salient concern—the focus of the phenomenal field (the cognitive and affective nexus). She draws Figure 1a. In this figure we see that she attributes attrition in Cantonese to her proficiency in English. Reduced Cantonese proficiency makes her interaction with her parents more difficult, both cognitively and emotionally. In the diagram, increased English proficiency contributes to an assimilation into American culture and identity, but suppresses identity based on Cantonese heritage. This too is seen as having a negative impact. It is important to note that the diagram can express relations that are not expressed in talk, but are in the phenomenal field. After the second part, she draws Figure 1b. Here two distinct but interlinked interventions are envisaged: relearning

Cantonese and forging a new self-identity. The estimated effect of relearning Cantonese is to improve the quality of the interaction with the parents as Cantonese proficiency improves. Any such improvements yield a happier state. The second intervention, namely, the goal to create a new self-narrative based on a Cantonese–American identity, helps to integrate her life story to date.

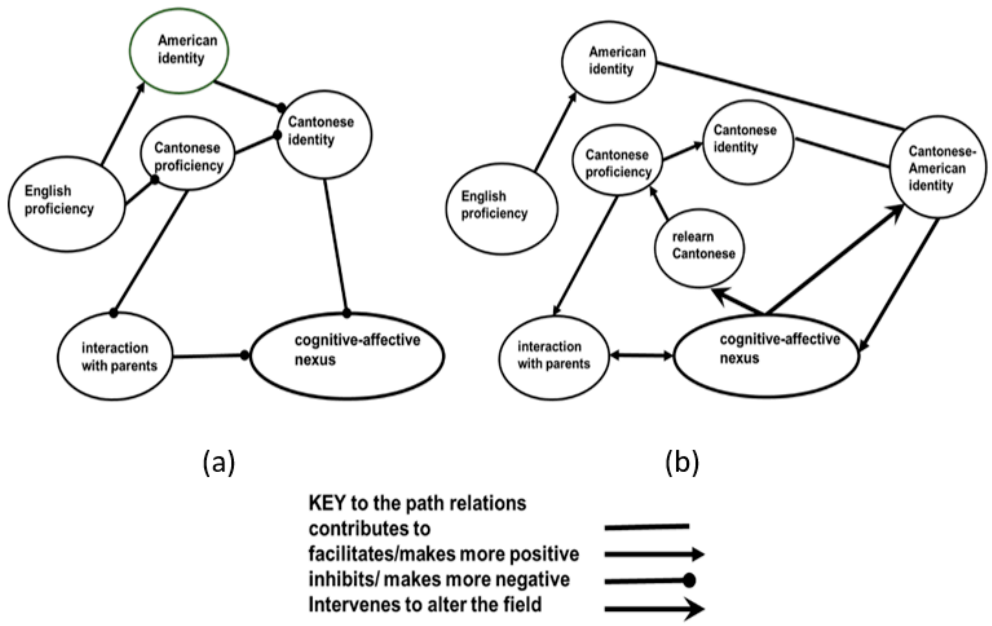

Figure 1. The Phenomenal Fields of an Illustrative Case: (**a**) before interventions, (**b**) after.

These two diagrams are snapshots of the principal relations of factors involved and so "arrest" the dynamical flow of experience. A comparison of the two snapshots provides an indication of the change in the phenomenal field arising out of the two interventions. Indeed, at the path level, the transition between the two snapshots depicts the (psychological) flow of the phenomenal field. Essentially formerly negative paths become positive (e.g., the path from Cantonese proficiency to the interaction with parents) or are deleted (e.g., the path from American identity to Cantonese identity), or become contributory to a novel factor (e.g., Cantonese identity to Cantonese–American identity). The flow arises from the process of resolving the cognitive–affective nexus: reflecting on the issue of concern drives the transformation.

2.2. Conversation, Path Diagrams and Cognitive–Affective and Linguistic Processes

In this part I consider the gist of the talk between our language attriter and her friend as she considers the communicative rupture with her parents and resolves to do something about it. The goal is to conjecture the cognitive–affective and linguistic processes involved as they relate to the paths in the diagrams before, in the following section, relating these to their neural bases.

Her utterances, capturing paths in the diagram, sometimes take the form of an argument comprising a claim and a reason (e.g., I will feel better because I am able to talk to them) or may issue as conditional statements (e.g., If I acknowledge my heritage I will feel whole). Such arguments arise primarily out of thinking counterfactually (i.e., if only I had ... or, if I were to ...). Cognitively, such thinking is a type of mental simulation and underlies emotions such as regret, wherein we envisage a better outcome if we had acted differently. In fact, we may choose actions that minimize anticipatory regret. We do so by placing ourselves in a hypothetical circumstance and mentally simulate alternative courses of action that lead to a preferred outcome. This yields an argument for action and provides the basis for responding to a question as to why we are intending to act in a particular way (cf.: Billig 1987; Green 2011; Vygotsky 1981).

My construal takes **mental simulation** as the core process, with a number of phases nested under it corresponding to separate stretches of talk as per the text. It is this mental simulation that captures the flow of the phenomenal field. I consider the talk (indicated in *italics* and not intended as vernacular) and offer a gloss of the main cognitive and affective processes involved (in **bold** and in brackets0). These glosses are high-level descriptions and are intended as indicative and simplifying. For example: **(imagine possible world)** covers both counterfactual "if only ... "and semi-factual "even if ... " thinking. The outcome of thinking counterfactually is the interventions to relearn Cantonese and to construct a new self-narrative. The gloss **(self-reflection)** includes the retrieval and use of autobiographical memories. The gloss **(affective attribution)** covers love, regret, blame and exoneration. Such attribution is a necessary part of thinking counterfactually, but the gloss allows a separate reference. These high-level glosses are combined on occasion to emphasize the relative weight of a process. A gloss at the beginning of a paragraph indicates the main process of the phase and, for simplicity's sake, ones at the end reference the various processes nested under it. The text in normal font that follows references the paths identified.

2.2.1. Part 1 of the Conversation Figure 1a Refers

> *Well I think about my interactions with my parents* **(self-reflection)**: *I feel an emotional distance because I hardly speak Cantonese anymore and have to use a translation app. Look I know they love me and know that they know I love them. I have become distant from my Cantonese heritage as I've assumed an American identity* **(affective attribution, mental state attribution** *nested under* **self-reflection)**

These utterances reflect the paths in Figure 1a to do with Cantonese proficiency and its impact both on her conversations with her parents, and her sense of identity as a loving daughter.

> *Am I to blame for this intersubjective rupture? I remember times when I was horrid to my parents and rebuked them: why had they never learned English?* **(self-reflection)** *and then think: if only I had maintained fluency in Cantonese but then (as a counter-argument) they wished me to speak English fluently and are proud of me and I am proud of myself too* **(imagine possible world, affective attribution, mental state attribution** nested under **self-reflection)**

These utterances justify the inhibitory path between English proficiency and attrition in Cantonese and address the emotional complexity involved.

2.2.2. Part 2 of the Conversation Figure 1b Refers

> *Well you know with the pandemic and all, I started to think. Is there a way forward out of this unhappy state?* **(imagine possible world)**
>
> *I could let matters go* **(imagine possible world)** *but I don't want to be a stranger to them* **(self-reflection).** *If I were to do nothing I will regret it* **(affective attribution)**. *Is there something I could do that will be better for myself and my parents?*

These utterances refer back initially to the state of affairs depicted in Figure 1a, and then look forward paths (interventions) in Figure 1b that could alter the cognitive–affective nexus for the better.

> A: *I will relearn Cantonese* **(imagine possible world):** *I think that I can improve my Cantonese. I know I can persevere* **(metacognitive processing)** *and if I do so it'll make talking to my parents so much better.*
>
> *And even if I do not succeed I will have tried and they will know it and that is a solace and no occasion for self-blame* **(metacognitive processing, affective attribution, mental state attribution** nested under **imagine possible world)**.

These utterances refer to a path emanating from the cognitive–affective nexus to a novel factor (relearning Cantonese) that enhances her interactions with her parents via a path coming from increased proficiency in Cantonese, and so helps resolve the issue of concern. They do not express the deletion of the negative path from English proficiency to Cantonese proficiency. Psychologically, though, this deletion signals a lack of opposition in her mind to the use of the two languages. In actual language use, given dual language activation, there will be an increased need for language control (see Kroll et al. 2015, for a review). The psychological value of the intervention is that it transforms her interactions with her parents and her future anticipated recollection of it.

B. *I will recognize my Cantonese heritage: I built an American identity at a cost of my personal history* **(self-reflection)**. *But if I work on my Cantonese heritage (e.g., my parents' life histories and culture) I do not want it to be in conflict with my American identity I will create a new Cantonese-American identity* **(imagine possible world** nested **under self-reflection).**

These utterances refer to a driving input from the cognitive–affective nexus to a novel factor—an imagined Cantonese–American identity in which her American identity no longer suppresses her heritage identity, but together with it contributes to a to-be-created identity. The path diagram also indicates something that is not directly expressed, that forging the new identity is envisaged to resolve the concern and also improve her interactions with her parents. Their life histories are a way for her to learn about her heritage, but also such conversations may well ground or embody her learning of Cantonese just as the initial learning of a language is learning a habit of conduct (Merleau-Ponty 1964).

2.3. The Neurophenomenology of Resolving the Cognitive–Affective Nexus

Neural dynamics can be captured at multiple scales, both spatial and temporal, reflecting self-organizing networks. Naming the coordination of regions as networks tuned to certain functions does not imply that the regions involved are specialized for such a function any more than the so-called word form area is specialized for word recognition. The functional profile of a region depends on the full set of its interconnections (Price and Friston 2005). The high-level descriptions likely cover a range of component processes, and so only detailed task comparisons may delimit the functional profile. However, it is convenient, and least tendentious here, to map the high-order process descriptions such as mental simulation to large-scale attentional networks. Table 1 references research that has associated the mental processes, such as mental simulation, that I have identified in the construction of the paths in the path diagrams with such large-scale networks or neural regions.

If the construal offered here is correct, there is analytic complexity to any analysis of neural data. The construal implicates a hierarchy of control in which one process is nested, at least transiently, within another (see also Green 2019). For example, in holding a conversation, we need to sustain attention on that goal whilst also calling on different networks as we change topics of conversation from talking about a shared external current scene to reporting a memory. Here nested under the goal to sustain attention to conversation, the attentional network mediating mental simulation, for example, must be sustained over some duration, as other networks and regions are coopted for other purposes. At a minimum, whilst perceptual and sensory processing continues, this entails that a network engaged in the fulfilment of some external task (e.g., the dorsal attentional network, **DAN**) is down-regulated or suppressed, at least temporarily, to allow internally focused attention to act. The default mode network (**DMN**) mediates such attention and its recruitment coincides with an increase alpha oscillation over parietal regions consistent with the suppression of **DAN** (e.g., Higgins et al. 2021). The **DMN** forms a core network for mental simulation and counter-factual thinking that is crucial to feelings of regret or self-blame (as expressed in her talk), and also in identifying suitable interventions. Such thinking requires distinguishing the present world from the imagined world, and is postulated to require a fronto-parietal and a cingulo-opercular network for its control (see van Hoeck et al. 2015). It also involves recruiting regions or networks that offer grounds for believing in the effectiveness of an intervention. Part of the reason she gives for undertaking to relearn Cantonese is the confidence she feels in her ability to persevere. Such a feeling is metacognitive and is associated with computations in rostral and dorsal prefrontal regions (Fleming and Dolan 2012).

Table 1. Resolving the cognitive–affective nexus: mapping to large-scale attentional networks.

High-Order Process	Neural Networks and Regions	Example References
mental simulation	**DMN**	Raichle (2015)
self-reflection	**DMN**	Spreng and Grady (2009)
		Sheldon et al. (2019)
	PCC, mPFC	Johnson et al. (2002)
imagine possible world	**DMN**	van Hoeck et al. (2015)
	FPC, C- O C	:
affective attribution	**amygdalae – basal ganglia**	:
	AI/plOFC; vmPFC/mOFC	:
mental state	**right dlPFC**	Cohen-Zimerman et al. (2021)
meta-cognitive processing	**rostral and dorsal PFC**	Fleming and Dolan (2012)

Key: DMN default mode network; **FPC** fronto-parietal control network; **C-O C** cingulo-opercular control network; **AI** anterior insula; **PFC** prefrontal cortex; **PCC** posterior cingulate cortex; **dlPFC** dorsolateral prefrontal cortex; **mPFC** medial prefrontal cortex; **vmPFC** ventromedial prefrontal cortex; **mOFC** medial orbito frontal cortex**; plOFC** posterior lateral orbito-frontal cortex.

At the heart of the present proposal is the driving role of affect. Affective processing is integral as it instigates the search for a preferred state (via mental simulation) and selects actions that plausibly lead to that state. van Hoeck et al. (2015) refer to an emotion and value processing network. It is the one mediating affective attribution and connects the amygdalae, striatum and frontal cortices, and in the proposal here is core to the cognitive–affective nexus and the drive to identify options expressed in the pathways of the path diagram to resolve a personal state of unhappiness. If this proposal has merit then it is reasonable to expect this network of regions to increase in activation as it provides a coordinating hub for identifying possible interventions. Such activation should then precede utterances that reference such interventions.

We can go further in our neurophenomenological approach. After the scan, we ask her what she was thinking and feeling at different moments in the scan. Retrospective experiential sampling of her experience at different time periods during the scan can deepen what can be inferred from her utterances (see Gonzalez-Castillo et al. 2021; Jääskeläinen et al. 2022; Jachs 2021, for discussion and use of such methods). We should expect retrospective sampling to reveal episodes of mental simulation subsuming periods of self-reflection and counterfactual thinking. Further, if she was asked to trace the felt psychological importance over successive time intervals (see Jachs 2021, for such a method) we might find high importance to be associated with the explicit recall of autobiographical memories under the dominance of the default mode network. Conceivably, as proposed above, the affective network is the primary driver of episodes in which she considers possible interventions, and so should precede episodes where she identifies the proposed interventions.

Analytically, the need is to explore the transient and flexible reorganization and coordination of the various neural networks implicated. Fortunately, researchers are actively exploring techniques such as Hidden Markov modeling to do so (e.g., Chen et al. 2017), which can be applied to both magnetoencephology data (Tibon et al. 2021) as well as to functional magnetic resonance imaging data (Quinn et al. 2018).

2.4. Further Tests and Predictions

I considered above how to explore the neurophenomenal field of our hypothetical attriter as she talked with a friend about the communicative rupture with her parents. Testing the proposal in this way may seem a promissory note if it involves full multimodal recording including body movement, facial expression and hand gestures, though that prospect may not be too far distant for certain neural-based signals, as wearable optically pumped magnetometers provide fine multichannel sensitivity for magnetoencephalography systems (e.g., Tierney et al. 2019). In this section I consider other ways to make use of the perceived structure of the phenomenal field together with retrospective experience sampling, and relate it to its neural basis.

We talk not only to others but to ourselves. Self-talk engages parts of the language-sensitive networks too (e.g., Jones and Fernyhough 2007; Grandchamp et al. 2019; Barber et al. 2021), and can concern all manner of topics. On the basis of such research, we can perform a type of resting state scan with our attriter. We ask her to silently think about and think through the issue of concern. After the scan, as previously, she draws a path diagram of the factors she considered important with its focus (the cognitive–affective nexus), and indicates how that diagram changed in the light of the intervention to relearn Cantonese and to forge a new sense of identity. Then, we make use of retrospective sampling to enhance our mapping from the subjective to the neural.

I have explored the value of path diagrams in understanding the neurophenomenal field of a hypothetical bilingual speaker. The approach can be extended with a novel between-subject protocol. We can ask bilingual speakers with common experiences and those from different community backgrounds and experiences to construct path diagrams with respect to a key issue of concern for them as bilingual speakers (see Beatty-Martínez et al. 2020, for the value of comparing bilingual speakers from different backgrounds in understanding behavioral data). Now, we record their neural responses as they listen to the narrative of our language attriter. We know from prior research that neural responses between speaker and listener couple or align most closely when listeners comprehend what is being said (Stephens et al. 2010; Liu et al. 2017). Under the novel protocol, we use path diagrams to predict time periods in the narrative that may have the most direct relevance to each individual., i.e., we predict periods of heightened cognitive–affective response. We can test these predictions by asking listeners retrospectively to profile their reactions at different time periods (e.g., as per Jachs 2021), and coordinate these with respect to their individual neural response. Under this type of protocol, the idea would be to examine whether or not the variety of neural response across participants as they listen to a common narrative can be explained by the phenomenal characterization of their personal bilingual histories. Studies of this type explore the neurophenomenal fields of different bilingual speakers.

3. Review: Characterizing the Neurophenomenal Field

The neural bases of language use in conversation between bilinguals will remain opaque to understanding until we capture the subjective experiences of the participants as well—we need to understand the neurophenomenal field. This paper proposes that we can use path diagrams (see Figure 1a,b) to capture the psychological structure of the phenomenal field: its cognitive–affective nexus (the focus), and the important factors and their path interrelations that bear on it. Such diagrams, together with further path qualities yielded by retrospective experience sampling, can inform the mapping between subjective experience—the flow of the phenomenal field—and its neural bases.

This paper has illustrated the approach by considering the neurophenomenal field of a hypothetical bilingual speaker reflecting on a matter of cognitive and affective significance for her in two contexts: one as she talked to a friend and a second as she thought about it silently. The matter of concern (the topic) was that she was an attriter in Cantonese, the language of her parents, and they spoke no English—a language in which she was proficient. Her path diagram indicated that by relearning Cantonese and forging a new self-identity (a Cantonese–American identity), she could resolve the cognitive–affective nexus. I treated mental simulation as a fundamental process in envisaging an alternative state of affairs from the present, and allowed the nesting of other processes (e.g., ones linked to the retrieval of personal memories) within it. By mapping such processes to high-level attentional network (such as the default mode network), it is possible to conjecture the neural bases of her talk. Retrospective experience sampling after the scanning allows such conjectures to be tested and refined. A third context envisaged bilinguals from either a common or a distinct background listening to her talk. Using the same basic procedures (path diagrams, retrospective experience sampling) allows assessment of the neurophenomenal fields of different bilingual speakers in response to an identical input.

Path diagrams aim to capture the psychological factors and their relations, as perceived by the bilingual, that bear on a key issue. The subjective flow of the phenomenal field is induced by resolving the cognitive–affective nexus through an intervention. It is a driving vector that can alter the polarity of the relations in the field—changing a negative relation, for example, to a positive one. Interventions issue in particular kinds of utterance (e.g., If I do X then Y will follow), and reflect the individual's understanding or model of their world.

Which kinds of interventions are considered are likely to depend not only on the polarity of the relations expressed in the path diagram, but also on their strength. The strength of a relation may commend it as a candidate intervention likely to have the greatest impact. Retrospective experience sampling, I have suggested, is a way to profile the likely multidimensional nature of the paths, and so amplify our ability to capture the neurophenomenal field. Other methods, involving more detailed enquiry with actual speakers (e.g., using Interpretative Phenomenological Analysis—Smith et al. 2009) may reveal phenomenological fields of greater complexity. The empirical question is then the extent to which the increased richness derived from such methods further enhances the explanation of the neural dynamics. It would also be valuable to have a more formal characterization of subjective experience in order to explore this question. The pioneering work of Székely (1965) offers a possible formalism for capturing the affective and cognitive content of inner experience.

Perhaps it goes without saying that just because a speaker recognizes a matter of concern does not entail that they wish to do anything about it. That decision depends on their stance with respect to the issue. Self-talk may yield acceptance of the state of affairs (che sera, sera). For the hypothetical individual, I considered such a stance was not part of the narrative of herself.

I illustrated a neurophenomenal approach with respect to a particular topic related to language use in a speaker who was an attriter in the language of her parents. Such diagrams also offer a way to identify individual differences in the subjective demands of learning a language later in life. Consider the urgent demand to learn a new language experienced by refugees from war living in a new culture. Such a demand will strongly characterize the cognitive–affective nexus of their phenomenal field. Allowing individuals to depict the key psychological factors that they perceive as stressing and de-stressing offers a way to identify potential interventions to support their language learning. Such interventions (e.g., physical exercise; joint music-making) may not be directly related to language learning, but directed at reducing the effects of stress on the network of cortical and subcortical regions involved in the cognitive control required for language learning (Green 2018; and see Sousa 2016, for a synoptic review of the effects of stress on the cortical and subcortical networks involved in cognitive control.) If successful, such efforts will be reflected in changes in the neurophenomenal field of such bilinguals.

Language concerns also arise for proficient speakers. As noted in the Introduction, a felt disjointedness when discussing feelings in one language rather than another may also merit an intervention in the phenomenal field (e.g., Wierzbicka 2004). In such circumstances, only a recognition of the fact of disjointedness may offer a way to circumvent irritation. Figure 2a,b below proposes an intervention in which the person shifts their pragmatic frame. When they use English, rather than their native tongue (Polish), they "act in English", meaning that they feel but do not worry about the discrepancy in emotional nuance, allowing for covering phrases such as "it's not quite how I'm feeling" to signal this discrepancy. It would be interesting to know whether such reframing alters the neurophenomenology of language use.

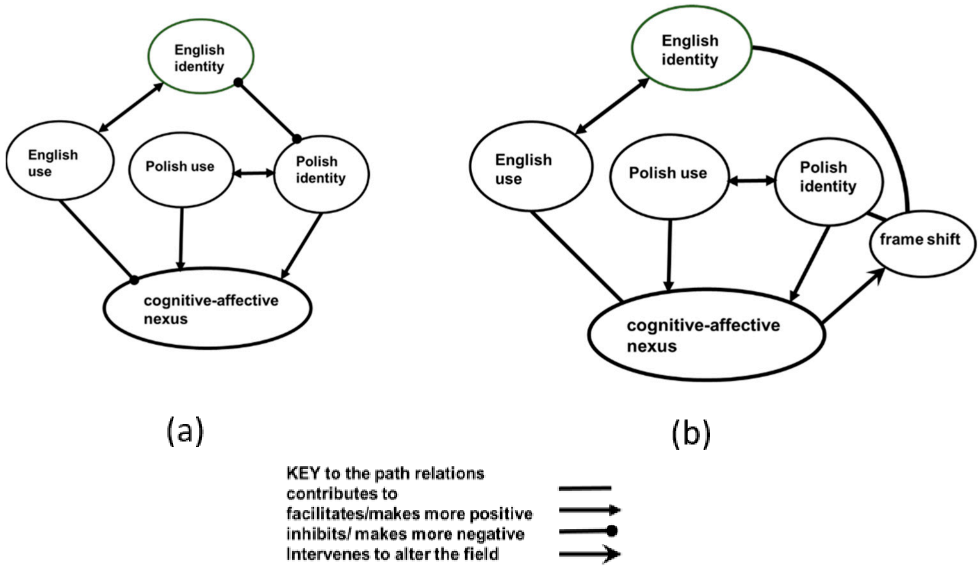

Figure 2. A frame shift intervention designed to accept the perceived incomparability of language use suffused by distinct cultural identities. (**a**) before the intervention, (**b**) after the intervention (please see text for exposition).

Depicting phenomenal fields in path diagrams is widely applicable beyond issues concerned with language use as such. Conversations cover a range of topics, and speakers may have a range of issues of concern that they wish to discuss. Understanding the neural basis of conversation between bilingual speakers requires not only tracking neural activity in language-eloquent networks and tracking the large-scale attention networks as different topics are covered (Green 2019), but also tracking the thoughts and feelings as such topics are discussed. Path diagrams together with experience sampling offer a way to do so.

In conclusion, and in line with Varela's (1996) exhortation for the necessity of neurophenomenology in the study of human consciousness, I hope I have provided grounds for believing that with the use of path diagrams, on-line scanning and experience sampling, we can explore the neural bases of subjective states during conversations between bilingual speakers, ushering in a neurophenomenology of the bilingual speaker.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I thank Judith Kroll and two anonymous reviewers for their comments on an earlier version of this paper.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

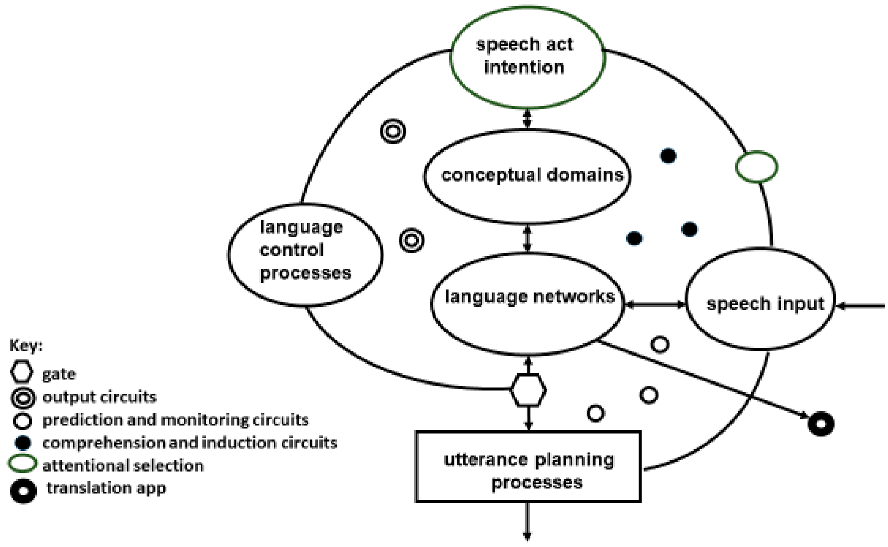

Figure A1. Mapping the speech act when a word or phrase is unknown in the target language (adapted from Green 2019).

Appendix A sketches the mapping of a speech act into overt speech in the context of conversation. It depicts the intimate relation between speech production processes and the comprehension of another's speech and own speech. The intention to speak about a particular topic recruits knowledge represented in a conceptual domain, and drives the activation of lexical concepts and constructions in functionally separable language networks. Outputs from the networks are gated for entry into the planning process as per language control processes (see Green 2019, for further details). Where no word or phrase is available in the target language, and switching between languages is precluded because there is no shared language, the speaker uses an external translation app. The path indicated in Figure 1 is that app use is initiated once the monitoring circuits detects the problem. Output from the app becomes speech input that is mapped into the language network before release via the gate into the utterance plan.

References

Andersen, Ditte, Signe Ravn, and Rachel Thomson. 2020. Narrative sensemaking and prospective social action: Methodological challenges and new directions. *International Journal of Social Research Methodology* 23: 367–75. [CrossRef]

Barber, Liam, Renate Reniers, and Rachel Upthegrove. 2021. A review of functional and structural neuroimaging studies to investigate the inner speech model of auditory verbal hallucinations in schizophrenia. *Translational Psychiatry* 11: 582. [CrossRef]

Beatty-Martínez, Anne L., Christian A. Navarro-Torres, Paola E. Dussias, María Teresa Bajo, Rosa E. Guzzardo Tamargo, and Judith F. Kroll. 2020. Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 1022–47. [CrossRef]

Billig, Michael. 1987. *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge: Cambridge University Press.

Chen, Jingyuan E., Mikail Rubinov, and Catie Chang. 2017. Methods and considerations for dynamic analysis of functional MR imaging data. *Neuroimaging Clinics of North America* 27: 547–60. [CrossRef]

Cohen-Zimerman, Shira, Harsh Khilwani, Gretchen N. L. Smith, Frank Krueger, Barry Gordon, and Jordan Grafman. 2021. The neural basis for mental state attribution: A voxel-based lesion mapping study. *Human Brain Mapping* 42: 65–79. [CrossRef] [PubMed]

Craik, Kenneth. 1943. *The Nature of Explanation*. Cambridge: Cambridge University Press.

Costa, Albert, Alice Foucart, Inbal Arnon, Melina Aparici, and Jose Apesteguia. 2014. "Piensa" twice: On the foreign language effect in decision making. *Cognition* 142: 236–54. [CrossRef] [PubMed]

Fleming, Stephen M., and Raymond J. Dolan. 2012. The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B* 367: 1338–49. [CrossRef] [PubMed]

Gonzalez-Castillo, Javier, Julia W. Y. Kam, Colin W. Hoy, and Peter A. Bandettini. 2021. How to interpret resting-state fMRI: Ask your participants. *Journal of Neuroscience* 41: 1130–41. [CrossRef] [PubMed]

Grandchamp, Romain, Lucie Rapin, Marcela Perrone-Bertolotti, Céline Pichat, Célise Haldin, Emile Cousin, Jean-Philippe Lachaux, Marion Dohen, Pascal Perrier, Maëva Garnier, and et al. 2019. The ConDialint Model: Condensation, Dialogality, and Intentionality dimensions of inner speech within a hierarchical predictive control framework. *Frontiers in Psychology* 10: 2019. [CrossRef]

Green, David W. 1997. Explaining and envisaging an ecological phenomenon. *British Journal of Psychology* 88: 199–217. [CrossRef]

Green, David W. 2001. Understanding microworlds. *Quarterly Journal of Experimental Psychology A* 54: 879–901. [CrossRef]

Green, David W. 2011. Arguments in mind. In *Science of Reason: A Festschrift for Jonathan St.B.T. Evans*. Edited by Ken Manktelow, David E. Over and Shira Elqayam. Brighton: Psychology Press, pp. 53–61.

Green, David W. 2018. The interactional challenge: L2 learning and use in the third age. In *Third Age Learners of Foreign Languages*. Edited by Danuta Gabryś-Barker. Bristol: Multilingual Matters, chp. 3. pp. 31–47.

Green, David W. 2019. Language control and attention during conversation: An exploration. In *The Handbook of the Neuroscience of Multilingualism*, 1st ed. Edited by John W. Schwieter. Hoboken: John Wiley & Sons Ltd., chp. 21. pp. 427–46.

Green, David W., and Ian C. McManus. 1995. Cognitive structural models: The perception of risk and prevention in coronary heart disease. *British Journal of Psychology* 86: 321–36. [CrossRef]

Green, David W., Ian C. McManus, and Jay Derrick. 1998. Cognitive structural models of unemployment and employment. *British Journal of Social Psychology* 37: 415–38. [CrossRef]

Higgins, Cameron, Yunzhe Liu, Diego Vidaurre, Zeb Kurth-Nelson, Ray Dolan, Timothy Behrens, and Mark Woolrich. 2021. Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron* 109: 882–93. [CrossRef]

Jääskeläinen, Iiro P., Jyrki Ahveninen, Vasily Klucharev, Anna N. Shestakova, and Jonathan Levy. 2022. Behavioral experience sampling methods in neuroimaging studies with movie and narrative stimuli. *Frontiers of Human Neuroscience* 16: 813684. [CrossRef]

Jachs, Barbara. 2021. The Neurophenomenology of Meditative States: Introducing Temporal Experience Tracing to Capture Subjective Experience States and Their Neural Correlates. Doctoral thesis, University of Cambridge, Cambridge, UK. [CrossRef]

Johnson, Sterling C., Leslie C. Baxter, Lana S. Wilder, James G. Pipe, Joseph E. Heiserman, and George P. Prigatano. 2002. Neural correlates of self-reflection. *Brain* 125: 1808–14. [CrossRef] [PubMed]

Johnson-Laird, Philip N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.

Jones, Simon R., and Charles Fernyhough. 2007. Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration. *Clinical Psychological Review* 27: 140–54. [CrossRef]

Kroll, Judith F., Paola E. Dussias, Kinsey Bice, and Lauren Perrotti. 2015. Bilingualism, mind, and brain. *Annual Review of Linguistics* 1: 377–94. [CrossRef]

Liao, Jenny. 2021. Forgetting my first language. *The New Yorker*, September 3. Available online: https://www.newyorker.com/culture/personal-history/forgetting-my-first-language (accessed on 5 October 2022).

Liu, Yichuan, Elise A. Piazza, Erez Simony, Patricia A. Shewokis, Banu Onaral, Uri Hasson, and Hasan Ayaz. 2017. Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Scientific Reports* 7: 43293. [CrossRef]

Mead, George H. 1959. *The Philosophy of the Present*. La Salle, IL: Open Court.

Merleau-Ponty, Maurice. 1964. *The Primacy of Perception*. Evanston, IL: Northwestern University Press.

Nikolic, Milena, and David A. Lagnado. 2015. There aren't plenty more fish in the sea: A causal network approach. *British Journal of Psychology* 106: 564–82. [CrossRef] [PubMed]

Pavlenko, Aneta. 2017. Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Current Opinion in Psychology* 17: 74–78. [CrossRef] [PubMed]

Pearl, Judea. 2000. *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

Pearl, Judea, and Dana Mackenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Price, Cathy J., and Karl J. Friston. 2005. Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology* 22: 262–75. [CrossRef] [PubMed]

Quinn, Andrew J., Diego Vidaurre, Romesh Abeysuriya, Robert Becker, Anna C. Nobre, and Marl W. Woolrich. 2018. Task-evoked dynamic network analysis through Hidden Markov modelling. *Frontiers in Human Neuroscience* 12: 603. [CrossRef]

Raichle, Marcus E. 2015. The brain's default mode network. *Annual Review of Neuroscience* 38: 433–47. [CrossRef] [PubMed]

Sheldon, Signy, Can Fenerci, and Lauri Gurguryan. 2019. A neurocognitive perspective on the forms and functions of autobiographical memory retrieval. *Frontiers in Systems Neuroscience* 13: 4. [CrossRef]

Smith, Jonathan A., Paul Flowers, and Michael Larkin. 2009. *Interpretative Phenomenological Analysis*. London: Sage.

Sousa, Nuno. 2016. The dynamics of the stress neuromatrix. *Molecular Psychiatry* 21: 302–12. [CrossRef]

Spreng, R. Nathan, and Cheryl L. Grady. 2009. Patterns of brain activity supporting autobiographical memory, prospection, and Theory of Mind, and their relationship to the Default Mode Network. *Journal of Cognitive Neuroscience* 22: 1112–23. [CrossRef] [PubMed]

Stephens, Greg J., Lauren J. Silbert, and Uri Hasson. 2010. Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences USA* 107: 14425–30. [CrossRef] [PubMed]

Székely, Edmond. 1965. *Basic Analysis of Inner Psychological Functions. British Journal of Psychology, Monograph Supplements (L. S. Hearnshaw, Ed.) XXXVII*. Cambridge: Cambridge University Press.

Tibon, Roni, Kamen A. Tsvetanov, Darren Price, David Nesbitt, Can A. N. Cam, and Richard Henson. 2021. Transient neural network dynamics in cognitive ageing. *Neurobiology of Aging* 105: 217–28. [CrossRef]

Tierney, Tim M., Niall Holmes, Stephanie Mellor, José D. López, Gillian Roberts, Ryan M. Hill, Elena Boto, James L. Legett, Vishal Shah, Matthew J. Brookes, and et al. 2019. Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography. *NeuroImage* 199: 598–608. [CrossRef]

van Hoeck, Nicole, Patrick D. Watson, and Aron K. Barbey. 2015. Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience* 9: 420. [CrossRef]

Varela, Francisco. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies* 3: 330–49.

Vygotsky, Lev. 1981. The genesis of higher mental functions. In *The Concept of Activity in Social Psychology*. Edited by J. Wertsch. Armonk, NY: Sharpe, pp. 144–88.

Wei, Li. 2018. Translanguaging as a practical theory of language. *Applied Linguistics* 39: 9–30. [CrossRef]

Wierzbicka, Anna. 2004. Preface: Bilingual lives, bilingual experience. *Journal of Multilingual and Multicultural Development* 25: 94–104. [CrossRef]

Wiliford, Kenneth, Daniel Bennequin, Karl Friston, and David Rudrauf. 2018. The projective consciousness model. *Frontiers in Psychology* 9: 02751. [CrossRef]

languages

MDPI

Article

The Nature and Function of Languages

Franco Fabbro [1,*], Alice Fabbro [2] and Cristiano Crescentini [1]

1 Department of Languages and Literatures, Communication, Education, and Society, University of Udine, 33100 Udine, Italy
2 School of Psychology and Education, Free University of Brussels, 1050 Brussel, Belgium
* Correspondence: franco.fabbro@uniud.it

Abstract: Several studies in philosophy, linguistics and neuroscience have tried to define the nature and functions of language. Cybernetics and the mathematical theory of communication have clarified the role and functions of signals, symbols and codes involved in the transmission of information. Linguistics has defined the main characteristics of verbal communication by analyzing the main tasks and levels of language. Paleoanthropology has explored the relationship between cognitive development and the origin of language in *Homo sapiens*. According to Daniel Dor, language represents the most important technological invention of human beings. Seemingly, the main function of language consists of its ability to allow the sharing of the mind's imaginative products. Following language's invention, human beings have developed multiple languages and cultures, which, on the one hand, have favored socialization within communities and, on the other hand, have led to an increase in aggression between different human groups.

Keywords: communication; symbols; neural recycling; cultural identities

Citation: Fabbro, Franco, Alice Fabbro, and Cristiano Crescentini. 2022. The Nature and Function of Languages. *Languages* 7: 303. https://doi.org/10.3390/languages7040303

Academic Editors: John W. Schwieter, Raquel Fernández Fuertes and Juana M. Liceras

Received: 16 May 2022
Accepted: 22 November 2022
Published: 28 November 2022

1. Introduction

For over two thousand years philosophers, theologians and poets have reflected on the nature of language (Heidegger 1959; Panikkar 2007). More recently, scientific disciplines, from linguistics to computer science, have also sought to clarify its characteristics and functions (Sapir 1921; Jakobson and Waugh 1979; Borden et al. 2006). Nevertheless, in spite of thousands of books and articles, both theoretical and experimental, the nature of language still remains rather enigmatic (Lieberman 2013; Scott-Phillips 2015; Corballis 2017a, 2017b).

In the *Encyclopedia Britannica*, language is defined as a symbolic system (composed of sounds, hand gestures or letters) created by a social group to facilitate human expression. According to this perspective, the functions of language include communication, cultural identity, play and imaginative and emotional expression (Britannica n.d.). The Italian encyclopedia *Treccani* also specifies that language is an exclusive faculty of human beings allowing for the expression of consciousness' contents through a conventional symbolic system (Treccani n.d.). These definitions, albeit general, highlight characteristic aspects of language: its communicative function, its symbolic nature and its ability to express and share consciousness' contents.

2. Language as a Communication System

Communication is a particular form of transport in which what is moved is neither matter nor energy but "information" (Escarpit 1976). Yet, information cannot exist without a material substrate, though it is not reducible to it (Longo 1998). For example, sending a telegram to New York saying "*all right*" would require the same amount of energy and matter as sending "*gar lilth*" instead. Both telegrams are composed of two words and the same eight letters arranged in a different order. Transmitting either message would require an equal amount of energy and matter; although, only the first would convey understandable information. Thus, information is to be regarded as something altogether different from its material supports (Wiener 1948).

Every communication system, be it verbal expression, telephone lines, radio channels or the internet, consists of at least three elements: an *emitter* (source or transmitter), a *receiver* and the *channel* carrying information from the emitter to the receiver. The information source (emitter) selects the "message" and encodes it in a "code" that is a set of "signs" or "symbols" which are represented by "signals" compatible and specific to the channel in use (Pierce 1961; Singh 1966). For example, the channel of a telegraph can only transmit electric current impulses of two possible durations: a short impulse (dot) and a long impulse (line). Accordingly, all letters of the alphabet have to be codified as a series of symbols made up of sets of lines and dots before being transmitted (Singh 1999).

To date, information's nature remains rather elusive (Dodig-Crnkovic and Burgin 2019). Information plays an essential role in the functioning of living systems (cells, organisms and societies) and in the regulation of some man-made devices, in particular, in self-regulating systems such as thermostats and electronic processors. Norbert Weiner, in one of the earliest definitions of information, describes it as being neither matter nor energy (Wiener 1948, p. 132; Montagnini 2015). According to Gregory Bateson (1972), information has to do with the notion of "difference": for example, maps are representations of territorial differences (roads, hills, mountains, cities). If all land was alike and presented no discernible features, there would be no information. Giuseppe Longo has thus proposed to define information as a "difference that generates a difference to somebody" (Longo 1998; Longo and Vaccaro 2013). Such definition necessarily implies the existence of an observer (man or machine) able to detect and/or produce differences.

Transferring information through a channel while limiting the interference of "noise", which tends to distort and shatter messages, is considered a crucial aspect of communication by mathematicians and engineers alike. In the monograph entitled *The Mathematical Theory of Communication* (1949), Claude Shannon and Warren Weaver maintain that the fundamental engineering problem of communication is to reproduce in one point (A), in an exact, or more or less approximate way, a message originating in another point (B) (Shannon and Weaver 1949). Such two points can be separated in "space" or "time". In this sense, communication then not only concerns the spatial transfer of information, but also its storage on physical supports.

Shannon was concerned with the technical aspects of communication, and in order to make mathematical observations, he isolated information from its semantic contents (Longo 1998, p. 28; Longo and Vaccaro 2013, pp. 22–23). In such a manner it was possible to relate it to the concepts of uncertainty and entropy. The level of information depends on the number of possible messages: if only one message is possible, there is no uncertainty and therefore no information. Moreover, information is related to probability, therefore, to a certain element of surprise. The smaller a message's likelihood, the more informative it is. By elaborating on these concepts, Shannon established an equation to quantify the amount of information contained in a message, regardless of its meaning. This quantity was related to the average logarithm of the improbability of the message. It was basically a measure of its unpredictability (Gleick 2011).

Establishing that information was not a concept of the physical discipline, and could not be related to either matter or energy, called for the invention of a new unit of measurement in order to define it. Shannon had the brilliant intuition to think of information as something allowing to answer a question with either a "yes" or a "no". The answer of such question can take one of two possible values: "1" (yes) or "0" (no). Thus, each question corresponds to a *Binary Digit* or a bit. For example, it is possible to precisely define a number between one and one hundred by asking a series of questions that can be answered with "yes" or "no". "Is the number greater than 50?" (Yes). "Is the number lesser than 75?" (No), and so on. According to this perspective, any question with a possible answer can be codified in sequences of "1" and "0" and can thus be measured in bits. Since language is a set of symbols (phonemes or letters of the alphabet), it too can be rendered in a series of bits.

For example, as each letter of the English alphabet can be coded in five bits, an average book of about four hundred letters approximately contains two million bits of information.

Conceptualizing language as a communication system (*Code Model*) is a major model of linguistic analysis. However, as language is not only a system for the transmission of information but performs other functions, it can be examined through other analytical frameworks (Scott-Phillips 2015).

3. Signs, Symbols and Codes

Information sources emit message units. These units are exchanged by means of a code (such as the genetic code, the alphabetic code, the Morse code, etc.) between the emitter (encoder) and the receiver (decoder) (de Saussure 1922). The code is a list of units, called symbols, that constitute the message. Symbols are signs in which the relationship between what is being represented and its representation is arbitrary (Mazzone 2005; Deacon 1997). The word *sign* derives from the Latin "*signum*" and describes "something referring to something else" (Peirce 1931–1935), as claimed medieval philosophers: *alquid stat pro aliquo* (Bettetini 1963; Mazzone 2005). The word *symbol* derives from the Greek *symbolon* (σύμβολον) which means: "to cast together", "to connect", "make coincide". The symbol for the Greeks originally designated an object, a tile, a fragment of ceramic or metal that was divided in the stipulation of an economic, emotional or spiritual contract. Each party kept a piece in token of their agreement. Upon meeting, the fragments of the *symbolon* were brought together to honor the bond and commemorate the economic, emotional or spiritual ties uniting them (Mazzone 2005). Aristotle emphasized above all the conventional and relational aspect of language's symbols (words) (*De Interpretatione*, (Aristotle 1962)). Symbols are entities in relation with one another, and for this reason the meaning of a word opposes and differs from the words surrounding it.

Each symbol of language has two faces: the signifier and the signified. The mathematical theory of communication has been interested above all in the signifier and its characteristics: encoding and decoding, resistance to noise, speed of transmission. To return to our earlier example, sending a message via telegraph requires that the letters of the alphabet be encoded in the symbols of the Morse code. The individual letters transformed into Morse symbols are then sent through a series of short (dot) and long (line) pulses, minimizing the effects of noise along the telegraph line. Other areas of general communication theory, such as semantics and pragmatics, are more interested in aspects related to meaning and the communicative context (Longo 2001). A fundamental aspect, which is often forgotten, is that at the origin and at the end of most of the experiences of coding or decoding a message, we always find "language". It is a very particular form of communication that presupposes thinking and speaking individuals, who have tacitly agreed upon an interpretative code among themselves through an action of social coordination (Singh 1966; Escarpit 1976; Mazzone 2005).

In a series of reflections developed in the biophysical context, Howard Pattee analyzed the most significant characteristics present in cultural (such as language) and biological (such as the genetic code) symbolic systems (Pattee 2008, 2015; Pattee and Kull 2009; Pattee and Rączaszek-Leonardi 2012). According to Pattee, "symbols" are "formal entities that stand for something else" (representing something else).

The first defining characteristic of symbols is that they are constituted by physical structures. In fact, all codes, rules and even the most abstract descriptions related to the symbolic dimension have well-defined physical bases. For example, DNA is formed by nucleotides, phonemes are made up of sound signals and even the letters of the alphabet consist of signs drawn on paper. The second characteristic concerns the aspect of reproducibility. The symbolic structures can be transmitted through their replication, which is articulated first in a process of "reading" and then in a process of "copying". The third characteristic refers to arbitrariness. All the symbolic structures, aside from being informative (highly improbable), present a complete arbitrariness between the "signifier" (phonemes, letters, nucleotides) and the "signified". The arbitrary relationship between symbols and their meaning depends on the history of that particular symbolic system (the history of languages or the history of the genetic code). Indeed, at the lowest level of organization, no symbol carries any meaning (i.e., no phoneme or no nucleotide refers to anything significant). The effects of the symbols are highlighted within a dynamic system capable of generating more or less complex structures. Symbolic systems are historically determined and represent "memories of possibility"; they are coordinated adaptations with biological, psychological or social reality.

4. The Symbolic Nature of Language

Compared to DNA and psyche, language is the symbolic domain that is most "external" (Fabbro 2021a, 2021b). It is a symbolic system constituted by several layers. At the most superficial level, sounds are symbols for phonemes. In turn, sequences of phonemes are symbols for words, while strings of words form symbols for sentences and, finally, ordered chains of sentences constitute symbols for narratives and stories.

All languages are made up of symbolic systems nested within one another and sharing some universal properties (Hockett 1960). The first is the *duality of structure*: all human languages use meaningless signs (phonemes), which combine into ordered sequences of sounds with meaning (words). The second characteristic refers to *arbitrariness*, which is one of the fundamental concepts of all symbolic systems. In language, this concept indicates that there is no physical similarity between a symbol (word) and the object that it represents. Within a linguistic community, phonemes, words and numerous grammatical aspects are passed on from one generation to the next (*transmission*) through usage and learning (de Saussure 1922; Thorpe 1972).

Like DNA, languages have a *linear code*. This implies that every verbal expression is made up at the most superficial level of a string of words that, at the deepest level, presents a hidden structure (syntax). Linearity manifests a fundamental aspect of spoken languages, namely the temporal arrangement of acoustic signals along a timeline (de Saussure 1922). An additional property of languages, which is also shared by DNA, is *discreteness*. The sound symbols of a language (phonemes) are represented by a discrete number of elements (in Italian = 30; in English = 44). The same can be said of words; there are no intermediate sentences between a sentence composed of "n" words and a sentence composed of "n + 1" words (Moro 2006).

Another important feature is *recursiveness*, which is the ability to infinitely repeat a process within the same structure. Since a sentence can consist of a nominal syntagm (SN) plus a verbal syntagm (VS) [S = NS + SV], and the verbal syntagm can consist of a verb plus a nominal syntagm or a verb plus another sentence [VS = V + NS or VS = V + S], it is possible to recursively expand a sentence by inserting another sentence at the level of the verbal syntagm. For example, "Marco is wearing a sweater" + "The sweater cost 100 euros" = "Marco is wearing a sweater that cost 100 euros". Recursiveness accounts for another property of languages: *openness*, that is, the possibility of producing sentences that are always new or have never been uttered before. As a result of recursiveness and openness, every speaker can generate an almost infinite number of sentences.

The rules that determine the organization of words within a sentence are called 'syntax'. Each language has specific syntactic rules. For some, such as Latin, word order has only rhetorical significance. In fact, in Latin there is little difference between the sentences: "*hominem videt femina*" or "*femina videt hominem*", while in English the order of the words is very important: the sentence "the child eats the chicken" means something very different from the sentence "the chicken eats the child" (Sapir 1921). Within human languages, there is an enormous variety of phonemes, words and grammatical rules. Nevertheless, there exist some restraints in their choice and implementation.

The repertoire of phonemes that humans can produce is limited. The phonoarticulatory organs and the nerve structures that coordinate them have structural and physiological constraints. Likewise, it is not possible to pronounce words that are too long (e.g., composed of 500 phonemes) because at a certain point the air emitted during exhalation runs out.

Based on these premises, many linguists, starting from Noam Chomsky, have argued that there are limits also in the possible syntactic rules. Therefore, the set of syntactic rules of all languages would be delimited by a system of categories, mechanisms and constraints called "universal grammar". Universal grammar seems to be related to the ways in which the brain develops, organizes and functions, which in turn are related to specific genetic information that has probably evolved over hundreds of thousands of years within a context of a musical nature (glossolalic singing) (Mithen 2005; Patel 2008; Fabbro 2018).

5. The Invention of Language

Humans belong to the class of *Mammals* and the order of *Primates*, which emerged about 80 million years ago and comprises about 400 species, including prosimians (tarsiers and lemurs), monkeys and anthropomorphic apes. The Hominid family, which includes our species, separated from the latter about 5–7 million years ago. The most significant characteristic that distinguishes hominids from other primates is the bipedal gait (Manzi 2017). Adapting to bipedal locomotion brought about a series of anatomical modifications, concerning the conformation of the lower limbs and feet, and important physiological transformations at the level of the respiratory system and central nervous system. The bipedal gait modified the respiratory rhythm allowing for an extended expiratory phase, a fundamental requirement to develop the ability to laugh, sing and speak (Provine 2000).

Homo sapiens, the only extant species of hominids, appeared in Africa about 300,000 years ago. Fossil remains attributed to *H. sapiens* have been found in Ethiopia (dated 195,000 years ago) and Morocco (dated about 300,000 years ago). The modern human presents a more gracile and slender structure than the Neanderthal man, with a brain volume of about 1400 cubic centimeters in *H. sapiens* and about 1600 cc in H. neanderthalensis. Paleoanthropologists believe that modern humans seem to have come from an eastern or southeastern region of Africa and then spread across the continent. They subsequently migrated, in several waves, to Eurasia and the rest of the world (Tattersall 2012).

For a long period of time, more than 200,000 years, H. sapiens did not produce any technological innovations: its lifestyle and style in manufacturing lithic tools was similar to that of other extant hominid species (*H. erectus*, *H. heildebergensis*, *H. neanderthalensis*, Denisova man). The only significant distinction consisted in a different organizational structure of human villages. In fact, modern humans present, in all hunter-gatherer cultures studied, a more numerous social structure than that of the other hominids. About 80–90 thousand years ago, *H. sapiens* began to manifest great creativity by producing ornaments, decorations and complex tools that involved the development of symbolic thinking. This was unprecedented. Numerous anthropologists, psychologists and linguists have wondered about the causes behind this qualitative leap in culture and technology. Some have concluded that what brought about this *cognitive revolution* was most probably the appearance of articulated language (Tattersall 2012; Lieberman 2013; Corballis 2015; Chomsky 2016).

Not much seems to have changed at the genetic, anatomical and physiological levels in *H. sapiens* since its appearance 80,000 years ago. For this reason, some neuroscientists and paleoanthropologists, including Israel Rosenfield and Ian Tattersall, have argued that a group of *H. sapiens* located in a southeastern region of Africa probably "invented" articulate language (Rosenfield 1992; Tattersall 2012; Fabbro 2018). Evidently this was not a "conscious invention", but rather a kind of social game probably developed by a sufficiently large group of children who lived together for a few generations. The hypothesis that articulate language was invented and did not evolve, as many biologists, linguists and psychologists have long argued (Pinker 1994; Dunbar 2014; Corballis 2002, 2017a), is supported by the recent discovery that at least two languages emerged from nowhere in a group of children in Nicaragua and in a community of Bedouins in the Negev desert (Senghas et al. 2004; Senghas 2005).

Nicaraguan sign language is the first language invented by a group of children to have been thoroughly studied (Tattersall 2012; Bausani 1974; Fabbro 1996). Towards the end of the 1970s, after the victory of the Sandinista revolution, a special school dedicated to the education of deaf and mute children was established in Managua, the capital of Nicaragua. Initially, 50 children were brought together, joined by more than 200 in 1981, a number that gradually grew in the years that followed. The school aimed to teach deaf–mute children to lip-read the Spanish language, a goal that was not achieved. Instead, just within the span of two generations, the children spontaneously invented Nicaraguan Sign Language.

The first generation shared a set of signs that they had developed in the domestic context of their families (*homesigns*). These signs did not represent an actual sign language yet, but rather a form of gestural communication. In contrast, the second generation of deaf children, younger in age, was able to develop a grammatically complete language on the basis of the signs shared by children of the previous generation. The younger children were able to categorize the gestures of the older students, generating a true grammar, i.e., a set of abstract categories capable of regulating the relationships between different symbolic units. Only younger children were able to transform "gestures" into "symbols". This confirms the hypothesis that a language can only be "acquired" in a complete form or "invented" by children who have not yet reached puberty (Senghas et al. 2004; Fabbro 2004).

The invention of articulated verbal communication (language) by (deaf–mute) children indicates that language has not evolved but has been invented (Rosenfield 1992). This means that some biological bases of language (concerning phonology and syntax) have evolved within vocal behaviors that are much more archaic than language, such as glossolalic singing. Spoken language is the most important technology serving the transmission of mental content. Other systems used for the transmission of mental contents are written language and mathematics. Overall, these all are technical and cognitive skills that, on the basis of pre-adaptation (*exaptation*) phenomena, have conquered brain territories that originally evolved to compensate for other functions (Dehaene and Cohen 2007; Dehaene 1997, 2009; Fabbro 2018, 2021a).

6. What Is Language's Purpose?

According to many authors, language is a form of communication (Miller 1975, 1987). However, it is possible to communicate effectively even without language: many animal species communicate very well even without speech. Recently, Daniel Dor argued that language is a technology aimed at sharing imagination (Dor 2014, 2015, 2016). According to this perspective, the task of the speaker is to provide clues about their own mental representations, while the addressee tries to reconstruct the mental representations of the speaker through a chain of interpretative processes (Scott-Phillips 2015). In fact, for every "literal meaning" of a word or a sentence, there are infinite possible modulations of meaning (also related to pragmatic aspects). This fact determines one of the most typical characteristics of language, namely the "pervasiveness of indeterminacy" (Scott-Phillips 2015). This is a limitation given that verbal expression does not allow a direct (literal) grasp of reality; at the same time, it allows a varied range of interpretative possibilities.

Like all technologies, language is a system for achieving a purpose, namely the construction of a network of psychic individualities that exchange the contents of their imaginations. It is an unconventional type of technology, similar to money, contracts and legal systems. Since human beings are social organisms, language's invention and its development were accomplished as collective processes. Individual minds can be viewed as the "nodes" of a metaphorical Web, in which language constitutes the software that each human being downloads into their own mind and uses to help achieve an imaginative community (Barabási 2002; Dor 2015).

Wilhelm von Humboldt (1836) was one of the first linguists to emphasize the central role that imagination plays within language. In his opinion, human languages are not tools for naming objects that have already been thought, but rather organs used for the formulation of thought. Similarly, Daniel Dor considers language as the "mother of all inventions" (Dor 2015). According to this perspective, it was the invention of language that really gave rise to human beings. Thus, the first and most important technological product of human beings is language. At the same time, language has changed us radically.

The unintentional invention of language has occurred repeatedly throughout human history (Fabbro 2018). Deaf and mute children in Nicaragua who spontaneously invented a new sign language were able to experience the passage of the "magical frontier of language" and compare it to their previous lives. The experience was described as inconceivable, disconcerting and astounding; what struck these children most was the realization of the abysmal loneliness of their lives prior to entering the sphere of language (Schaller 1991; Senghas et al. 2004).

The relationships between language and technology are much closer than is commonly assumed. One aspect shared by both language and technology is recursiveness. In fact, all technological tools are constituted by components that are assembled according to a hierarchical structure. Technologies are comprised of technological components that contain smaller components within them, made up of even more elementary components (such as screws, vanes, bolts, etc.). Thus, modern technology is similar to a language that is open to the creation of new structures and functions (Arthur 2009). For these reasons, we believe that it is not a coincidence that the cultural and technological explosion that followed the cognitive revolution, some 80,000 years ago, is related to the most important invention made by humankind, namely the invention of language. Finally, some interesting evolutionary psychology studies, developed by Robin Dunbar (1996, 2014), have analyzed the possible role that vocal communication and language play in strengthening social bonds and reducing stress.

7. Languages as the Bedrock of Cultural Identities

It is likely that the first language split into several others just under a few hundred years after its invention. In fact, language can arise only from other human languages. The emergence of dialects is favored by a combination of geographic isolation and of linguistic variation (at phonological, syntactical and lexical levels) naturally occurring within generations (von Humboldt 1836; Locke and Bogin 2006). Languages facilitate relationships within the members of a group, yet they also contribute to the segregation of communities (Fabbro 1999; Pagel and Mace 2004; Pagel 2009). There is evidence suggesting that languages may act as biological barriers genetically isolating populations; indeed, the tendency to prefer partners speaking the same language still prevails among human beings (Spielman et al. 1974; Cavalli-Sforza 1996; Fabbro 1996, 1999; Lieberman 2013).

The acquisition of language, both in written and spoken form, sculpts the brain in specific ways and is affected by the existence of critical periods (or sensitive periods) (Fabbro 2004). Generally, it is possible to learn a second language well only before puberty and preferably before the age of seven. After the brain structures of the implicit memory system (particularly procedural memory) involved in the acquisition of language and syntax have matured, it is generally not possible to completely acquire a second language at the first language level (Paradis 1994, 2009; Cargnelutti et al. 2019). These observations suggest that the true "territory" of a language is not geographical, but rather neurological and mental (Fabbro 2018). Moreover, perfect language acquisition is only possible within early childhood.

The use of vocal signals, learned within specific critical periods, is a rather widespread biological phenomenon: it is present in many species of passerines (canaries, finches, etc.) and in some mammals, such as dolphins and killer whales (Riesch et al. 2012). In killer whales, the development of pod-specific cultural habits (related to singing and feeding) has mediated their speciation into numerous subspecies (Riesch 2016). Cultural and genetic isolation is one of the mechanisms for producing biological diversity, and diversity is the ground on which life originates and develops.

Language's advent determined the emergence of human cultures. Human beings within each culture have developed more or less different narratives, customs and traditions. The different tribes and populations of *H. sapiens* were rather technologically and traditionally homogenous before the invention of language. Its invention generated considerable diversity in ornaments, decorations and tattoos among different language groups. However, it is likely that linguistic and cultural diversity in hunter-gatherer societies also fostered an increase in violence (Fabbro et al. 2022).

Violence is a behavior of destructive and systematic aggression, aimed primarily at the elimination of isolated individuals or groups of the same species. In the second half of the last century, some ethologists have documented the presence of destructive behaviors between different groups of chimpanzees (Wrangham and Peterson 1996; Wrangham et al. 2006; Kelly 2005). These behaviors did not seem to be driven by food scarcity, but rather competition to access larger territories and greater food and sexual resources. In humans, language differences appear to have fueled the tendency for inter-group violence, as in almost all past human cultures those who spoke foreign languages were considered "subhuman beings" liable be subjugated or killed.

In fact, one of the most characteristic abilities of human beings, even of those who have not studied phonetics or phonology, consists of the ability to recognize from the pronunciation whether an individual belongs to their linguistic community or not. This ability, which seems to develop more during adolescence, is independent of the communication of information. According to this perspective, language plays a significant role as a marker of group identities (Locke and Bogin 2006; Ritt 2017). These aspects of human socialization and communication indicate that language, like many other aspects of cognition, has both strengths and limitations, which must be properly understood and regulated (Fabbro 2021a, 2021b; Fabbro et al. 2022).

Author Contributions: Conceptualization, F.F. and C.C.; writing—original draft preparation, F.F.; writing—review and editing, F.F., A.F. and C.C.; supervision, F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

Aristotle. 1962. *On Interpretation*. Translated by Harold P. Cook. Cambridge: Harvard University Press.

Arthur, W. Brian. 2009. *The Nature of Technology: What It Is and How it Evolves*. New York: Free Press.

Barabási, Albert-Laszlâo. 2002. *Linked: The New Science of Networks*. New York: The Perseus Book.

Bateson, Gregory. 1972. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Northvale: Jason Aaronson Inc.

Bausani, Alessandro. 1974. *Le lingue Inventate. Linguaggi Artificiali, Linguaggi Segreti, Linguaggi Universali [Invented Languages. Artificial Languages, Secret Languages, Universal Languages]*. Roma: Astrolabio.

Bettetini, Gianfranco. 1963. *Il segno. Dalla Magia al Cinema [The sign. From Magic to Cinema]*. Milano: Edizione i 7.

Borden, Gloria J., Katherine S. Harris, and Raphael Lawrence. 2006. *Speech Science Primer. Physiology, Acoustics and Perception of Speech*. Baltimore: William & Wilkins.

Britannica. n.d. Available online: https://www.britannica.com/topic/language (accessed on 7 April 2022).

Cargnelutti, Elisa, Barbara Tomasino, and Franco Fabbro. 2019. Language Brain Representation in Bilinguals With Different Age of Appropriation and Proficiency of the Second Language: A Meta-Analysis of Functional Imaging Studies. *Frontiers in Human Neuroscience* 13: 154. [CrossRef] [PubMed]

Cavalli-Sforza, Luigi Luca. 1996. *Geni, Popoli e Lingue [Genes, Peoples and Languages]*. Milano: Adelphi.

Chomsky, Noam. 2016. *Il Mistero del Linguaggio. Nuove Prospettive [The Mystery of Language. New Perspectives]*. Translated by di Maria Greco. Milano: Feltrinelli.

Corballis, Michael. 2002. *From Hand to Mouth: The Origins of Language*. Princeton: Princeton University Press.

Corballis, Michael. 2015. *The Wandering Mind: What the Brain Does When You're Not Looking*. Chicago: University of Chicago Press.

Corballis, Michael. 2017a. *The Truth about Language: What It Is and Where It Came From*. Chicago: University of Chicago Press.

Corballis, Michael. 2017b. Language Evolution: A Changing Perspective. *Trends in Cognitive Sciences* 21: 229–36. [CrossRef] [PubMed]

de Saussure, Ferdinand. 1922. *Course in General Linguistics*. Translated by Wade Baskin. New York: Columbia University Press.

Deacon, Terrence. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton & Company.

Dehaene, Stanislas. 1997. *The Number Sense*. New York: Oxford University Press.

Dehaene, Stanislas. 2009. *Reading in the Brain*. New York: Penguin.

Dehaene, Stanislas, and Laurent Cohen. 2007. Cultural recycling of cortical maps. *Neuron* 56: 384–98. [CrossRef] [PubMed]
Dodig-Crnkovic, Gordana, and Mark Burgin. 2019. *Philosophy and Methodology of Information*. New Jersey: World Scientific.
Dor, Daniel. 2014. The instruction of imagination: Language and its evolution as a communication technology. In *The Social Origins of Language*. Edited by Daniel Dor, Chris Knight and Jerome Lewis. Oxford: Oxford University Press, pp. 105–25.
Dor, Daniel. 2015. *The Instruction of Imagination: Language as a Social Communication Technology*. Oxford: Oxford University Press.
Dor, Daniel. 2016. From experience to imagination: Language and its evolution as a social communication technology. *Journal of Neurolinguistics* 43: 107–19. [CrossRef]
Dunbar, Robin Ian MacDonald. 1996. *Grooming, Gossip and the Evolution of Language*. Cambridge: Harvard University Press.
Dunbar, Robin Ian MacDonald. 2014. *Human Evolution*. London: Pelican Books.
Escarpit, Robert. 1976. *Teoria dell'informazione e della Comunicazione [General Information Theory and Communication]*. Translated by di Maria Grazia Rombi. Roma: Editori Riuniti.
Fabbro, Franco. 1996. *Il Cervello Bilingue. Neurolinguistica e Poliglossia [The Bilingual Brain. Neurolinguistics and Polyglossy]*. Roma: Astrolabio.
Fabbro, Franco. 1999. *The Neurolinguistics of Bilinguism: An Introduction*. Hove: Psychology Press.
Fabbro, Franco. 2004. *Neuropedagogia delle Lingue. Come Insegnare le Lingue ai Bambini [Neuropedagogy of Languages. How to Teach Languages to Children]*. Roma: Astrolabio.
Fabbro, Franco. 2018. *Identità culturale e violenza. Neuropsicologia delle lingue e delle Religioni [Cultural Identity and Violence. Neuropsychology of Languages and Religions]*. Torino: Bollati Boringhieri.
Fabbro, Franco. 2021a. *Che cos'è la psiche. Filosofia e neuroscienze [What Is the Psyche. Philosophy and Neuroscience]*. Roma: Astrolabio.
Fabbro, Franco. 2021b. *I Fondamenti Biologici della Filosofia [The Biological Foundations of Philosophy]*. Milano and Udine: Mimesis Edizioni.
Fabbro, Franco, Alice Fabbro, and Cristiano Crescentini. 2022. Neurocultural identities and the problem of human violence. In *Evil in the Modern World*. Edited by Laura Dryjanska and Giorgio Pacifici. Berlin/Heidelberg: Springer, pp. 131–36.
Gleick, James. 2011. *The Information. A History, A Theory, A Flood*. New York: Pantheon Books.
Heidegger, Martin. 1959. *On the Way to Language*. Translated by Peter Hertz. San Francisco: Harper.
Hockett, Charles. 1960. The origin of speech. *Scientific American* 203: 88–111. [CrossRef]
Jakobson, Roman, and Linda R. Waugh. 1979. *The Sound Shape of Language*. Bloomington: Indiana University Press.
Kelly, Raymond C. 2005. The evolution of lethal intergroup violence. *Proceedings of the National Academy of Sciences* 102: 15294–298. [CrossRef] [PubMed]
Lieberman, Philip. 2013. *The Unpredictable Species: What Makes Humans Unique*. Princeton: Princeton University.
Locke, John L., and Barry Bogin. 2006. Language and life history: A new perspective on the development and evolution of human language. *Behavioral and Brain Sciences* 29: 259–80. [CrossRef] [PubMed]
Longo, Giuseppe O. 1998. *Il nuovo Golem. Come il Computer Cambia la Nostra Cultura [The New Golem. How the Computer Changes Our Culture]*. Roma and Bari: Laterza.
Longo, Giuseppe O. 2001. *Homo Technologicus*. Roma: Meltemi.
Longo, Giuseppe O., and Andrea Vaccaro. 2013. *Bit Bang. La nascita della filosofia digitale [Bit Bang. The Birth of Digital Philosophy]*. Adria: Apogeo.
Manzi, Giorgio. 2017. *Ultime notizie sull'evoluzione umana [Latest News on Human Evolution]*. Bologna: Il Mulino.
Mazzone, Marco. 2005. *Menti simboliche. Introduzione agli studi sul linguaggio [Symbolic Minds. Introduction to Language Studies]*. Roma: Carocci.
Miller, George A. 1975. *The Psychology of Communication*. New York: Basic Books.
Miller, George A. 1987. *Language and Speech*. New York: W.H. Freeman & Company.
Mithen, Steven. 2005. *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. Cambridge: Harvard University Press.
Montagnini, Leone. 2015. Information versus matter and energy. La concezione dell'informazione in Wiener e le sue conseguenze sull'oggi [Information versus matter and energy. Wiener's conception of information and its consequences for today]. *Biblioteche oggi* 33: 41–61.
Moro, Andrea. 2006. *I confine di Babele. Il cervello e il mistero delle lingue impossibili [The Boundaries of Babel. The Brain and the Mystery of Impossible Languages]*. Milano: Longanesi.
Pagel, Mark. 2009. Human language as a culturally transmitted replicator. *Nature Reviews Genetics* 10: 405–15. [CrossRef] [PubMed]
Pagel, Mark, and Ruth Mace. 2004. The cultural wealth of nations. *Nature* 428: 275–78. [CrossRef] [PubMed]
Panikkar, Raimon. 2007. *Lo spirito della parola [The Spirit of the Word]*. Torino: Bollati Boringhieri.
Paradis, Michel. 1994. Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism. In *Implicit and Explicit Learning of Second Languages*. Edited by Nick Ellis. London: Academic Press, pp. 393–419.
Paradis, Michel. 2009. *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins.
Patel, Aniruddh. 2008. *Music, Language, and the Brain*. Oxford: Oxford University Press.
Pattee, Howard. 2008. Physical and Functional Conditions for Symbols, Codes, and languages. *Biosemiotics* 1: 147–68. [CrossRef]
Pattee, Howard. 2015. Cell phenomenology: The first phenomenon. *Progress in Biophysics and Molecular Biology* 119: 461–68. [CrossRef] [PubMed]

Pattee, Howard Hunt, and Kalevi Kull. 2009. A biosemiotic conversation: Between physics and semiotics. *Sign Systems Studies* 37: 311–31. [CrossRef]

Pattee, Howard Hunt, and Joanna Rączaszek-Leonardi. 2012. *Laws, Language and Life: Howard Pattee's Classic Papers on the Physics of Symbols with Contemporary Commentary*. New York: Springer.

Peirce, Charles Sanders. 1931–1935. *Collected Papers of Charles Sanders Peirce*. Edited by Charles Harstorne and Paul Weiss. Cambridge: Harvard University Press, vol. I–IV.

Pierce, John R. 1961. *An Introduction to Information Theory: Symbols, Signals and Noise*. New York: Dover Publications.

Pinker, Steven. 1994. *The Language Instinct. How the Mind Creates Language*. New York: Harper.

Provine, Robert. 2000. *Laughter: A Scientific Investigation*. New York: Viking.

Riesch, Rüdiger. 2016. Killer whales are speciating right in front of us. *Scientific American* 315: 54–56. [CrossRef] [PubMed]

Riesch, Rüdiger, Lance G. Barrett-Lennard, Graeme M. Ellis, John K. B. Ford, and Volker B. Deecke. 2012. Cultural traditions and the evolution of reproductive isolation: Ecological speciation in killer whales? Population divergence in killer whales. *Biological Journal of the Linnean Society* 106: 1–17. [CrossRef]

Ritt, Nikolaus. 2017. Linguistic Pragmatics from an Evolutionary Perspective. In *The Routledge Handbook of Pragmatics*. Edited by Anne Barron, Yueguo Gu and Gerard Steen. London: Routledge, pp. 490–502.

Rosenfield, Israel. 1992. *The Strange, Familiar, and Forgotten: An Anatomy of Consciousness*. New York: Vintage.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & Company.

Schaller, Susan. 1991. *A Man without Words*. Berkeley: University of California Press.

Scott-Phillips, Thom. 2015. *Speaking Our Minds. Why Human Communication is Different, and How Language Evolved to Make It Special*. New York: Palgrave Macmillan.

Senghas, Ann. 2005. Language emergence: Clues from a new Bedouin Sign Language. *Current Biology* 15: R463–R465. [CrossRef] [PubMed]

Senghas, Ann, Sotaro Kita, and Asli Ozyurek. 2004. Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305: 1779–82. [CrossRef] [PubMed]

Shannon, Claude Elwood, and Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press.

Singh, Jagjit. 1966. *Great Ideas in Information Theory, Language and Cybernetics*. New York: Dover Publications.

Singh, Simon. 1999. *The Code Book. The Secret History of Codes and Codebreaking*. New York: Anchor.

Spielman, Richard S., Ernest C. Migliazza, and James V. Neel. 1974. Regional Linguistic and Genetic Differences among Yanomama Indians. *Science* 184: 637–44. [CrossRef] [PubMed]

Tattersall, Ian. 2012. *Masters of the Planet: The Search for Our Human Origins*. New York: St. Martin's Griffin.

Thorpe, William. 1972. The Comparison of Vocal Communication in Animals and Man. In *Non-Verbal Communication*. Edited by Robert A. Hinde. Cambridge: Cambridge University Press, pp. 27–48.

Treccani. n.d. Available online: https://www.treccani.it/vocabolario/linguaggio/ (accessed on 7 April 2022).

von Humboldt, Wilhelm. 1836. *On Language: On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species*. Translated by Peret Heath. Cambridge: Cambridge University Press.

Wiener, Norbert. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*. New York: Wiley.

Wrangham, Richard W., and Dale Peterson. 1996. *Demonic Males: Apes and the Origins of Human Violence*. Boston: Mariner Book.

Wrangham, Richard W., Michael L. Wilson, and Martin N. Muller. 2006. Comparative rates of violence in chimpanzees and humans. *Primates* 47: 14–26. [CrossRef] [PubMed]

languages

Review

What Can Aphasia Tell Us about How the First-Acquired Language Is Instantiated in the Brain?

Mira Goral [1,2,3]

[1] Speech-Language-Hearing Sciences, Lehman College, The City University of New York, 250 Bedford Park Blvd., Bronx, NY 10468, USA; mira.goral@lehman.cuny.edu
[2] Speech-Language-Hearing Sciences, Graduate Center, The City University of New York, Bronx, NY 10016, USA
[3] MultiLing Center for Society Across the Lifespan, University of Oslo, 0313 Oslo, Norway

Abstract: Recent neurolinguistic theories converge on the hypothesis that the languages of multilingual people are processed as one system in the brain. One system for the multiple languages is also at the core of a translanguaging framework of multilingualism—a framework that focuses on each speaker's complete linguistic repertoire rather than on the separate languages they know. However, evidence from neuroimaging studies suggests at least some nonoverlapping activations of the first-acquired language (L1) and other (non-L1) languages of multilingual people, especially when the age of acquisition and/or levels of proficiency differ across the languages. Neurolinguistic studies of acquired language disorders have demonstrated that in multilingual people who experience language impairments due to brain lesion, L1 may be less impaired or better recovered than non-L1. This paper explores the evidence available to date from the study of acquired language impairment regarding this potential primacy of the first-acquired language. Findings suggest that L1 may be better preserved in many instances of language impairment, challenging the theory of a single system for multiple languages.

Keywords: first-acquired language; L1; non-L1; aphasia; primary progressive aphasia; dementia; neuroimaging; bilingual; multilingual; adults

Citation: Goral, Mira. 2022. What Can Aphasia Tell Us about How the First-Acquired Language Is Instantiated in the Brain? *Languages* 7: 283. https://doi.org/10.3390/languages7040283

Academic Editor: John W. Schwieter

Received: 14 May 2022
Accepted: 21 October 2022
Published: 4 November 2022

1. Introduction

Two main sources of neurolinguistic evidence have informed our understanding of language processing in the brain: lesion studies and neuroimaging studies. The complexity of the human language system, combined with great interindividual variability, has made the mapping of those neuronal networks responsible for language challenging. Despite these challenges, researchers have put forward models of language processing in the brain, implicating extensive neuronal networks in both cerebral hemispheres (e.g., Fridriksson et al. 2018; Price 2010), though such models are still debated. Further complicating matters is the even greater interindividual variabilities associated with multilingualism. Multilingual people, who comprise the majority of the world's population, are those who use more than one language (e.g., Grosjean 2022). Indeed, one of the questions that has preoccupied many researchers and that has remained unresolved is whether people who know and use multiple languages engage the same neuronal system for all their languages. Scholars have hypothesized that the system subserving language adapts to manage all languages as they are learned and mastered. Thus, the same neuronal networks are engaged for (any) language processing, rather than that separate systems subserve different languages of the same individual (Abutalebi and Green 2007, 2016).

This idea may challenge the notion that the first-acquired language, sometimes termed the mother tongue, has a special status in the brains of multilingual people, especially when other languages are learned later in life or do not reach the highest levels of proficiency. A special status of the first-acquired language has been hypothesized and examined in both sets of literature: lesion studies and neuroimaging studies. The following sections explore what the literature tells us about processing differences of the first-acquired language (L1) and other languages (non-L1) of multilingual people. I first briefly review evidence from neuroimaging studies with neurotypical multilingual people.

2. Functional Neuroimaging Studies of L1 vs. Non-L1

Since the early studies of functional neuroimaging, and with renewed interest in neuroplasticity, a large number of neurolinguistic studies set out to examine similarities and differences in brain activation of the first language (L1) and second language (L2) of bilingual people (e.g., Abutalebi et al. 2001; Cao et al. 2014; Golestani et al. 2006; Illes et al. 1999; Jeong et al. 2007; Liu et al. 2007, 2021; Nakada et al. 2001; Nelson et al. 2009; Oh et al. 2019; Perani et al. 2003; Saur et al. 2009; Tan et al. 2003; Vingerhoets et al. 2003, to name a few).

In such studies, brain activation is measured—using functional magnetic resonance imaging (fMRI) or positron emission tomography (PET)—while the participants are performing a language task in each of their languages, to compare activation patterns associated with each language. Activation patterns are also compared across different groups of bilingual participants while performing a language task in their non-L1 to examine the effects of the age of language appropriation and degrees of proficiency on activation patterns. The following two examples illustrate this approach.

In one study, while in the scanner, participants were presented with verbs in the present tense in their second language and were asked to first covertly conjugate the verbs in the past tense, and then to say the verbs in their past tense forms (Oh et al. 2019). The authors were interested in comparing regular and irregular verbs in English, the participants' L2. Group results for the regions activated during the overt conjugation of the two verb types were computed and contrasted to examine differences between the two conditions and across three participant groups: early bilingualism with high proficiency in both languages, late bilingualism with high proficiency in both languages, and late bilingualism with low proficiency in L2 (Oh et al. 2019).

In Wartenburger et al. (2003), for instance, bilingual participants completed a grammaticality judgement task in each of their two languages (German and Italian), and the performance of three bilingual groups (early bilingualism with high proficiency in both languages, late bilingualism with high proficiency in both languages, late bilingualism with low proficiency in L2) was compared across the two languages. Other studies have employed a variety of language processing and production tasks, such as verbal fluency (e.g., Perani et al. 2003), listening to sentences (e.g., Saur et al. 2009), silent narrative production (e.g., Tu et al. 2015), and word rhyming judgment (e.g., Cao et al. 2014).

This large literature, studying speakers of different language combinations (e.g., English–Mandarin, Cantonese–Mandarin, Catalan–Spanish, English–Spanish, German–Italian), evinced an inconclusive answer to the question of whether the two languages engage completely overlapping or somewhat differential neuronal networks. Individual studies revealed that processing in non-L1 activates the same language areas found to be active during L1 processing (e.g., Illes et al. 1999; Tan et al. 2003), but a number of studies have found also significant (albeit often small) differences in activation patterns emerging for each of the languages (e.g., Liu et al. 2007; Perani et al. 2003). Differences in the results obtained from individual studies may be attributed to variation in structural differences between the language pairs examined, variation in the language components engaged by the tasks used (e.g., semantics, syntax, phonology), and heterogeneity in the bilingual profiles of the participants (e.g., age of L2 learning, degree of L2 proficiency). Several recent systematic reviews and meta-analyses of neuroimaging studies with bilingual people attempt to extract findings that traverse the individual studies and their methodological variation. These review studies converge on the finding of wider brain activation patterns in the participants' non-L1 compared with their L1, although this finding is qualified by the levels of proficiency and the age of language appropriation of the participants studied and the linguistic components examined.

The most recent of these reviews is a meta-analysis study published by Cargnelutti et al. (2019) of 57 papers that reported fMRI (53) and PET (4) studies of L1 and L2 processing in bilingual participants. The results demonstrated that for participants who learned their second language after age 3 and before age 5 (the early bilingual group), performing language tasks in both L1 and L2 activated primarily left frontal regions, with no significant differences between the activation patterns of L1 and L2. Specifically, activation associated with L1 included the left precentral gyrus, left inferior frontal gyrus, left posterior medial frontal gyrus, left inferior temporal gyrus, and left middle temporal gyrus. Similarly, L2 activated the left precentral gyrus, left inferior frontal gyrus, left posterior medial frontal gyrus, and also left superior parietal lobule. For the late bilingual group (those who learned their L2 after age 6), the meta-analysis revealed that both languages activated the expected regions associated with language, namely, left frontal and left temporal regions (including the precentral gyrus, posterior medial frontal gyrus, and inferior frontal gyrus), as well as regions associated with executive functioning, such as the dorsolateral prefrontal cortex. Within-group comparisons revealed no activation clusters associated exclusively with L1. L2, however, was associated with greater activation than L1 in the left inferior frontal gyrus and left posterior medial frontal gyrus. In their review, Cargnelutti et al. aimed to examine the role of the age of L2 acquisition on L1 and non-L1 neuronal processing and selected the papers accordingly; they did not address which language tasks were performed by the participants included in the studies or the specific language pairs represented in the sample.

Interestingly, evidence for greater dissimilarities between the two languages in early rather than in later bilingual participants, when controlling for proficiency levels, was reported in at least one study, published after the 2019 review, although overall similar regions were activated for both L1 and L2 (including left inferior and middle frontal cortex, left superior temporal gyrus, and bilateral occipital cortices) (Ou et al. 2020).

The greater activation during L2 processing reported in Cargnelutti et al. (2019) was also found in a previous meta-analysis study (Liu and Cao 2016). One of the questions Liu and Cao (2016) set out to answer was, again, whether the age of L2 appropriation affects the degree of similarities in brain activations associated with L1 and L2. To answer this question, they examined 13 studies with early bilingual participants and 27 studies with late bilingual participants, all highly proficient in both languages. Their results revealed greater areas of activation associated with L2 than with L1 for both early and late bilingual participants, with greater differences between L1 and L2 processing for the late bilingual group. The regions associated with greater L2 than L1 activation included the left insula, left precentral gyrus, left middle frontal gyrus, and left inferior frontal gyrus in both the early and late groups, plus the left superior frontal gyrus and left medial frontal gyrus in the late group only. The additional regions included areas associated with both language processing and executive functioning (e.g., left inferior frontal gyrus) or primarily with executive functioning (e.g., left medial frontal gyrus). Liu and Cao attempted to match the language tasks the studies used for the comparisons between L1 and L2 activation. Moreover, they were interested in the role of orthographic similarities in L1 and L2 and found differences in L2 activation depending on the degree of orthographic transparency (e.g., greater bilateral temporal and precentral gyri activation for L2 that was more transparent than L1 vs. greater left frontal gyri activation for L2 that was less transparent than L1).

Language-specific activation has been consistently found in studies that used cortical stimulation as their method. Direct cortical stimulation involves applying direct current to an exposed cortical surface during neurosurgery. Cortical stimulations allow for a perhaps more refined localization of language processing in the brain than do neuroimaging studies, such as fMRI, although the language tasks are typically limited in scope (as they are performed immediately pre-op). Giussani et al. (2007) reviewed studies that used cortical stimulation to map language-related brain regions prior to neurosurgical operations in multilingual patients. They included in their review seven studies, reporting on 70 patients, mostly using naming tasks. All studies reported common but also language-specific sites; language-specific performance was associated with left posterior temporoparietal regions and left frontal regions.

The greater activation associated with L2 compared with L1 revealed in the review papers can point to the differential engagement of neuronal networks by each language, providing the answer "not completely" to the question of the overlapping neuronal involvement of L1 and non-L1. Indeed, the finding of greater activation in L2 is consistent with findings from skill learning. It has been shown that performing a new task activates larger regions compared with performing the same task after becoming skilled in it (e.g., Gobel et al. 2011; Patel et al. 2013). Thus, performing a task in an L2 may be associated with greater activation than performing a similar task more proficiently in L1.

Moreover, the greater activation found for L2 can be interpreted, as Abutalebi and Green argue (Abutalebi and Green 2007; Green and Abutalebi 2013), as associated with activation and inhibition mechanisms, rather than with differences in language processing per se. Abutalebi and Green identified a control network involving cortical and subcortical areas, likely responsible for the mechanisms that allow bilingual people to select one language or more during their communication with speakers who do or do not share all their languages. In papers published in 2007 and 2008 (Green and Abutalebi 2008), these authors reviewed neuroimaging studies of L1 and L2 processing and demonstrated the association between cognitive control processes and activation in regions including the left prefrontal cortex, the left caudate, and the inferior parietal lobules bilaterally. In a 2016 paper, the authors focused on control processes as fundamental to language use, and updated their model of the neural basis for the control mechanism to include the dorsal anterior cingulate cortex/pre-supplementary motor area, the left prefrontal cortex, the left caudate, the inferior parietal lobules bilaterally, the right prefrontal cortex, the thalamus and the putamen of the basal ganglia, and the cerebellum (Abutalebi and Green 2016). The regions that have evinced greater activation during L2 processing in the systematic reviews, such as the left inferior frontal gyrus, including the dorsolateral prefrontal cortex and left posterior medial frontal gyrus, and the left superior parietal lobule, are all part of the control network proposed by Abutalebi and Green. However note that there is some overlap in the regions that are included in the control network and in the language network, such as the left inferior frontal gyrus.

The conviction that language, any language, is processed in one system of neuronal networks whether the person speaks one language or multiple languages (e.g., Abutalebi and Green 2016; Marangolo et al. 2009) is consistent with the framework of translanguaging (Otheguy et al. 2015). A translanguaging theory of multilingualism reminds us that individuals who know multiple languages use components from one or more of their languages at any given moment according to the communication situation and their interlocutors, appropriately selecting those, typically without awareness or explicit decisions. These individual idiolects, the complete linguistic repertoire of each person, are likely represented as a single system in the individual's brain. These ideas are also consistent with evidence from psycholinguistic studies for nonselective lexical access in multilingual people. That is, experimental studies have repeatedly shown cross-language influences even when the language context and task at hand were designed to elicit processing in only one language. (The large body of literature on nonselective lexical activation is beyond the scope of this paper, but see, for example, Kroll et al. 2012, 2014).

Taken together, many researchers agree that the complete linguistic system of each person is instantiated in one, albeit complex, network system in the brain (e.g., Green and Abutalebi 2013; Perani and Abutalebi 2005) and that named languages are sociopolitical constructs rather than (neuro)linguistic ones (e.g., Vogel and Garcia 2017). However, neurolinguistic evidence for a potential special status of the first-acquired language persists (e.g., Cargnelutti et al. 2022). Similarly, psycholinguistic evidence for the need to inhibit a dominant (often the first-acquired) language while processing non-L1 (or a less-proficient language) (e.g., Green 1998; Misra et al. 2012) is consistent with the possibility of a unique status for that first-acquired language. The notion of L1 inhibition has been examined in a variety of studies in relation to nonselective lexical activation (e.g., Kroll et al. 2012), language switching processing cost (e.g., Costa and Santesteban 2004), and the extra activation during L2 processing (Abutalebi and Green 2007).

Nevertheless, within the framework of a single system subserving multiple languages, one might predict that an acquired lesion in the brain will affect all languages—indeed, the single language system—of a multilingual person. I turn next to evidence from studies of acquired language disorders in multilingual adults.

3. Brain Lesion Studies of L1 vs. Non-L1

The promise that the manifestation of acquired language disorders resulting from brain lesion in bilingual people can shed light on the underlying brain organization in bilingualism has intrigued the neurolinguistic community for quite some time. Comparable impairment of all the languages of a multilingual person is consistent with the assumption of a single neuronal network subserving all languages: even a single focal lesion to the system will affect all languages. Moreover, lesion in the networks hypothesized to be responsible for language control may manifest in impairment in language selection. I turn next to review the evidence from two types of acquired language impairments: stroke related and dementia related.

3.1. Stroke-Induced Aphasia

Aphasia is an acquired language disorder resulting from brain lesion, most commonly a stroke. The language impairments among people with aphasia vary in degree of severity and in scope. Typically, all language modalities are impaired (spoken, written, signed), and most people with aphasia experience word-finding difficulties that interfere with their ability to communicate. In severe aphasia, all linguistic systems appear impaired and communication is greatly affected, such that people with severe aphasia have compromised capability to understand spoken or written language and to contribute meaningfully to a communicative situation. On the other end of the severity continuum, people with mild aphasia retain much of their communicative abilities but experience difficulty retrieving words during language production (anomia) and may have difficulty comprehending and contributing meaningfully to a fast, multi-interlocutor conversation. The degree of severity and the profile of the impairment are associated with the size and the site of the stroke-induced lesion (e.g., McNeil and Pratt 2001; Papathanasiou and Coppens 2021; Turkeltaub 2019).

Early reports of stroke-induced aphasia in multilingual people highlighted cases with differential impairment across languages (Paradis 1983, 2004). In such reported cases, one language was markedly more impaired than another, or the person was able to communicate in only one of their previously used languages. The focus on differential impairment, possibly a result of a publication bias, led to a discussion of factors that can account for the patterns observed. Ribot (1881) suggested that the first-acquired language would have greater resilience compared with all other languages. In contrast, Pitres (1895) observed that the language most used at the time of the aphasia onset may be the more accessible to the patients. Nevertheless, early reviews seem to agree that the majority of multilingual people with aphasia show comparable impairment in their languages, or relative impairments that are consistent with their relative levels of proficiency prior to the aphasia onset (Albert and Obler 1978; Fabbro 2001; Paradis 2004).

A recent systematic review and meta-analysis that included 65 studies with 130 cases found a small but significant advantage to the first language of multilingual people with aphasia (Kuzmina et al. 2019). That is, L1 tended to be less impaired than L2, for both language comprehension and language production tasks. This was especially true for late bilingual participants, those who learned their L2 after age 7. (Note that to answer the question and compare the degree of impairments in L1 vs. L2, the authors excluded from the analyses six cases of simultaneous bilingualism, for whom there were two L1s rather than L1 and an L2.) A recent meta-analysis of language intervention effects in multilingual people with aphasia also revealed that treatment in L1 yielded stronger within-language benefits than treatment in non-L1, regardless of the linguistic modality targeted or tested (Lehtonen et al. 2022). (In their analyses, Lehtonen et al. included both later and early bilingual participants, including those who had two L1s.) These results are consistent with those of a previous review of treatment efficacy in bilingual people with aphasia (Faroqi-Shah et al. 2010) and with the better recovery of L1 reported in Kuzmina et al. (2019).

Despite these findings, multiple reports have also demonstrated the less expected pattern, that is, cases for whom the first language is not better preserved than other non-L1 languages. Cases illustrating this outcome include those reported in Filiputti et al. (2002), Goral et al. (2013), and Kiran and Roberts (2010). For example, one of the four participants described in Kiran and Roberts (2010), P1, demonstrated greater difficulty communicating in her L1, Spanish, than her L2, English, in all language modalities, after her stroke. Growing up in the U.S., she was exposed first to Spanish and then to English, has used both languages daily, and rated her proficiency higher in English than in Spanish. The higher proficiency in L2 prior to the stroke may have influenced the better recovery of that language after the stroke. In the case reported in Goral et al. (2013), the first-acquired language of the participant was the least recovered of her three languages. Lower abilities in L1 were observed in all language modalities and tasks. That first-acquired language was arguably the least used in the years following her aphasia onset. The participant's greatest recovery of English, the language of the environment and the one spoken in her immediate family, is consistent with Pitres' law. Therefore, we find an interaction between the laws of Ribot and Pitres: For many bilingual people, the first-acquired language also remains their most proficient and dominant language, and, in aphasia, often the most resilient. However, given the dynamic nature of language proficiency (e.g., Lerman et al. 2020; Navarro and Rossi 2022), some people develop higher proficiency in their non-L1, and when the first-acquired language is not used much for years around the aphasia onset, a more used non-L1 may be the more recovered one.

In the literature on another type of acquired language impairments, dementia, we find reports that corroborate Ribot's law even when language use would have predicted Pitres'. I turn now to the literature on language impairment and language preference in multilingual people with dementia. I address two sets of studies: those that examine language impairment in primary progressive aphasia and those that study dementia of the Alzheimer's type.

3.2. Dementia-Related Aphasia

3.2.1. Primary Progressive Aphasia Studies

Primary progressive aphasia (PPA) is a type of dementia that primarily affects language abilities before affecting other cognitive abilities. Although the etiology of PPA is not well understood, a variety of etiologies have been suggested, including genetic mutation (e.g., Premi et al. 2013). People with PPA vary in the manifestation of the acquired impairment. Language difficulties include anomia, slowed rate of language production, semantic deficits, and reading impairment, among others. Variability in these impairments is associated with three main variants: nonfluent, semantic, and logopenic (Gorno-Tempini et al. 2011). A growing number of studies have been published on individuals who are multilingual and manifest language difficulties due to PPA (e.g., Costa et al. 2019; Ellajosyula et al. 2020; Kambanaros and Grohmann 2012; Lind et al. 2018; Malcolm et al. 2019; Zanini et al. 2011). Several case reports converge on the finding of greater resilience of the first-acquired language in bilingual individuals with PPA.

For example, Lind et al. (2018) collected data in two time points from a bilingual English–Norwegian speaker with PPA. At the first testing time, 12 months after diagnosis, there was a clear advantage to the first-acquired language, with better word retrieval performance in L1 in naming tasks and in conversation. However, at the second testing time, 30 months after diagnosis, both languages appeared to be impaired, with marked word-finding difficulties in single-word retrieval tasks and in a conversation context in both.

Similarly, Zanini et al. (2011) reported greater impairments—in all linguistic components—in the L2 (Italian) of their Friulian–Italian bilingual participant who was highly proficient in both her languages. A similar pattern was also observed for an English–Hebrew bilingual woman who was highly proficient in both her languages prior to her diagnosis (Lerman et al. Forthcoming), for a Portuguese–French bilingual with PPA (Machado et al. 2010), and for two multilingual patients with PPA who demonstrated greater impairment in their non-L1 as their symptoms worsened (Mendez et al. 2004). In three cases of Spanish–English bilingual people with PPA, all of whom reported using English regularly at work and in other domains and two of whom endorsed better skills in and greater use of their English than their Spanish at the time before the diagnosis, better performance was observed in their Spanish, L1, than in English, their L2, on tasks including object and action naming and recognition and elicited language production (Goral Forthcoming).

Largely comparable impairments were documented in one case of a Chinese–English bilingual person (Filley et al. 2006) at a later stage of the PPA progression (the authors reported that the participant had noted language decline for 6 years prior to participation in their study). Similarly, comparable deterioration over a 1-year period was reported in Druks and Weekes (2013) for a Hungarian–English late bilingual person with PPA, despite greater impairment in L2 at the start of the study. The time after diagnosis may play a crucial role in the finding of comparable vs. differential impairments in L1 and non-L1 in PPA (Lind et al. 2018).

Two recent review papers of language impairment in bilingual people with PPA support the potential better preservation of L1. A review of 13 published cases with PPA (Malcolm et al. 2019) confirmed that in no case was L1 more impaired than non-L1. Of the 13 individuals, 8 were highly proficient in both (or all) their languages, 3 were L2 dominant, and 1 was dominant in L1 (for 1 individual, this information was not available). Another review paper reporting on 33 bilingual/multilingual participants with PPA found largely comparable impairments in the languages of the participants, although the authors highlight the great interindividual variability in the sample (Costa et al. 2019). The authors reported the dominant language for the participants, which was almost evenly divided between L1 (n = 15) and L2 (n = 14), with 3 individuals reporting equal dominance and 1 for whom no information was available. Interestingly, in over 50% of the sample, language impairment was noted in L2 first; from the data the authors present, it is difficult to say whether this finding was driven primarily by the individuals reporting L1 dominance. Moreover, greater impairment in naming in L2 was observed for the patients with the agrammatic variant of PPA (n = 14), not all of whom reported L1 dominance, whereas parallel naming impairments were evident for the other two variants. Overall impairment was largely comparable across the languages of all the participants. A recent group study of 16 bilingual people with PPA that examined the question of comparable versus differential levels of impairments reported greater resilience of L1 in both word comprehension and word retrieval (Ellajosyula et al. 2020). However, Ellajosyula et al. determined L1 to be the language of greater proficiency, as rated by the participants or their caregivers; this was the first-acquired language of 14 of the 16 participants.

Taken together, the studies of multilingual people with PPA suggest largely comparable impairment across the languages, as language abilities deteriorate with the progression of the dementia. In some cases, greater impairment may be experienced in non-L1, even for individuals who have achieved high proficiency levels and have used their non-L1 languages extensively throughout their lives.

3.2.2. Alzheimer's Disease Dementia Studies

Dementia of the Alzheimer's type is the most common type of dementia, affecting multiple cognitive abilities, including language. Early language difficulties associated with Alzheimer's disease (AD) include word-retrieval difficulties and impaired discourse production and comprehension (e.g., Mendez and Cummings 2003; Vonk et al. 2020). The anomia observed in people with AD manifests in empty speech (use of light words, such as "this" and "that one"), circumlocutions, and neologisms (nonwords) (e.g., Kavé and Goral 2016; Nicholas et al. 1985).

Evidence supporting the idea that L1 may be more resilient in bilingual individuals with dementia has been reported in the literature, but only in a few case reports (e.g., McMurtray et al. 2009). McMurtray et al. (2009) reported on two Japanese–English highly proficient bilingual individuals who had used both their languages regularly. Both individuals started preferring using their L1, Japanese, while gradually stopping the use of their L2, English, prior to the time of the evaluation when they were diagnosed with dementia. The study did not report their language performance in their two languages, but the authors proposed the idea that the language preference shift may be taken as an early marker for the onset of dementia. Brice et al. (2014) examined the performance of a Spanish–English bilingual person with AD at multiple testing times and demonstrated that his performance was, over time, better in his L1, Spanish, than in his highly proficient English, as measured by general language and cognition assessment tools (the Global Deterioration Scale and the Mini-Mental Status Examination). In four Japanese–Portuguese individuals with AD, language performance did not seem to differ in L1 and L2 (Meguro et al. 2003), but language specific characteristics affected the participants' reading performance in the two languages.

Three group studies, reporting on highly proficient bilingual people with dementia, lend minimal support to a greater resilience of a first-acquired language. Salvatierra et al. (2007) examined verbal fluency performance in a group of 11 Spanish–English bilingual people with mild to moderate AD. As a group, the participants retrieved more words in Spanish, their L1. Differences in performance between the AD group and a cognitively healthy control group were found for semantic (but not phonemic) fluency, and for both Spanish and English. In contrast, Manchon et al. (2014) found comparable impairments in the two languages of their 13 highly proficient bilingual individuals (L1 German, Spanish, or Italian; L2 French) with probable AD, who were tested on a variety of language tests. Veenstra et al. (2014) examined object naming in 26 highly proficient Frisian–Dutch bilingual participants with mild to moderate dementia associated with probable AD and found comparable levels of accuracy performance for both languages. They did, however, find a significant correlation between the age of acquisition of the words and naming accuracy, consistent with the idea that earlier-acquired words would be more resilient to dementia-related deterioration.

Gollan et al. (2010) compared Spanish–English bilingual participants (all native speakers of Spanish) with and without AD-related dementia on an object picture naming test. English-dominant participants with AD (n = 16) named fewer pictures than a matched group of participants without AD. The difference was more pronounced in English than in Spanish. The Spanish-dominant participants with AD (n = 13) in the study also named fewer pictures than their matched control group, but there was no interaction between group and language such that the participants with AD named fewer items in both their L1 and L2 compared with the participants without dementia. The authors interpreted their results to suggest that testing in the dominant language of bilingual people with dementia is more revealing than testing in their nondominant language.

However, the results of Gollan et al. (2010) may be specific to the population they enrolled, namely, native speakers of Spanish who have been immersed in English and may have become more proficient in their early-acquired L2, English, than in their first-acquired Spanish. Moreover, as the authors acknowledge, the test they used was designed in English, and the two versions, the English and Spanish ones, may not have been comparable. It is possible that the Spanish-dominant participants in the Gollan et al. (2010) study would have demonstrated better performance in their L1, Spanish, but because of the test characteristics, an equal degree of impairment was found for both their L1 and L2 compared with the participants without dementia. In a later study, Ivanova et al. (2014) observed that their 18 participants with probable AD, all childhood bilingual individuals with high proficiency in both their languages, also demonstrated greater decline in picture-naming performance in their nondominant language than in their dominant language, suggesting again that the degree of language use interacts with the age of language appropriation in the process of decline.

I hypothesize that the different patterns of results emerging for PPA vs. AD dementia may be related to the underlying impairment in the two types of dementia. In PPA, language is primarily impaired, at least in the early stages of neurodegeneration, while other cognitive abilities remain preserved to a great extent. If there is a differential linguistic impairment of L1 and non-L1 with acquired brain lesion, we may expect this difference to emerge in multilingual people with PPA. An exception here would be the presence of semantic degradation, which would affect all languages. Semantic degradation is evident in the semantic variant of PPA and in later stages of dementia of the AD type, and appears to affect all languages. This is consistent with the assumption that the underlying semantic network is shared for all languages of multilingual people, supported by a large body of literature (see, for example, Francis 2020; Kroll and Tokowicz 2005; Kroll et al. 2010).

In AD, declining cognitive abilities, such as executive functioning and cognitive control (Baudic et al. 2006; Bondi et al. 2009), may underlie the observed language impairment, at least at the early stages of the acquired language impairments (e.g., Kempler and Goral 2008; Rogers and Friedman 2008). Such cognitive impairment is likely to affect task performance, regardless of the language examined, and thus, we may expect no differences in the decline patterns observed for L1 and non-L1. Evidence from caregiver reports supports reversion to the first-acquired language in later stages of AD-related dementia (e.g., Mendez et al. 1999), presumably when the language system itself is impaired due to further deterioration in brain integrity.

Finally, consistent with the prediction of impaired language control following lesions affecting the control network is the finding that some multilingual individuals with an acquired brain lesion mix their languages unintentionally (e.g., Fabbro et al. 2000; Svennevig et al. 2019). In some instances, atypical language switching behavior is observed, such as switching to a language that the interlocutor does not understand. Such language selection errors have been interpreted as the result of faulty language control mechanisms (Fabbro et al. 2000; Fyndanis and Lehtonen 2021). In other cases, mixing words and phrases from two (or more) languages or switching midutterance from one language to another—often but not always to L1—can be attributed to attempts to resolve word-finding difficulties (Goral et al. 2019).

4. Concluding Remarks

In search of an answer to the question of whether the processing of the first-acquired language is unique as compared with later-learned languages, I have reviewed research evidence from neuroimaging studies of neurotypical multilingual adults and from studies of multilingual adults with acquired brain lesion. Evidence from cases of acquired language impairment is consistent with a more robust instantiation of the first-acquired language in the brain of multilingual people. This was evident in the better preservation and recovery of the first-acquired language in people who experience a stroke and consequently acquire aphasia. Similarly, evidence from primary progressive aphasia, a progressive acquired language impairment due to neurodegeneration, suggests that the first-acquired language deteriorates less or after non-L1 languages do. In dementia associated with Alzheimer's disease, in which cognitive impairment is prevalent, language impairment appears to progress in parallel in both languages of bilingual people, although in this case too, a reverting to the use of the first-acquired language in later stages of brain deterioration may be evident.

This evidence from lesion studies aligns, albeit only partially, with evidence from neuroimaging studies with neurotypical bilingual people. Whereas many researchers converge on the idea that the same neuronal system is engaged in the processing of all languages of multilingual people, neuroimaging studies have consistently found differential brain activation during the performance of a variety of language tasks in the first-acquired language vs. non-L1 languages. The activation differences between L1 and L2 have been found across studies that used different language tasks and included participants with varying degrees of language proficiency and varying ages of language acquisition. Differences between languages appear more consistently in studies that used more refined methods, such as cortical stimulation, than the broader activation captured in fMRI and PET studies, which may point to the shortcomings of the latter methods. In current neuroimaging studies, interpretation of the association between the signals measured and the language activities used can be affected by the regions analyzed, the tasks and subtractions used, among other variables (e.g., Fedorenko and Kanwisher 2009; Hayes and Huxtable 2012; Indefrey 2006).

The activation differences for L1 and non-L1 found in the neuroimaging studies have been interpreted primarily as indicating the engagement of control mechanisms necessary for the appropriate activation and inhibition of a language while using another, a skill that bilingual people develop as they adapt to using one or more of their languages in different communicative situations. Indeed, much of the additional activation observed for L2 over L1 centered in regions associated with the hypothesized control mechanisms. Almost all of the extra activation associated with non-L1 could be attributed to control mechanisms, supporting the argument that multiple languages share their underlying neuronal networks. We would expect greater prevalence of unintentional language switching and mixing following acquired brain lesion that affects the control network, a phenomenon that has been reported but is not highly prevalent. A control deficit explanation of all differential impairments is also less consistent with the systematically better preservation of L1, rather than any one of the languages, evident in the literature. The data reviewed here suggest that regardless of whether the first-acquired and later-learned languages may share overlapping or not completely overlapping neuronal networks, L1 may be instantiated in the brain in a manner that affords it greater resilience in the presence of brain damage.

It takes a shift in language dominance to change the expected pattern of better recovery in L1; language is dynamic, and patterns of language use clearly affect language processing in the brain (Green and Abutalebi 2013; Libben and Schwieter 2019). However, evidence suggests that even when non-L1 becomes highly proficient and predominantly used, L1 may be more resilient against brain-related deterioration. The special status of the first-acquired language evident in the lesion studies is consistent with findings from other disciplines, including, for example, the differential patterns found for emotional processing in the first-acquired language (e.g., Hadden et al. 2020; Liu et al. 2022; Sulpizio et al. 2019).

The greater resilience of a first-acquired language may be related to the notion of automaticity. It is possible that processing in a first-acquired language is automatic in a way that later-learned languages, even when achieving high levels of proficiency, never reach (e.g., McLeod and McLaughlin 1986; Segalowitz 2008). High proficiency in a language is associated with faster processing and more accurate performance than that measured for lower levels of proficiency in non-L1, but performance in non-L1, even at high proficiency, still retains a degree of variability that can indicate nonautomatic processing (Segalowitz 2010). The additional demands associated with the nonautomatic processing can be the source of differential activation patterns documented in neuroimaging studies of bilingualism and with differential effects of acquired brain damage. I attempted to bring together evidence from these two types of neurolinguistic studies to explore the potential special status of a first-acquired language. Whereas the two sets of data converge, to a degree, the review highlights a number of ways in which the discipline can move forward.

Avenues for future directions include investigations of the relationship between underlying brain impairment and the manifestation of the language deficits in each language, especially with respect to the patterns of language proficiency and use across the individual's lifespan and in the period prior to the acquired brain changes. This has been explored in the literature on stroke-related aphasia and, to a lesser extent, in dementia among multilingual people, but the heterogeneity inherent in this population makes draw-ing conclusions difficult. In this respect, dissociating the effects of specific cognitive and linguistic deficits and their influence on each language can help resolve inconsistent results for differing types of dementias and aphasias. The implications of such dissociations for models of bilingual processing and control mechanisms could then be interrogated. For example, the role of language inhibition in the manifestation of selective impairment in the languages of bilingual people with varying types of acquired brain lesions warrants further investigation. As well, the implication of shared neuronal networks for applied theories of translanguaging could strengthen the evidence supporting the shift toward translanguaging instruction and assessment in typical and atypical multilingual individuals. Furthermore, as our detection tools improve, a more fine-grained understanding of brain activation patterns as they relate to language processing should help refine our conclusions about multiple language processing in the brain. Triangulating results from a variety of measures and disciplines will help us understand the potential primacy of the first-acquired language and its implications.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: I thank my colleagues, including Gary Libben, for fruitful discussions. I thank the reviewers for their helpful input.

Conflicts of Interest: The authors declare no conflict of interest.

References

Abutalebi, Jubin, and David W. Green. 2007. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* 20: 242–75. [CrossRef]

Abutalebi, Jubin, and David W. Green. 2016. Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition* 19: 689–98. [CrossRef]

Abutalebi, Jubin, Stefano F. Cappa, and Daniela Perani. 2001. The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition* 4: 179–90. [CrossRef]

Albert, Martin L., and Loraine K. Obler. 1978. *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism*. New York: Academic Press.

Baudic, Sophie, Gianfranco D. Barba, Marie C. Thibaudet, Alain Smagghe, Philippe Remy, and Latchezar Traykov. 2006. Executive function deficits in early Alzheimer's disease and their relations with episodic memory. *Archives of Clinical Neuropsychology* 21: 15–21. [CrossRef]

Bondi, Mark W., David P. Salmon, and Alfred W. Kaszniak. 2009. The neuropsychology of dementia. In *Neuropsychological Assessment of Neuropsychiatric and Neuromedical Disorders*. Edited by Igor Grant and Kenneth M. Adams. Oxford: Oxford University Press, pp. 159–98.

Brice, Alejandro E., Sarah E. Wallace, and Roanne G. Brice. 2014. Alzheimer's dementia from a bilingual/bicultural perspective: A case study. *Communication Disorders Quarterly* 36: 55–64. [CrossRef]

Cao, Fan, Say Y. Kim, Yanni Liu, and Lin Liu. 2014. Similarities and differences in brain activation and functional connectivity in first and second language reading: Evidence from Chinese learners of English. *Neuropsychologia* 63: 275–84. [CrossRef]

Cargnelutti, Elisa, Barbara Tomasino, and Franco Fabbro. 2019. Language brain representation in bilinguals with different age of appropriation and proficiency of the second language: A meta-analysis of functional imaging studies. *Forntiers in Human Neuroscience* 13: 154. [CrossRef]

Cargnelutti, Elisa, Barbara Tomasino, and Franco Fabbro. 2022. Effects of linguistic distance on second language brain activations in bilinguals: An exploratory coordinate-based meta-analysis. *Frontiers in Human Neuroscience* 15: 744489. [CrossRef]

Costa, Albert, and Mikel Santesteban. 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. [CrossRef]

Costa, Ana S., Regina Jokel, Alberto Villarejo, Sara Llamas-Velasco, Kimiko Domoto-Reilley, Jennifer Wojtala, Kathrine Reetz, and Álvaro Machado. 2019. Bilingualism in primary progressive aphasia: A retrospective study on clinical and language characteristics. *Alzheimer Disease and Associated Disorders* 33: 47–53. [CrossRef]

Druks, Judit, and Brendan S. Weekes. 2013. Parallel deterioration to language processing in a bilingual speaker. *Cognitive Neuropsychology* 30: 578–96. [CrossRef] [PubMed]

Ellajosyula, Ratnavalli, Jwala Narayanan, and Karalyn Patterson. 2020. Striking loss of second language in bilingual patients with semantic dementia. *Journal of Neurology* 267: 551–60. [CrossRef] [PubMed]

Fabbro, Franco. 2001. The bilingual brain: Bilingual aphasia. *Brain and Language* 79: 201–10. [CrossRef] [PubMed]

Fabbro, Franco, Miran Skrap, and Salvatore Aglioti. 2000. Pathological switching between languages after frontal lesions in a bilingual patient. *Journal of Neurology Neurosurgery and Psychiatry* 68: 650–52. [CrossRef] [PubMed]
Faroqi-Shah, Yasmeen, Tobi Frymark, Robert Mullen, and Beverly Wang. 2010. Effect of treatment for bilingual individuals with aphasia: A systematic review of the evidence. *Journal of Neurolinguistics* 23: 319–41. [CrossRef]
Fedorenko, Evelina, and Nancy Kanwisher. 2009. Neuroimaging of language: Why hasn't a clearer picture emerged? *Language and Linguistics Compass* 3: 839–65. [CrossRef]
Filiputti, Dario, Alessandro Tavano, Lorenza Vorano, Guido De Luca, and Franco Fabbro. 2002. Nonparellel recovery of languages in a quadrilingual aphasic patient. *International Journal of Bilingualism* 6: 395–410. [CrossRef]
Filley, Christopher M., Gail Ramsberger, Lise Menn, Jiang Wu, Bessie Y. Reid, and Allan L. Reid. 2006. Primary progressive aphasia in a bilingual woman. *Neurocase* 12: 296–99. [CrossRef]
Francis, Wendy S. 2020. Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism* 24: 464–77. [CrossRef]
Fridriksson, Julius, Dirk-Bart den Ouden, Argye E. Hillis, Gregory Hickok, Chris Rorden, Alexandra Basilakos, Grigori Yourganov, and Leonardo Bonilha. 2018. Anatomy of aphasia revisited. *Brain* 141: 848–62. [CrossRef]
Fyndanis, Valantis, and Minna Lehtonen. 2021. Cognitive control and pathological language mixing in aphasia. In *Multilingualism Across the Lifespan*. Edited by Robert Blackwood and Unn Røyneland. London: Routledge, pp. 209–30.
Giussani, Carlo, Franck-Emmanuel Roux, Vincent Lubrano, Sergio M. Gaini, and Lorenzo Bello. 2007. Review of language organisation in bilingual patients: What can we learn from direct brain mapping? *Acta Neurochirurgica* 149: 1109–16. [CrossRef] [PubMed]
Gobel, Eric W., Todd B. Parrish, and Paul J. Reber. 2011. Neural correlates of skill acquisition: Decreased cortical activity during a serial interception sequence learning task. *Neuroimage* 58: 1150–57. [CrossRef] [PubMed]
Golestani, Narly, François Xavier Alario, Sébastien Meriaux, Denis Le Bihan, Stanislas Dehaene, and Christophe Pallier. 2006. Syntax production in bilinguals. *Neuropsychologia* 44: 1029–40. [CrossRef] [PubMed]
Gollan, Tamar H., David P. Salmon, Rosa I. Montoya, and Eileen Da Pena. 2010. Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. *Neuropsychologia* 48: 1356–66. [CrossRef] [PubMed]
Goral, Mira. Forthcoming. Language shift in bilingual people with primary progressive aphasia.
Goral, Mira, Maryam Naghibolhosseini, and Peggy S. Conner. 2013. Asymmetric inhibitory treatment effects in multilingual aphasia. *Cognitive Neuropsychology* 30: 564–77. [CrossRef]
Goral, Mira, Monica Norvik, and Bård U. Jensen. 2019. Variation in code-switching in multilingual aphasia. *Journal of Clinical Linguistics and Phonetics* 33: 915–29. [CrossRef]
Gorno-Tempini, Maria L., Argye E. Hillis, Sandra Weintraub, Andew Kertesz, Mario Mendez, Stefano E. E. Cappa, Jennifer Ogar, Jonathan D. Rohrer, Stuart Black, Bradley F. Boeve, and et al. 2011. Classification of primary progressive aphasia and its variants. *Neurology* 76: 1006–14. [CrossRef]
Green, David W. 1998. Mental control of the bilingual lexico- semantic system. *Bilingualism: Language and Cognition* 1: 67–81. [CrossRef]
Green, David W., and Jubin Abutalebi. 2008. Understanding the link between bilingual aphasia and language control. *Journal of Neurolinguistics* 21: 558–76. [CrossRef]
Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef]
Grosjean, François. 2022. *The Mysteries of Bilingualism: Unresolved Issues*. Hoboken: Wiley.
Hadden, Lowri M., Ben Alderson-Day, Mike Jackson, Charles Fernyhough, and Richard P. Bentall. 2020. The auditory-verbal hallucinations of Welsh-English bilingual people. *Psychology and Psychotherapy: Theory, Research and Practice* 93: 122–33. [CrossRef] [PubMed]
Hayes, Dave J., and Adrianne G. Huxtable. 2012. Interpreting deactivations in neuroimaging. *Frontiers in Psychology* 3: 27. [CrossRef] [PubMed]
Illes, Judy, Wendy S. Francis, John E. Desmond, John D. Gabrieli, Gary H. Glover, Russell Poldrack, Anthony D. Christine J. Lee, and Anthony D. Wagner. 1999. Convergent cortical representation of semantic processing in bilinguals. *Brain and Language* 70: 347–63. [CrossRef]
Indefrey, Peter. 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. [CrossRef]
Ivanova, Iva, David P. Salmon, and Tamar H. Gollan. 2014. Which Language Declines More? Longitudinal versus Cross-Sectional Decline of Picture Naming in Bilinguals with Alzheimer's Disease. *Journal of the International Neuropsychological Society* 20: 534–46. [CrossRef] [PubMed]
Jeong, Hyeonjeong, Motoaki Sugiura, Yuko Sassa, Tomoki Haji, Nobuo Usui, Masatoi Taira, Kaoru Horie, Shigeru Sato, and Ryuta Kawashima. 2007. Effect of syntactic similarity on cortical activation during second language processing: A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping* 28: 194–204. [CrossRef] [PubMed]
Kambanaros, Maria, and Kleanthes K. Grohmann. 2012. BATting multilingual primary progressive aphasia for Greek, English, and Czech. *Journal of Neurolinguistics* 25: 520–37. [CrossRef]
Kavé, Gitit, and Mira Goral. 2016. Word retrieval on structured tests and in connected speech in healthy older adults and in individuals with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology* 38: 958–66. [CrossRef]

Kempler, Daniel, and Mira Goral. 2008. Language and Dementia: Neuropsychological aspects. *Annual Review of Applied Linguistics* 28: 1–18. [CrossRef]
Kiran, Swathi, and Patricia M. Roberts. 2010. Semantic feature analysis treatment in Spanish–English and French–English bilingual aphasia. *Aphasiology* 24: 231–61. [CrossRef]
Kroll, Judith F., and Natasha Tokowicz. 2005. Models of bilingual representation and processing: Looking back and to the future. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Annette M. B. de Groot. Oxford: Oxford University Press, pp. 531–54.
Kroll, Judith F., Janet G. Van Hell, Natasha Tokowicz, and David W. Green. 2010. The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition* 13: 373–81. [CrossRef]
Kroll, Judith F., Paula E. Dussias, Cari A. Bogulski, and Jorge R. V. Kroff. 2012. Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In *The Psychology of Learning and Motivation*. Edited by Brian H. Ross. Cambridge: Elsevier Academic Press, pp. 229–62. [CrossRef]
Kroll, Judith F., Susan C. Bobb, and Noriko Hoshino. 2014. Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science* 23: 159–63. [CrossRef]
Kuzmina, Ekaterina, Mira Goral, Monica Norvik, and Brendan Weekes. 2019. What influences language impairment in bilingual aphasia? A meta-analytic review. *Frontiers in Psychology* 10: 445. [CrossRef] [PubMed]
Lehtonen, Minna, Monica Norvik, Jan Antfolk, and Mira Goral. 2022. Cross-language generalization in treatment of multilingual aphasia: A meta-analysis. Poster presented at the Learning and Neuroplasticity Meeting, Äkäslompolo, Finland, April 2–5.
Lerman, Aviva, Dorit Mais, Yael Nissani, and Taryn Malcolm. Forthcoming. Preserving lexical retrieval skills across languages in a bilingual person with logopenic primary progressive aphasia. *Aphasiology*. [CrossRef]
Lerman, Aviva, Mira Goral, and Loraine K. Obler. 2020. The complex relationship between pre-stroke and post-stroke language abilities in multilingual aphasia. *Aphasiology* 34: 1319–40. [CrossRef]
Libben, Gary, and John W. Schwieter. 2019. Lexical organization and reorganization in the multilingual mind. In *The Handbook of the Neuroscience of Multilingualism*. Edited by John W. Schwieter. Malden and Oxford: Wiley-Blackwell, pp. 297–313.
Lind, Marianne, Hanne G. Simonsen, Ingeborg S. B. Ribu, Bente A. Svendsen, Jan Svennevig, and Kees de Bot. 2018. Lexical access in a bilingual speaker with dementia: Changes over time. *Clinical Linguistics & Phonetics* 32: 353–77. [CrossRef]
Liu, Hengshuang, and Fan Cao. 2016. L1 and L2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. *Brain and Language* 159: 60–73. [CrossRef]
Liu, Linyan, John W. Schwieter, Fenqi Wang, and Huanhuan Liu. 2022. First and second languages differentially affect rationality when making decisions: An ERP study. *Biological Psychology* 169: 108265. [CrossRef] [PubMed]
Liu, Xiaojin, Liu Tu, Xiaoxi Cheng, Junjing Wang, Meng Li, Zhi Lu, and Ruiwang Huang. 2021. Effect of AoA-L2 on L1 and L2 networks in early and late bilinguals. *International Journal of Bilingualism* 25: 1616–34. [CrossRef]
Liu, Ying, Susan Dunlap, Julie Fiez, and Charles Perfetti. 2007. Evidence for neural accommodation to a writing system following learning. *Human Brain Mapping* 28: 1223–34. [CrossRef]
Machado, Alvaro, Margarida Rodrigues, Sonia Simoes, Isabel Santana, and Joao Soares-Fernandes. 2010. The Portuguese who could no longer speak French: Primary progressive aphasia in a bilingual man. *Journal of Neuropsychiatry and Clinical Neurosciences* 22: E31–E32. [CrossRef] [PubMed]
Malcolm, Taryn, Aviva Lerman, Marta Korytkowska, Jet M. J. Vonk, and Loraine K. Obler. 2019. Primary Progressive Aphasias in Bilinguals and Multilinguals. In *The Handbook of the Neuroscience of Multilingualism*. Edited by J. Schwieter. Hoboken: Wiley.
Manchon, Mélanie, Karin Buetler, Françoise Colombo, Lucas Spierer, Frédéric Assal, and Jean-marie M. Annoni. 2014. Impairment of both languages in late bilinguals with dementia of Alzheimer type. *Bilingualism: Language and Cognition* 18: 90–100. [CrossRef]
Marangolo, Paula, Christina Rizzi, Patrice Peran, Fabrizio Piras, and Umberto Sabatini. 2009. Parallel recovery in a bilingual aphasic: A neurolinguistic and fMRI study. *Neuropsychology* 23: 405–09. [CrossRef] [PubMed]
McLeod, Beverly, and Barry McLaughlin. 1986. Restructuring or automaticity? Reading in a second language. *Language Learning* 36: 109–23. [CrossRef]
McMurtray, Aaron, Erin Saito, and Beau Nakamoto. 2009. Language preference and development of dementia among bilingual individuals. *Hawaii Medical Journal* 68: 223–26. [PubMed]
McNeil, Malcolm R., and Shiela R. Pratt. 2001. Defining aphasia: Some theoretical and clinical implications of operating from a formal definition. *Aphasiology* 15: 901–11. [CrossRef]
Meguro, Kenichi, Mirna L. H. Senaha, Paulo Caramelli, Junichi Ishizaki, Rosa Y. S. Chubacci, Mitsue Meguro, Hideo Ambo, Ricardo Nitrini, and Atsushi Yamadori. 2003. Language deterioration in four Japanese–Portuguese bilingual patients with Alzheimer's disease: A trans-cultural study of Japanese elderly immigrants in Brazil. *Psychogeriatrics* 3: 63–68. [CrossRef]
Mendez, Mario F., and Jeffry L. Cummings. 2003. *Dementia: A Clinical Approach*, 3rd ed. Amsterdam and New York: Butterworth-Heinemann Medical.
Mendez, Mario F., Kent M. Perryman, Marcel O. Pontón, and Jeffry L. Cummings. 1999. Bilingualism and dementia. *The Journal of Neuropsychiatry and Clinical Neurosciences* 11: 411–12. [CrossRef]
Mendez, Mario F., Samira Saghafi, and David G. Clark. 2004. Semantic dementia in multilingual patients. *The Journal of Neuropsychiatry and Clinical Neurosciences* 16: 381–81. [CrossRef]

Misra, Maya, Taomei Guo, Susan C. Bobb, and Judith F. Kroll. 2012. When bilinguals choose a single word to speak: Electrophysiological evidence for global inhibition in bilingual word production. *Journal of Memory and Language* 67: 224–37. [CrossRef]

Nakada, Tsutomu, Yukihiko Fujii, and Ingrid L. Kwee. 2001. Brain strategies for reading in the second language are determined by the first language. *Neuroscience Research* 40: 351–58. [CrossRef]

Navarro, Ester, and Eleonora Rossi. 2022. Using latent variable analysis to capture variability in bilingual language experience. *PsyArXiv*. [CrossRef]

Nelson, Jessica R., Ying Liu, Julie Fiez, and Charles A. Perfetti. 2009. Assimilation and accommodation patterns in ventral occipitotemporal cortex in learning a second writing system. *Human Brain Mapping* 30: 810–20. [CrossRef] [PubMed]

Nicholas, Marjorie, Loraine K. Obler, Martin L. Albert, and Nancy Helm-Estabrooks. 1985. Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research* 28: 405–10. [CrossRef] [PubMed]

Oh, Tomasina M., Steven Graham, Philina Ng, Ing B. Yeh, Bernard P. L. Chan, and Andrew M. Edwards. 2019. Age and Proficiency in the bilingual brain revisited: Activation patterns across different L2-learner types. *Frontiers in Communication* 4: 39. [CrossRef]

Otheguy, Ricardo, Ofelia García, and Wallis Reid. 2015. Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6: 281–307. [CrossRef]

Ou, Jian, Wenlong Li, Yang Yang, Nizhuan Wang, and Min Xu. 2020. Earlier second language acquisition is associated with greater neural pattern dissimilarity between the first and second languages. *Brain and Language* 203: 104740. [CrossRef]

Papathanasiou, Ilias, and Patrick Coppens. 2021. *Aphasia and Related Neurogenic Communication Disorders*, 3rd ed. Burlington: Jones & Bartlett Learning.

Paradis, Michel. 1983. *Readings in Aphasia in Bilinguals and Polyglots*. Montreal: Didier.

Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Patel, Ronak, R. Nathan Spreng, and Gary R. Turner. 2013. Functional brain changes following cognitive and motor skills training: A quantitative meta-analysis. *Neurorehabilitation and Neural Repair* 27: 187–99. [CrossRef]

Perani, Daniela, and Jubin Abutalebi. 2005. The neural basis of first and second language processing. *Current Opinion in Neurobiology* 15: 202–6. [CrossRef]

Perani, Daniela, Jubin Abutalebi, Eraldo Paulesu, Simona Brambati, Paula Scifo, Stefano F. Cappa, and Ferruccio Fazio. 2003. The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping* 19: 170–82. [CrossRef]

Pitres, Albert. 1895. Etude sur l'aphasie chez les polyglottes. *Revue de Medecine* 15: 873–99, [Translated by M. Paradis. *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Marcel Didier, pp. 26–48].

Premi, Enrico, Stefano Gazzina, Marco Bozzali, Silvana Archetti, Antonella Alberici, Mara Cercignani, Angelo Bianchetti, Roberto Gasparotti, Marinella Turla, Carlo Caltagirone, and et al. 2013. Cognitive reserve in *Granulin*-related frontotemporal dementia: From preclinical to clinical stages. *PLoS ONE* 8: e74762. [CrossRef] [PubMed]

Price, Cathy J. 2010. The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of The New York Academy of Sciences* 1191: 62–88. [CrossRef] [PubMed]

Ribot, Théodule. 1881. *Les Maladies de la Mémoire*, 2nd ed. Paris: Baillère.

Rogers, Sean L., and Rhonda B. Friedman. 2008. The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia* 46: 12–21. [CrossRef] [PubMed]

Salvatierra, Judy, Monica Rosselli, Amarillis Acevedo, and Ranjan Duara. 2007. Verbal fluency in bilingual Spanish/English Alzheimer's disease patients. *American Journal of Alzheimer's Disease & Other Dementias* 22: 190–201.

Saur, Dorothee, Annette Baumgaertner, Anja Moehring, Christian Büchel, Matthias Bonnesen, Michael Rose, Mariachristina Musso, and Jürgen M. Meisel. 2009. Word order processing in the bilingual brain. *Neuropsychologia* 47: 158–68. [CrossRef]

Segalowitz, Norman. 2008. Automaticity and second languages. In *The Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Hoboken: Wiley, pp. 382–408.

Segalowitz, Norman. 2010. *Cognitive Bases of Second Language Fluency*. London: Routledge.

Sulpizio, Simone, Michelle Toti, Nicola Del Maschio, Albert Costa, Davide Fedeli, Remo Job, and Jubin Abutalebi. 2019. Are you really cursing? Neural processing of taboo words in native and foreign language. *Brain and Language* 194: 84–92. [CrossRef]

Svennevig, Jan, Pernille Hansen, Hanne G. Simonsen, and Anne M. D. Landmark. 2019. Code-switching in multilinguals with dementia: Patterns across speech contexts. *Clinical Linguistic and Phonetics* 33: 1009–30. [CrossRef]

Tan, Li H., John A. Spinks, Ching M. Feng, Wai T. Siok, Charles A. Perfetti, Jinhu Xiong, Peter T. Fox, and Jia-Hong Gao. 2003. Neural systems of second language reading are shaped by native language. *Human Brain Mapping* 18: 158–66. [CrossRef]

Tu, Liu, Junjing Wang, Jubin Abutalebi, Bo Jiang, Ximin Pan, Meng Li, Wei Gao, Yuchen Yang, Bishan Liang, Zhi Lum, and et al. 2015. Language exposure induced neuroplasticity in the bilingual brain: A follow-up fMRI study. *Cortex* 64: 8–19. [CrossRef]

Turkeltaub, Peter E. 2019. A taxonomy of brain–behavior relationships after stroke. *Journal of Speech, Language, and Hearing Research* 62: 3907–22. [CrossRef]

Veenstra, Wencke S., Mark Huisman, and Nick Miller. 2014. Age of acquisition and naming performance in Frisian-Dutch bilingual speakers with dementia. *Dementia & Neuropsychologia* 8: 249–59. [CrossRef]

Vingerhoets, Guy, John Van Borsel, Cathelijne Tesink, Maurits van den Noort, Karel Deblaere, Ruth Seurinck, Pieter Vandemaele, and Eric Achten. 2003. Multilingualism: An fMRI study. *NeuroImage* 20: 2181–96. [CrossRef] [PubMed]

Vogel, Sara, and Ofelia Garcia. 2017. *Translaguaging. Oxford Research Encyclopedia of Education*. Oxford: Oxford University Press.

Vonk, Jet M. J., Vincent Bouteloup, Jean F. Mangin, Bruno Dubois, Frédéric Blanc, Audrey Gabelle, Mathieu Ceccaldi, Cédric Annweiler, Pierre Krolak-Salmon, Catherine Belin, and et al. 2020. Semantic loss marks early Alzheimer's disease-related neurodegeneration in older adults without dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 12: e12066.
Wartenburger, Isabell, Hauke R. Heekeren, Jubin Abutalebi, Stefano F. Cappa, Arno Villringer, and Daniela Perani. 2003. Early setting of grammatical processing in the bilingual brain. *Neuron* 37: 159–70. [CrossRef]
Zanini, Sergio, Valentina Angeli, and Alessandro Tavano. 2011. Primary progressive aphasia in a bilingual speaker: A single-case study. *Clinical Linguistics & Phonetics* 25: 553–64.

languages

MDPI

Article

Emotion Word Processing in Immersed Spanish-English/ English-Spanish Bilinguals: An ERP Study

Anna B. Cieślicka * and Brenda L. Guerrero

Department of Psychology and Communication, Texas A&M International University, Laredo, TX 78041, USA
* Correspondence: anna.cieslicka@tamiu.edu

Abstract: We conducted a lexical decision task to measure Spanish-English/English-Spanish bilinguals' behavioral (RT) and electrophysiological (EPN, Early Posterior Negativity and LPC, Late Positive Complex) responses to English emotion words and their Spanish translation equivalents. Bilingual participants varied in age of acquisition (AoA of Spanish/English: early, late), language status (L1 Spanish, L1 English) and language dominance (English-dominant, Spanish-dominant, balanced) but were all highly immersed bicultural individuals, uniformly more proficient in English than Spanish. Behavioral data showed faster and more accurate responses to English than Spanish targets; however, the emotion effect was only present for Spanish, with positive Spanish words recognized significantly faster than those that were negative or neutral. In the electrophysiological data, the emotion response was affected by language of the target stimulus, with English targets eliciting larger EPN amplitudes than Spanish targets. The reverse effect was found on the LPC component, where Spanish targets elicited a higher positivity than English targets. Dominance did not turn out to be a significant predictor of bilingual performance. Results point to the relevance of proficiency in modulating bilingual lexical processing and carry implications for experimental design when examining immersed bilinguals residing in codeswitching environments.

Keywords: emotion words; bilingual; early posterior negativity; late positive complex; proficiency; dominance

Citation: Cieślicka, Anna B., and Brenda L. Guerrero. 2023. Emotion Word Processing in Immersed Spanish-English/English-Spanish Bilinguals: An ERP Study. *Languages* 8: 42. https://doi.org/10.3390/languages8010042

Academic Editor: John W. Schwieter

Received: 5 September 2022
Revised: 19 January 2023
Accepted: 20 January 2023
Published: 31 January 2023

1. Introduction

Emotion words can be categorized into emotion-label and emotion-laden, where the former name a specific emotional state (e.g., angry, overjoyed), while the latter do not directly refer to an emotion but elicit it (e.g., *kitty, war*; Altarriba and Basnight-Brown 2010). Such words differ along the dimension of valence (positive, negative, or neutral) and arousal, or the amount of physiological response (high or low) they evoke (Lang et al. 1997). Research on emotion processing has consistently demonstrated the so called "emotion effect", i.e., the differential processing of emotionally-relevant content relative to non-emotional, neutral material (see Citron 2012).

The emotion effect may manifest as a faster response to emotionally valenced words in a lexical decision task (e.g., Estes and Adelman 2008; Kousta et al. 2009; Kuchinke et al. 2005; Larsen et al. 2006; Schacht and Sommer 2009b), enhanced priming for emotion relative to neutral words (Altarriba 2006; Altarriba and Canary 2004), faster lexical access of emotion words in reading (Kissler and Herbert 2013), better recall (e.g., Altarriba and Bauer 2004; Anooshian and Hertel 1994; Ayçiçeği-Dinn and Caldwell-Harris 2009; Rubin and Friendly 1986), stronger attentional blink effect in response to emotion vs. neutral words (Colbeck and Bowers 2012), slower naming latencies in the Stroop task (Eilola et al. 2007; Sutton et al. 2007), increased galvanic skin response (GSR) to emotion words in psychophysiological studies (e.g., Harris et al. 2003), or a larger amplitude of an event-related potential (ERP) response in electrophysiological studies (e.g., Hofmann et al. 2009; Holt et al. 2005; Kissler et al. 2009; Zhang et al. 2014).

1.1. Electrophysiological Correlates of Emotion Word Processing

Two ERPs have emerged as major indices of emotion effects: the early posterior negativity (EPN) and the late positive complex (LPC). The EPN component, recorded mainly at occipito-temporal sites, is viewed as an index of early (i.e., automatic) lexical access as it starts immediately after the onset of the lexicality effect in the word recognition task. Peaking between 200 and 300 ms, it reflects an enhanced attention to emotional content at early processing stages (Herbert et al. 2006; Junghöfer et al. 2001; Kissler et al. 2007, 2009; Optiz and Degner 2012; Schacht and Sommer 2009a, 2009b; Schupp et al. 2003). The EPN emotion effect has been consistently found in studies using varied methodologies, such as silent reading (Kissler and Herbert 2013; Kissler et al. 2007, 2009), grammatical decision (e.g., Kissler et al. 2007, 2009) or lexical decision (Citron et al. 2011; Palazova et al. 2011; Schacht and Sommer 2009b; Scott et al. 2009), and it seems independent of the stimulus presentation rate or the nature of the task (Herbert et al. 2008; Kissler et al. 2006, 2009; but see Rellecke et al. 2011). EPNs have been reported for emotion words of different grammatical categories, such as nouns (Kissler et al. 2007), verbs (Schacht and Sommer 2009b), and adjectives (Herbert et al. 2006, 2008).

The LPC component, which begins at approximately 400–500 ms post-stimulus and lasts for a few hundred milliseconds, is primarily recorded at centro-parietal electrodes and reflects higher-order cognitive stages of more elaborate semantic processing (Citron 2012; Cuthbert et al. 2000; Fischler and Bradley 2006; Kissler et al. 2009; Palazova et al. 2013; Schupp et al. 2003). The LPC is sensitive to the dimension of valence, in that its amplitude increases for positively (or negatively) valenced words relative to neutral words (Hofmann et al. 2009; Kanske and Kotz 2007); however, increased LPCs have also been reported for high arousal neutral over emotionally valenced words (e.g., Citron et al. 2011; Recio et al. 2014; see also Yao et al. 2016). Unlike the EPN, the LPC is affected by task requirements. It manifests in tasks requiring explicit processing of the emotional content or deep semantic processing, such as an overt valence categorization task (e.g., Delaney-Busch et al. 2016) but not in shallow tasks, such as orthographic judgment to spelling patterns (Fischler and Bradley 2006), same/different font judgment (Schacht and Sommer 2009b), or a semantic categorization task (e.g., Delaney-Busch et al. 2016).

While the electrophysiological literature is generally consistent when it comes to emotion effects present in both early and late ERP components, the findings differ regarding differential processing of emotion over neutral words. An enhanced EPN has been found in response to positively valenced versus neutral or versus both negative and neutral words (Chen et al. 2015; Palazova et al. 2011; Recio et al. 2014), or in response to positive high-arousal and negative low-arousal words (Citron et al. 2013) and to emotionally arousing pleasant words over neutral words (e.g., Schacht and Sommer 2009b). Other research has shown an increase in the EPN elicited by both positive and negative versus neutral verbal stimuli (Herbert et al. 2008; Kissler and Herbert 2013; Kissler et al. 2007, 2009; Optiz and Degner 2012; Palazova et al. 2011; Schacht and Sommer 2009b). In later time windows, positive words have been associated with a larger LPC in comparison to negative or neutral words (Herbert et al. 2006, 2008; Kissler and Herbert 2013; Kissler et al. 2009; Palazova et al. 2011; Recio et al. 2014; Schapkin et al. 2000; Zhang et al. 2014); however, other studies have yielded results showing a larger LPC in response to negative relative to neutral words (e.g., Bayer et al. 2010; Hofmann et al. 2009), to negative words as compared to both positive and neutral words (Bernat et al. 2001; Delaney-Busch et al. 2016; Kanske and Kotz 2007), or an increased LPC amplitude in response to both negative and positive vs. neutral words (e.g., Conrad et al. 2011). These inconsistencies may be attributed to a number of factors that have been shown to modulate emotion processing, such as word frequency (e.g., Kissler et al. 2007; Kuchinke et al. 2007; Scott et al. 2009), concreteness (e.g., Hinojosa et al. 2014; Imbir et al. 2016; Kanske and Kotz 2007; Palazova et al. 2013; Yao et al. 2016), grammatical class (Palazova et al. 2011; Schacht and Sommer 2009b), arousal level of the emotionally valenced word (e.g., Citron et al. 2011, 2013; Delaney-Busch et al. 2016; Hofmann et al. 2009; Recio et al. 2014), task demands (e.g., Fischler and Bradley 2006; Hinojosa et al. 2010; Kissler et al. 2009; Schacht and Sommer 2009b), the origin (automatic vs. reflective) of the word's emotional content (Imbir et al. 2016), and individual differences (Citron 2012; Gibbons 2009; Mueller and Kuchinke 2016).

1.2. Emotion Word Processing in Bilinguals

The question of emotion word processing is even more complex with bilinguals who express and perceive emotions in more than one language. Early research into second language (L2) emotion processing (e.g., Bond and Lai 1986) has suggested the possibility that L2 emotion word processing is characterized by more distance than L1, primarily because of the strong coupling between cognition and emotion and the fact that emotional connotations of words are established during the person's cognitive growth (see Harris 2015). Hence, L1 emotion words are intrinsically linked to a person's emotional responses, unlike L2 words that have been acquired later in life. This diminished L2 emotionality has been referred to as *disembodied cognition* (Pavlenko 2012) or *reduced emotional resonance in L2* (Toivo and Scheepers 2019). Indeed, a number of studies into bilingual emotion word processing have found attenuated effects for bilinguals' L2 as compared to L1 emotion words (e.g., Anooshian and Hertel 1994; Caldwell-Harris 2014; Colbeck and Bowers 2012; Gonzales-Reigosa 1976; Harris et al. 2003; Harris 2004; Sheikh and Titone 2016).

For example, Iacozza et al. (2017) recorded pupillary responses of Spanish-English bilinguals engaged in reading emotionally-charged English or Spanish sentences. After each sentence, participants were instructed to rate its emotional impact. While pupillary responses showed a significantly larger effect for L1 (Spanish) compared to English sentences, explicit ratings of emotionality were comparable across both languages. Iacozza et al. (2017) suggest that the sympathetic nervous system might show an attenuated emotional response in L2 depending on the language context. While an automatic, implicit measure of emotional reactivity, such as pupil dilation, is likely to show differential effects in the native and foreign language, a more explicit measure, such as subjective rating of the material's emotional impact, might reveal no such effects. Attenuated response of the sympathetic nervous system to L2 emotion stimuli was also reported by Jankowiak and Korpal (2017), who presented late proficient Polish (L1)–English (L2) bilinguals with emotionally-laden spoken and written L1/L2 narratives. The GSR results showed a reduced response to L2 compared to L1, and the effect was further constrained by the modality of the presentation with visual stimuli eliciting a more pronounced skin conductance level than the auditory stimuli.

However, other research has shown comparable effects for bilinguals' L1 and L2 emotion word processing (e.g., Conrad et al. 2011; Eilola et al. 2007; Eilola and Havelka 2011; Ferré et al. 2010; Harris et al. 2006; Kim 1993; Optiz and Degner 2012; Ponari et al. 2015; Sutton et al. 2007), suggesting that the decreased sensitivity of bilinguals to L2 emotional content might be more nuanced than initially assumed and affected by such bilingual participant characteristics as age of acquisition (AoA), level of proficiency, amount of exposure to L2, and language dominance (e.g., Ayçiçeği and Harris 2004; Eilola et al. 2007; Harris 2004; Harris et al. 2003, 2006; Kazanas and Altarriba 2016; Sutton et al. 2007). For example, using a modified Stroop paradigm, Eilola et al. (2007) found equal emotional response for both L1 and L2 in late Finnish-English bilinguals who were highly proficient in their L2 (see also Sutton et al. 2007).

In addition to proficiency, language dominance has also been shown to affect bilingual emotion processing. In Harris et al.'s (2003) GSR study, native speakers of Turkish who learned English after 12 years of age were presented with L1 and L2 emotionally valenced words, including taboo words and childhood reprimands. Results showed a significant difference for reprimands in L1 vs. L2, with reprimands in Turkish eliciting a significantly stronger GSR than in English. Contrary to the expectation that taboo words in one's native language would elicit stronger responses than similar L2 taboo words learned later in life, reactivity to taboo words in both languages was highly comparable, suggesting that L1 is not necessarily a more emotional language in cases where L2 becomes more dominant (see also Ayçiçeği and Harris 2004).

Indeed, Caldwell-Harris (2014, 2015) suggests that all of these bilingual participant factors are interrelated and converge in modulating emotional processing. For example, high proficiency is typically causally linked with early acquisition and correlated with frequency of use or amount of exposure to language. Increased exposure and frequency of use are, in turn, relevant for dominance, which is also crucially dependent on the immersive learning as when a bilingual resides in the L2-speaking country. Based on her review of studies into the emotionality differences between multilinguals' languages, Caldwell-Harris (2015) emphasizes the need to account for those modulating factors, suggesting that emotional processing differences are the most pronounced when the person's L1 is dominant and L2 is less proficient and learned later.

One question raised in the ERP bilingual emotion literature is whether L2 emotion words can evoke the early EPN response compatible to that elicited by L1 or whether the emotional content of L2 words only becomes available at later stages of processing. Because the EPN and LPC components reflect, respectively, early (automatic) versus late (higher-order) lexical processes, they can provide an insight into whether L1 and L2 emotion word processing differ. If both components are comparably affected by the emotional valence of words, regardless of whether the words are presented in L1 or L2, this would be indicative of early access of emotional content in both L1 and L2. However, if EPN responses are more pronounced for L1, this would support the idea of an attenuated emotional response in L2. In turn, the LPC, reflecting more elaborate semantic processing strategies, might be observed for L2 even when no EPN response was recorded. This is because of the possibility that L2 words have been by then re-translated into L1 and it is the L1 word's emotional content that causes an increased amplitude in LPC.

The existing ERP emotion processing studies have shown mixed results so far. For example, in Chen et al.'s (2015) lexical decision experiment, Chinese-English bilinguals who were late learners of L2 English showed an enhanced EPN effect only in response to L1 positive words during the time windows of 250–300 ms and 300–350 ms. In turn, valence effects for L2 did not emerge until 400–500 ms post-stimulus, and their topography differed from the EPN component, suggesting no EPN effect for L2 words. Similarly, no effects were found for L2 in later time windows. The LPC effects were shown only for L1 in the time windows of 500–550 ms and 550–600 ms, such that larger amplitudes were recorded for neutral than for positive words. The only marginally significant emotion effect for L2 was found in the time window between 400–500 ms, with neutral words eliciting a marginally larger negativity than positive words.

Conversely, other studies showed the presence of emotion effects for both L1 and L2 on both early and late ERP components (Conrad et al. 2011; Kissler and Bromberek-Dyzman 2021; Optiz and Degner 2012). In addition, the timing of the early EPN component has been shown to differ across L1 and L2 in that the processing of emotion words in L2 may be delayed relative to L1. For example, Conrad et al. (2011) conducted an ERP study with Spanish-German and German-Spanish bilinguals matched on L2 proficiency. Results revealed that, regardless of the language status, emotionally valenced words evoked a larger amplitude of an EPN and LPC as compared to neutral words for both bilingual groups. These findings suggest that emotion word processing in L1 and L2 does not differ qualitatively, although quantitatively the EPN response was delayed by 50–100 ms for L2 relative to L1. Conrad et al. (2011) interpret this time shift as indicative of a general delay in L2 visual recognition processes rather than a delayed L2 emotion recognition per se.

Further support for qualitatively comparable L1 and L2 emotion effects was demonstrated by Optiz and Degner (2012), who asked German-French and French-German bilinguals to perform a go/no-go lexical monitoring task. Participants were presented with valenced and neutral L1/L2 words and asked to determine whether a pseudoword was orthographically similar to real words in the respective target language. Results showed an amplified EPN for both positive and negative compared to neutral words, regardless of the language status. As in Conrad et al.'s (2011) experiment, the timing of the EPN differed across L1 and L2, in that emotional processing of L2 words was delayed relative to L1. Optiz and Degner (2012) attribute this delay to costs of interference resolution in highly proficient L2 users. Since L1 and L2 lexicons in such proficient speakers are highly integrated, access to a word's emotional content results in automatic activation of both L2 and L1 lexical representations, hence incurring extra processing costs. However, in a recent study, Kissler and Bromberek-Dyzman (2021) failed to find any timing differences in the onset of EPN or LPC for L1 vs. L2 emotion responses in German-English bilinguals.

Previous studies into L2 emotion word processing have focused on comparing bilinguals against monolinguals (e.g., Kim 1993), late AoA L2 learners with different L2s (e.g., Conrad et al. 2011; Optiz and Degner 2012), unabalanced late bilinguals' performance in their L1 and L2 (Kissler and Bromberek-Dyzman 2021), or low proficient bilinguals performing in their L2 (e.g., Chen et al. 2015). Participants in these studies were consistently dominant in their L1. More recently, Vélez-Uribe and Rosselli (2021) examined emotion processing in Spanish-English bilinguals varying along the dimension of dominance and proficiency. While the balanced group comprised individuals with comparable levels of proficiency in their L1 and L2, the unbalanced bilinguals were more proficient in English than Spanish. Participants were asked to perform an emotion rating task in both Spanish and English. Results revealed a significant language effect on both the EPN and LPC components. The EPN data showed a larger amplitude for words in Spanish than English and the main effect of valence, i.e., an enhanced EPN response to positive vs. neutral and neutral vs. negative words. The LPC data showed overall larger amplitudes for words in English than in Spanish. In addition, significant differences emerged between balanced vs. unbalanced groups. Whereas balanced bilinguals showed comparable emotion effects for both English and Spanish, the unbalanced group manifested differences in the LPC amplitudes for Spanish words, such that positive targets recorded an enhanced positivity relative to negative and neutral targets. Vélez-Uribe and Rosselli (2021) explain these differences by suggesting that emotional content in the more proficient language might be processed identically by balanced and unbalanced bilinguals but that the processing patterns might diverge for the less proficient language.

1.3. The Present Study

The present study aims to further explore the time course of bilingual emotion word processing by focusing on bilingual participants who are not only highly proficient in their L2 but for whom L2 often becomes their dominant language. We use a lexical decision task to measure proficient Spanish-English and English-Spanish bilinguals' reaction time (RT) and electrophysiological (EPN and LPC) responses to English emotion-label and emotion-laden words and their Spanish translation equivalents. Our bilinguals offer a unique opportunity to assess the contributions of the various factors modulating L1 and L2 ERP emotion effects. Specifically, our participants reside in a highly immersive environment, a US-Mexican border town, where both languages are spoken interchangeably. While the majority of them learned Spanish as their L1 or were exposed to both Spanish and English simultaneously, their early educational experience in US-based schools and subsequent English-only academic environments led to many of these bilinguals becoming dominant in English. In addition, they fall on a continuum of AoA, in that some of them learned English in early childhood, while others learned it after they had acquired Spanish. We ask the following questions: (1) Do L1 and L2 emotion word processing differ qualitatively and/or quantitatively in highly proficient immersed bilinguals who are routinely exposed to both languages and reside in a bilingual community? (2) How do language dominance and AoA modulate L1/L2 emotion word processing?

Given the inconsistent research with bilinguals (Conrad et al. 2011; Kissler and Bromberek-Dyzman 2021; Optiz and Degner 2012), our question regarding quantitative/qualitative differences between L1 vs. L2 emotion processing is purely exploratory. Of note, Spanish-English/English-Spanish bilinguals examined here are typically more proficient in English than in Spanish, regardless of their L1. It is therefore likely that ERP responses might be more pronounced for emotion words in the more proficient language

(English) than for emotion words in Spanish. In addition, L1 and L2 emotion processing might be affected by language dominance. For example, for Spanish-English bilinguals dominant in Spanish, EPN and LPC responses to emotionally valenced words might be more enhanced when presented in their dominant language (Spanish) than in the weaker language (English). If, however, English has become the dominant language for a Spanish-English bilingual, the reverse might be expected, with English (L2) emotion words evoking a larger EPN and LPC response than Spanish (L1).

As for the role of AoA, early bilingualism might be indicative of greater proficiency on account of length of exposure (Caldwell-Harris 2014). We therefore expected to see a more pronounced emotion effect for English emotion words in early than in late English L1 speakers who are dominant in English. Such bilinguals not only learned English early in their life and thus benefit from an increased length of exposure relative to late learners, but they have maintained greater dominance and proficiency in English over Spanish. Early Spanish learners might show an enhanced emotion effect for Spanish words; however, if they became dominant and more proficient in English, this effect might be diminished. On the other hand, late L2 learners might show an attenuated emotion effect similar to Chen et al.'s (2015) results.

Overall, while the research questions asked here are largely exploratory, the novelty of our study lies in the fact that most of the existing literature has examined bilinguals dominant in their L1 who have learned their L2 in more formal settings, whereas the population we investigate is unique. Specifically, it consists of bilinguals whose L2 was acquired in the immersive environment and became their more proficient language. These bilinguals are not only highly proficient in English, but they reside in the bilingual and bicultural community characterized by dense codeswitching practices. Exploring L1 and L2 emotion word processing in such bilinguals might help to shed new light on the interplay of the various participant characteristics in modulating the emotion effect.

2. Materials and Methods

2.1. Participant

The participants were 27 bilinguals (9 male, 18 female, *M age* = 21.25, *SD* = 5.65) recruited from the student population of a South Texas university. Informed consent was obtained from all subjects involved in the study. Data from one participant were discarded due to the excessive amount of muscular and ocular artifacts (40%) in the EEG recordings. Participants were all right-handed (Oldfield 1971), with normal or corrected-to-normal vision. Proficiency was established based on the adapted version of the Language History Questionnaire (LHQ; Li et al. 2006), while language dominance was assessed with the Bilingual Dominance Scale (BDS) (Dunn and Tree 2009). A total of 16 participants reported Spanish as their L1, 4 grew up as simultaneous bilinguals, and 6 learned English as L1. Out of the 16 L1 Spanish participants, only 2 were born in Mexico, while the remaining were second and third generation immigrants who were born, raised, and educated in the US. All L1 English and simultaneous bilinguals were born and raised in the US. Regardless of their native language, the majority of the bilinguals became dominant in English. Specifically, based on the BDS, 16 bilinguals were categorized as English-dominant, 5 as balanced, and 5 as Spanish-dominant. Proficiency self-ratings revealed that, overall, bilingual participants rated English significantly higher than Spanish in terms of speaking, reading, understanding, and writing. In total, 15 bilinguals reported losing fluency in Spanish and all but 3 had over 7 years of schooling in English. Self-ratings for English were therefore consistently significantly higher than for Spanish in all bilingual groups (see Table 1 for summary of the participant characteristics).

Table 1. Participants' language background information. Proficiency rating was measured with a 7-point Likert scale where 1 = very poor, 7 = native like. Asterisks show significant differences in proficiency ratings for English and Spanish.

L1 STATUS	**L1 English (N = 6)**		**L1 Spanish (N = 16)**		**Both (N = 4)**			
Mean Age	22.8		20.6		21.3			
Dominance	Eng-dom (N = 6) Span-dom (N = 0)		Eng-dom (N = 6) Span-dom (N = 5) Balanced (N = 5)		Eng-dom (N = 4) Span-dom (N = 0)			
AoA English	Early (N = 6) Late (N = 0)		Early (N = 6) Late (N = 10)		Early (N = 2) Late (N = 2)			
AoA Spanish	Early (N = 2) Late (N = 4)		Early (N = 16) Late (N = 0)		Early (N = 4) Late (N = 0)			
Global	L1 English		L1 Spanish		Both		Mean All	
Proficiency	Eng	Span	Eng	Span	Eng	Span	Eng	Span
	7.0 ***	3.4	6.5 *	5.7	5.8 **	4.3	6.4 **	4.5
Speaking	7.0	3.0	6.5	5.6	6.0	4.3	6.5	4.3
Reading	7.0	3.5	6.6	5.6	6.0	4.0	6.5	4.4
Understanding	7.0	4.7	6.5	6.0	5.7	5.0	6.4	5.2
Writing	7.0	2.5	6.5	5.4	5.3	3.7	6.3	3.9

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

With regard to AoA, bilinguals were categorized as early and late. The term *early bilingual* is used in the literature to describe an individual who has been exposed to L2 from a very early age simultaneous with their cognitive and linguistic growth, for example as when a child grows up in a bilingual home with parents speaking two languages interchangeably. In turn, a *late bilingual* is a person whose exposure to L2 started after the foundations of their L1 have already been established and who starts learning L2 either through classroom instruction or immigration. The early category included participants who acquired language before the age of 5, while the late category included two subgroups: participants who learned Spanish/English between the ages of 6–9 and those who learned their L2 between 10–16 (see Heredia and Cieślicka 2014). Since only two participants reported learning their L2 between the ages of 10–16, the two late subcategories were collapsed for further analyses. Overall, for AoA of English, fourteen bilinguals reported learning English before the age of 5 and twelve learned English between the ages of 6–9. For Spanish, 22 were early and 4 were late bilinguals.

2.2. Stimuli

The stimuli were 240 English emotion-label and emotion-laden words selected from the Affective Norms for English Words (ANEW) database (Bradley and Lang 1999) and their Spanish translations obtained from the Spanish adaptation of ANEW (Redondo et al. 2007). Spanish translations that were cognates or cross-language homographs were excluded. Within each language, emotion words varied along the dimension of valence (positive, negative, and neutral), with 80 words for each condition. Ratings of valence in English differed significantly between positive (M = 7.27; CI [7.11, 7.44]), negative (M = 2.84; CI [2.61, 3.07]), and neutral (M = 5.56; CI [5.31, 5.8]) words, F (2, 237) = 404, $p_s < 0.001$. Ratings of arousal differed between positive (M = 5.53; CI [5.36, 5.70]) and neutral (M = 4.19; CI [3.99, 4.38]; $p_{Tukey} < 0.001$) and between negative (M = 5.70; CI [5.45, 5.95]) and neutral words ($p_{Tukey} < 0.001$), with no difference between positive and negative words ($p_{Tukey} = 0.51$). Similarly, for Spanish, stimuli ratings of valence differed significantly between positive (M = 7.22; CI [7.06, 7.38]), negative (M = 2.23; CI [2.10, 2.36]), and neutral (M = 5.12; CI [5.01, 5.24]) targets, F (2, 237) = 1277, $p_s < 0.001$. Arousal ratings were significantly different between Spanish positive (*M* = 6.01; CI [5.85, 6.16]) and neutral (*M* = 4.57; CI [4.47, 4.68]; $p_{\text{Tukey}} < 0.001$), as well as between negative (*M* = 6.20; CI [6.04, 6.35]) and neutral targets ($p_{\text{Tukey}} < 0.001$), with no difference between the positive and negative ($p_{\text{Tukey}} = 0.15$). Spanish and English arousal and valence ratings did not differ significantly across positive, negative, and neutral categories (all $p_s > 0.05$).

Stimuli were also matched according to word length, grammatical category, frequency, and concreteness (all $p_s > 0.05$; see Table 2 for summary of stimuli characteristics). Word frequencies were selected from the SUBTLEX-ESP database (Cuetos et al. 2011) for Spanish and SUBTLEX-US (Brysbaert et al. 2012) for English. Concreteness ratings were derived from Brysbaert et al. (2014) for English and from Hinojosa et al. (2016) for Spanish words. Concreteness was controlled for, such that for each language, half the words were concrete and half were abstract. Thus, the final list included 240 English words (40/40 positive concrete/abstract, 40/40 negative concrete/abstract, 40/40 neutral concrete/abstract). The Spanish translations followed the same procedure.

Table 2. Means for valence, arousal, concreteness, word length (number of letters) and word frequency for English and Spanish word stimuli. Square brackets represent SE.

	English			**Spanish**		
	Negative	**Neutral**	**Positive**	**Negative**	**Neutral**	**Positive**
Valence [1]	2.84 [0.12]	5.56 [0.13]	7.27 [0.08]	2.23 [0.07]	5.12 [0.06]	7.24 [0.08]
Arousal [2]	5.70 [0.13]	4.19 [0.1]	5.53 [0.09]	6.20 [0.79]	4.57 [0.05]	6.01 [0.08]
Frequency [3]	61.6 [8.3]	62.0 [9.2]	62.0 [7.5]	63.1 [5.1]	64.8 [6.2]	62.4 [12.7]
Concreteness [4]	4.59 [1.05]	5.41 [1.05]	4.94 [1.05]	4.67 [0.94]	5.47 [0.87]	4.80 [1.10]
Length [5]	5.42 [0.24]	5.33 [0.27]	5.79 [0.27]	6.49 [0.28]	5.91 [0.30]	6.59 [0.29]

[1] Based on ANEW and REDONDO valence norming: 1-very negative, 9-very positive. [2] Based on ANEW and REDONDO arousal norming: 1-not at all arousing, 9-highly arousing. [3] Based on SUBTLEX-US and SUBTLEX-ESP ($SUBTL_{WF}$)-frequency per 1 million of occurrences. [4] Based on concreteness ratings from Hinojosa et al. (2016) for Spanish: 1-very abstract, 9-very concrete; English concreteness values based on Brysbaert et al. (2014)—the scale was recoded to be comparable with Spanish: 1-abstract, 5-concrete. [5] Range = 2–5 syllables. Range = 4–11 letters.

An additional set of 160 nonwords (80 English and 80 Spanish) was created using Wuggy (http://crr.ugent.be/programs-data/ (accessed on 1 March 2020)), an experimental software that creates nonwords by changing letters from the provided set of language-specific items. The resulting nonwords were pronounceable in English/Spanish, orthographically legal, and matched with the experimental stimuli in terms of length.

Lists were counterbalanced using a Latin square design, and participants were randomly assigned to each list. Two experimental lists were needed to counterbalance the design and to ensure that a participant did not see the same emotion word in both English and Spanish. Each list included 160 nonwords and 240 emotion words, 120 of which were English and 120 were Spanish (see Supplementary Materials for a complete list of stimuli).

2.3. Procedure

Stimuli presentation was controlled by E-Prime 2.0 (Schneider et al. 2002), which automatically randomized stimuli for each participant. Participants were seated approximately 60–70 cm from a 21-inch computer screen with both index fingers resting on the Chronos response box. They were instructed to read the letter strings appearing on the computer screen and to respond, as fast and as accurately as possible, by pressing a "YES" button on the response box if the string of letters was a Spanish or English word or "NO" if the string of letters was a nonword. Response sides were counterbalanced across participants. On each trial, a fixation cross was displayed for 800 ms, followed by the stimulus. Stimuli were presented centrally in black letters (font: Arial, size: 20) against a white background and remained on the screen until participants responded. After each trial, a "BLINK NOW" message in a white screen in black capital letters appeared for 1000 ms, allowing participants to blink and relax their muscles.

2.4. EEG Recordings

EEG was recorded from 64 scalp sites using a Biosemi Active Two headcap (10/20) layout and referenced to electrode Cz. The common mode sense (CMS) active and the driven right leg (DRL) passive electrodes were used as ground electrodes. To minimize artifacts related to eye movements, bipolar horizontal and vertical electrooculography (EOG) activity was recorded with additional electrodes attached under and next to the eyes. Electrode impedances were kept below 5 kΩ. The EEG signals were recorded continuously at a sampling rate of 2048 Hz. Preprocessing steps were performed using *MATLAB* (MATLAB R2022a, The Mathworks, Inc., Natick, Massachusetts, United States), EEGLAB (v. 2021.1; Delorme and Makeig 2004), and the ERPLAB Toolbox (v.9.00; Lopez-Calderon and Luck 2014). The data were first visually inspected for abnormalities, with sections of data showing excessive muscular artefacts being manually rejected. Abnormal channel activity was detected with the help of the trimOutlier plugin (Lee & Myakoshi SCCN, INC, UCSD) and by plotting the channels in EEGLAB. No more than six channels were rejected in each dataset (M = 6; min = 0, max = 6)

Each dataset was next filtered offline with a 0.1 Hz high-pass (slope 12 dB/octave) and 30 Hz low-pass (slope 24 dB/octave) IIR Butterworth filter. Subsequently, to correct for vertical and horizontal EOG artefacts, the Independent Component Analysis (ICA; Makeig et al. 1995) was run on the EEG data. The mean number of rejected ICs per participant was 2.32 (SD = 0.97; min = 1, max = 4). EEG continuous signal was next segmented into epochs of 900 ms, starting 200 ms prior to stimulus onset. The pre-stimulus period of 200 ms was used for baseline correction. Ocular and muscular artifacts were corrected using Artifact Detection function (peak-to-peak moving window; threshold: +/−100 µV; window size: 200 ms; window step: 100 ms) in ERPLAB and further subjected to visual inspection. Epochs containing ocular/muscular artifacts or amplitudes exceeding ±100 µV were rejected (see Table 3 for summary of the percentage of accepted epochs per condition).

Table 3. Percentage of accepted epochs per condition in the EPN and LPC analyses.

Condition	Percentage of Accepted Epochs
English Positive	82.6%
English Negative	78.7%
English Neutral	77.2%
Spanish Positive	78.3%
Spanish Negative	80.6%
Spanish Neutral	80.8%

2.5. Statistical Analyses

2.5.1. Behavioral Data Analysis

RTs exceeding 3.0 standard deviations above the mean were excluded and analyzed as errors (2%). All analyses were performed on correct responses. Trimmed RT and accuracy data were analyzed with a linear mixed-effects model (LMM) using the buildmer package (v. 1.9, Voeten 2020; see also Matuschek et al. 2017) in R (v. 4.1.0., R Core Team 2021) and javomi (v. 2.0, The Jamovi Project 2021). The variables of interest were language of the target stimulus (English/Spanish) and valence (positive, negative, neural). Among bilingual participant factors, we originally planned to assess the effect of dominance. However, as is typical of our student population, while many bilinguals are heritage speakers who were exposed to Spanish since birth on account of growing up in a Spanish-speaking household, the majority became dominant in English by virtue of residing in the US and attending US schools (see Section 2.1). Because of the uneven number of participants in each dominance group, we entered participants' dominance score as a continuous variable.

Overall, the following fixed effects were included in each model: (1) language (English, Spanish); (2) valence (positive, negative, neutral); (3) dominance score; and (4) their interactions. The fixed effects were coded using deviation coding. As determined by buildmer, maximal models with a full random-effect structure were first computed. These included random intercept by subject and item and random slope by-subject and by-item (Barr et al. 2013; see also Matuschek et al. 2017). Maximally converged models were then run in jamovi and gamlj-General Analyses for Linear Models in jamovi (Version 2.4.7). The final structure, summary and variance components for each of the models are available in the Supplementary Materials.

2.5.2. Electrophysiological Data Analysis

The ERP data were segmented into three time windows defined *a priori* based on previous research: 200–300 ms and 300–400 ms for the EPN; and 500–700 ms for the LPC, (see, e.g., Kissler and Bromberek-Dyzman 2021; Scott et al. 2009). The two different time windows for the early component were chosen to address possible latency shifts in emotion effects recorded for L1 vs. L2, given the L2 delay effect reported in previous emotion word processing studies (e.g., Conrad et al. 2011; Optiz and Degner 2012). Based on the previous literature (e.g., Chen et al. 2015; Conrad et al. 2011; Kissler et al. 2007), the following electrodes were selected for this early component analysis: F7/8, PO3/PO4, P1/P2, P3/P4, P6, P7/P8, CP1, CP5, FC5/FC6, T7/8, O1, Oz. For the LPC effect, which is most salient at the centro-parietal sites, the following electrodes were chosen: CP1/CP2, Pz, P1/P2, P3/P4, P5/P6, PO3/PO4, CPz, Oz, O1/O2. Mean EPN and LPC amplitudes were analyzed with repeated measures (RM) ANOVAs, conforming to a 2 (language of the target stimulus: English, Spanish) by 3 (valence: positive, negative, neutral) within-subjects independent variables, with dominance score as a covariate. P-values were adjusted using Greenhouse–Geisser correction for violations of sphericity, and the Bonferroni correction was applied for multiple testing in all post hoc comparisons.

To fully address our research question regarding the effect of AoA on ERP emotion effects, we would need to compare the amplitudes elicited by Spanish/English emotion words between early/late learners of Spanish/English. However, as reported earlier (see Section 2.1), the majority of our bilinguals (22) were early Spanish learners, with only 4 participants in the late AoA Spanish group. Because of the unequal group size for the AoA of Spanish/English, this variable was not entered into the overall analysis.

3. Results

3.1. RT Data

We found a fixed effect of language ($\beta = -83.28$, $SE = 14.7$, $df = 30.09$, $t = -5.67$, $p < 0.001$), with slower responses to Spanish ($M = 963$ ms, 95% CI [878, 1048]) than English targets ($M = 777$ ms, 95% CI [707, 847]), and of valence ($\beta = 41.52$, $SE = 9.22$, $df = 439.72$, $t = 4.5$, $p < 0.001$), with faster responses to positive ($M = 810$ ms, 95% CI [736, 884]) than to negative ($M = 911$ ms, 95% CI [838, 985]) and neutral words ($M = 889$ ms, 95% CI [815, 963]). In addition, the analysis revealed a language x valence interaction ($\beta = -28.5$, $SE = 9.22$, $df = 439.71$, $t = -3.09$, $p < 0.01$). This interaction showed that positive Spanish targets were responded to faster ($M = 870$ ms, 95% CI [782, 958]) than negative ($M = 1033$ ms, 95% CI [945, 1121]) and neutral Spanish targets ($M = 986$ ms, 95% CI [898, 1074]). Regardless of valence, Spanish targets took longer to respond than English (negative Spanish: $M = 1033$ ms, 95% CI [945, 1121] vs. negative English: $M = 790$ ms, 95% CI [716, 864]; neutral Spanish: $M = 986$ ms, 95% CI [898, 1074] vs. neutral English: $M = 791$ ms, 95% CI [717, 865]; positive Spanish: $M = 870$ ms, 95% CI [782, 958] vs. positive English: $M = 749$ ms, 95% CI [675, 823]). Dominance failed to yield significant effects (see Figure 1 for summary of the RT data).

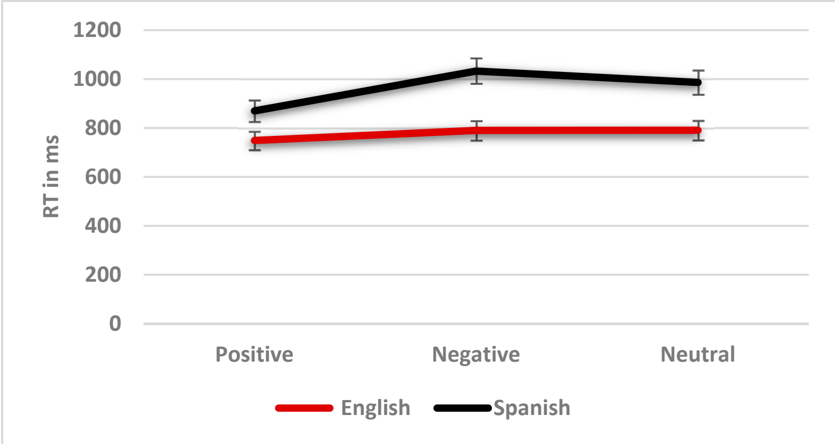

Figure 1. Mean RTs in milliseconds for English and Spanish positive, negative, and neutral words recorded in the LDT task. Error bars depict 95% confidence interval.

3.2. Accuracy Data

The accuracy analysis showed a fixed effect of language, $b = 0.57$, $SE = 0.24$, $z = 2.38$, $p < 0.001$, whereby English targets ($M = 98\%$, 95% CI [97, 98]) were responded to with greater accuracy than Spanish targets ($M = 89\%$, 95% CI [88, 90]). The analysis also yielded a fixed effect of valence, $b = -0.62$, $SE = 0.18$, $z = -3.52$, $p < 0.001$, such that positive targets ($M = 96\%$, 95% CI [95, 97]) were responded to with greater accuracy than both negative ($M = 92\%$, 95% CI [91, 93]) and neutral targets ($M = 92\%$, 95% CI [91, 93]). Mirroring the RT data, where English words elicited faster responses than Spanish words, regardless of valence, the accuracy analysis also showed higher response accuracy for English relative to Spanish targets (negative English: $M = 97\%$, 95% CI [96, 99] vs. negative Spanish: $M = 87\%$, 95% CI [86, 89]; neutral English: $M = 97\%$, 95% CI [95, 98] vs. neutral Spanish: $M = 87\%$, 95% CI [85, 88]; positive English: $M = 99\%$, 95% CI [97, 100] vs. positive Spanish: $M = 93\%$, 95% CI [92, 95]) (see Table 4 for the summary of the RT and accuracy data). Unlike the RT analysis, where no effects of dominance were obtained, here we found a fixed effect of dominance, $b = -0.05$, $SE = 0.01$, $z = -3.88$, $p < 0.001$ and a dominance x language interaction, $b = 0.05$, $SE = 0.01$, $z = 3.19$, $p < 0.01$, with English-dominant bilinguals responding more accurately to English (99%) than to Spanish targets (89.6%, $p < 0.001$). Conversely, Spanish-dominant and balanced bilinguals responded more accurately to Spanish targets (Spanish-dominant: 98.7%; balanced: 98.6%) than did English-dominant bilinguals (89.6%, $p < 0.001$).

Table 4. Means of response latencies in milliseconds and accuracy results (percentage of correct responses; SE in parentheses) for Spanish and English emotion words recorded in the lexical decision task (LDT).

Language of Target	English		Spanish	
	RT (ms)	ACC (%)	RT (ms)	ACC (%)
Positive	749 (36.5)	99 (0.008)	870 (31.6)	93 (0.008)
Negative	790 (36.5)	97 (0.008)	1033 (43.3)	87 (0.008)
Neutral	791 (36.5)	97 (0.008)	986 (43.4)	87 (0.008)

3.3. EEG Results

3.3.1. EPN

In the early EPN time window, the RM ANOVA revealed a main effect of language, $F(1,24) = 5.92$, $p = 0.017$, $\eta^2_p = 0.21$, with more pronounced amplitudes following English targets ($M = -0.98$ µV, 95% CI [−1.9, −0.1]) than Spanish targets ($M = 0.68$ µV, 95% CI [−0.68, 2.02]) (see Figure 2). Likewise, in the late EPN time window, there was a main effect of language, $F(1,25) = 4.77$, $p = 0.04$, $\eta^2_p = 0.160$, with English targets eliciting larger amplitudes ($M = -1.02$ µV, 95% CI [−3.1, 1.03]) than Spanish targets ($M = 0.09$ µV, 95% CI [−1.54, 1.72]) (Figure 3). No effect of valence was found. Likewise, dominance failed to yield significant effects in either early or late EPN time windows.

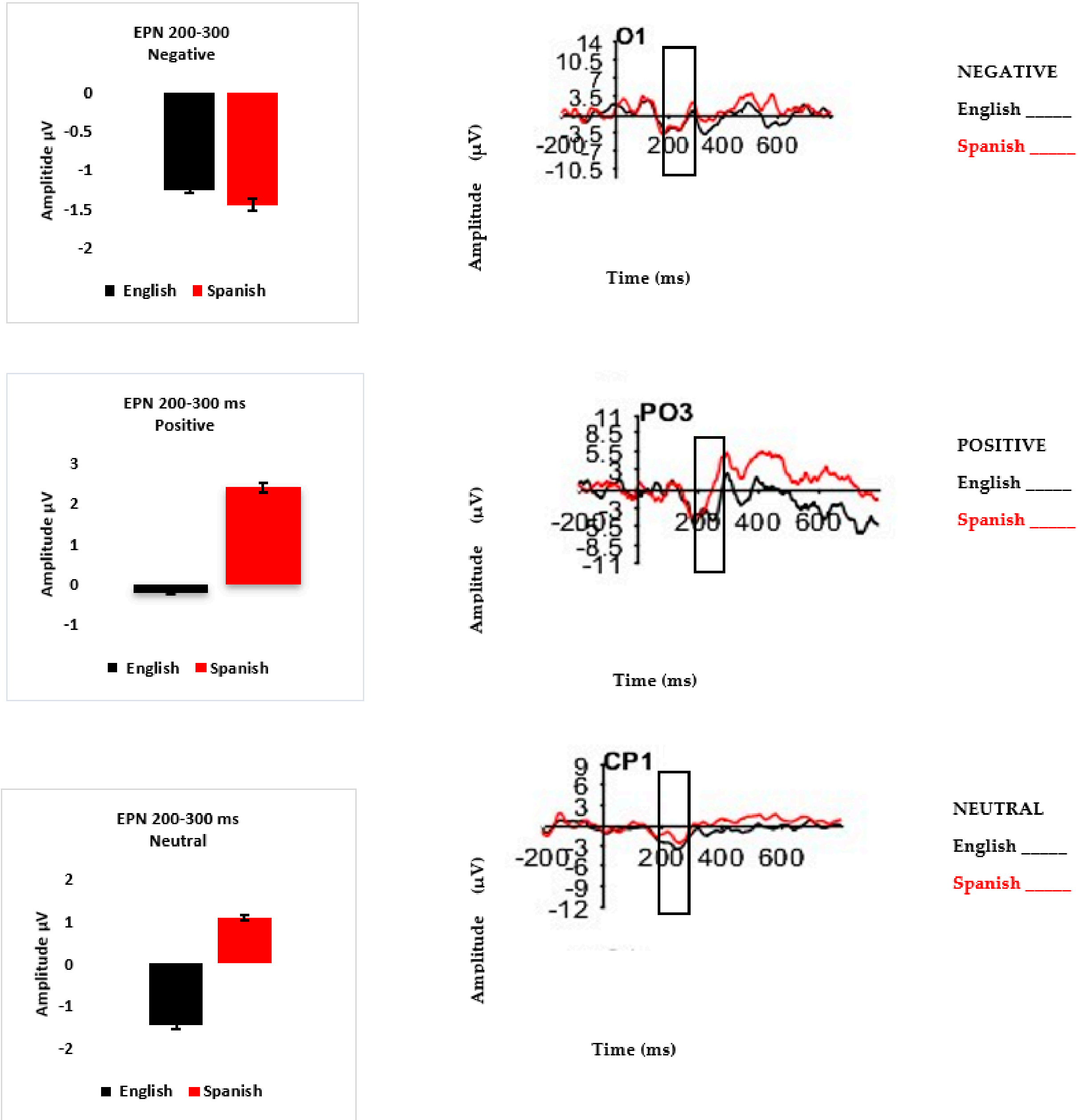

Figure 2. Amplitudes recorded for EPN 200–300 ms as a function of language of the target stimulus and valence. Representative electrodes O1, PO3, and CP1 illustrate differences in the EPN responses between English and Spanish negative, positive, and neutral targets. Bar plots show posterior EPN activity averaged across the posterior electrodes and the entire time window for EPN 200–300 ms. Error bars are standard errors.

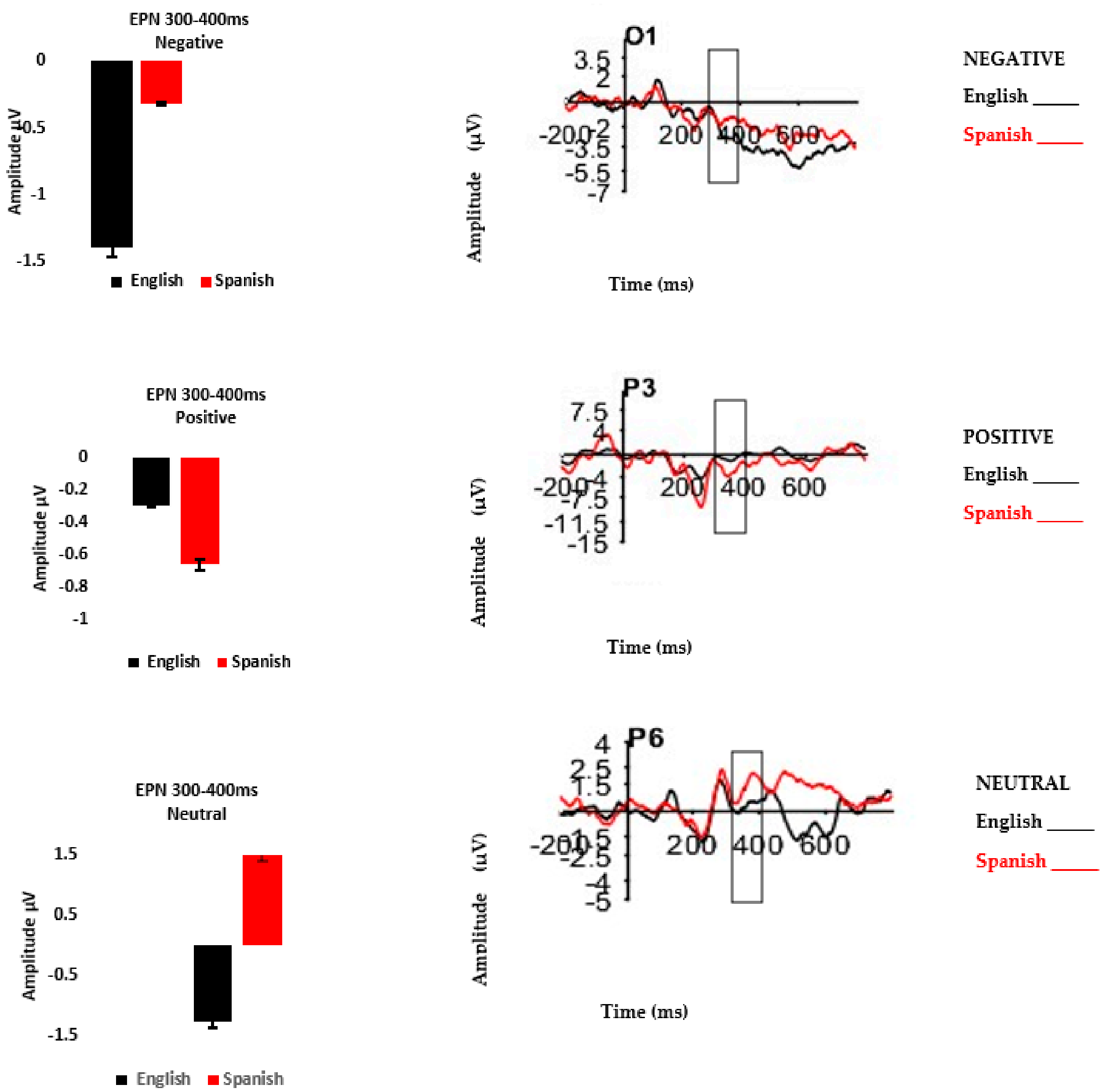

Figure 3. Amplitudes recorded for EPN 300–400 ms as a function of language of the target stimulus and valence. Representative electrodes O1, P3, and P6 illustrate differences in the EPN responses between English and Spanish negative, positive, and neutral targets. Bar plots show posterior EPN activity averaged across the posterior electrodes and the entire time window for EPN 300–400 ms. Error bars are standard errors.

3.3.2. LPC

In the LPC time window, there was a main effect of language, $F(1,24) = 7.07$, $p = 0.01$, $\eta^2_p = 0.23$, which revealed that Spanish targets evoked a more pronounced positivity ($M = 0.1$ µV, 95% CI [−1.06, 1.26]) than English targets ($M = -2.14$ µV, 95% CI [−3.99, −0.28]) (see Figure 4). Neither valence, $F(2,48) = 1.63$, $p = 0.93$, $\eta^2_p = 0.003$, nor language × valence interaction, $F(2,48) = 0.07$, $p = 0.93$, $\eta^2_p = 0.05$, turned out to be significant. Similar to the results reported in the early time windows, dominance did not show any significant effects.

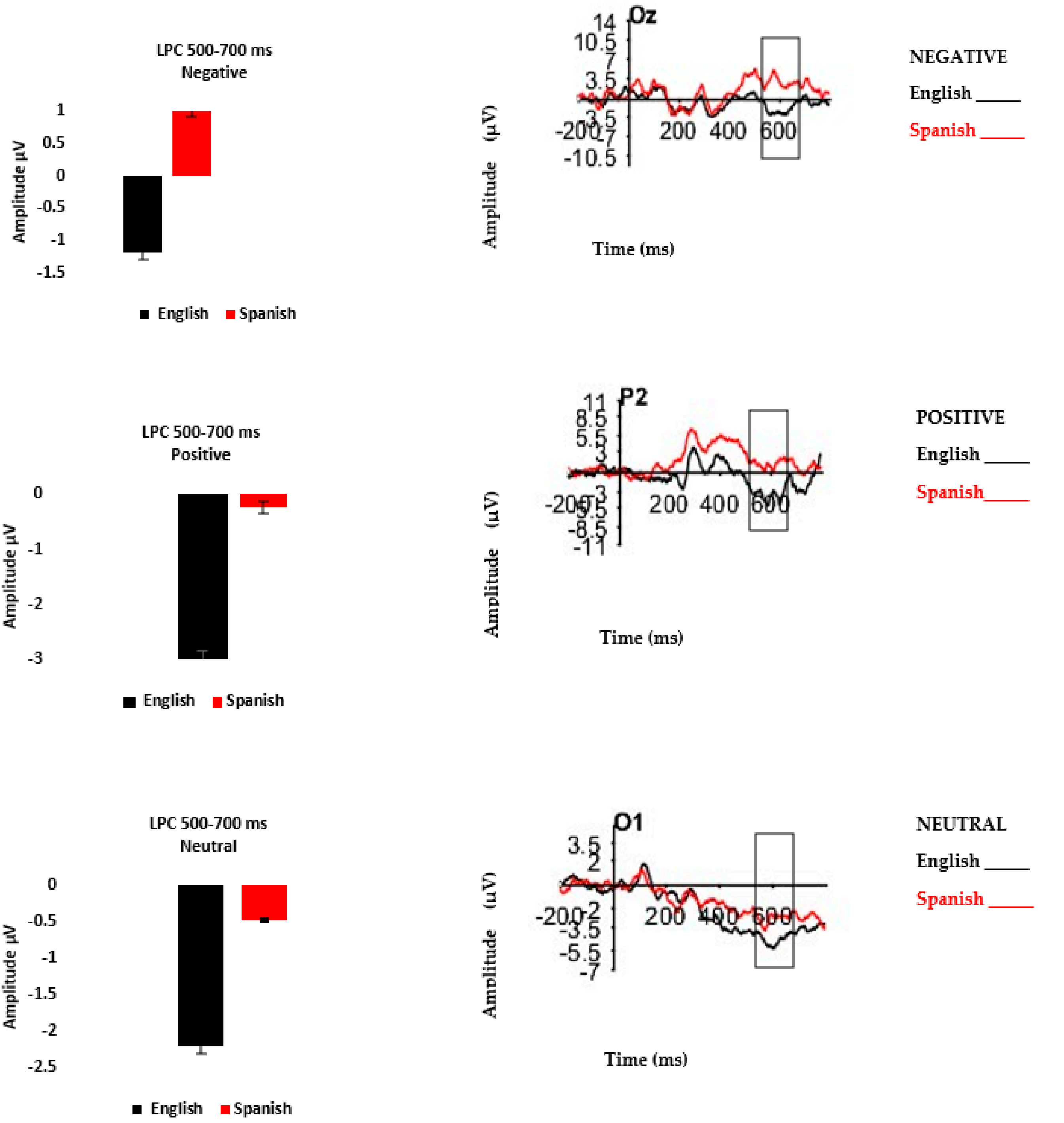

Figure 4. Amplitudes recorded for LPC 500–700 ms as a function of language of the target stimulus. Representative electrodes Oz, P2, and O1 illustrate differences in the LPC responses between English and Spanish negative, positive, and neutral targets. Bar plots show LPC activity averaged across the centro-parietal electrodes and the entire time window for LPC 500–700 ms. Error bars are standard errors.

4. Discussion

The current study aimed to shed light on the dynamics of emotion word processing in immersed Spanish-English/English-Spanish bilinguals who reside in a bilingual community and are routinely exposed to both languages in everyday personal and professional interactions. While the majority of the participants learned Spanish as their L1 or grew up speaking both languages, regardless of their L1, all bilinguals uniformly reported significantly higher proficiency in English than in Spanish. The participants were presented with English/Spanish emotion-label and emotion-laden (positive, negative) and neutral words along with nonwords and asked to make a lexical decision, while their RTs and ERP (EPN and LPC) responses were recorded. We asked the following research questions: (1) Do L1/L2 emotion processing differ qualitatively and/or quantitatively in highly proficient immersed bilinguals who are routinely exposed to both languages and reside in a bilingual community? (2) How do language dominance and AoA modulate L1/L2 emotion word processing?

4.1. Behavioral Results

The behavioral data revealed that English targets were responded to significantly faster than the Spanish, and this effect held true for both emotionally valenced and neutral stimuli. The accuracy data further confirmed the RT results, showing that English targets were responded to with a significantly greater accuracy than the Spanish. Dominance did not emerge as significant in the RT data. Lack of the modulating effect of dominance in our RT data is compatible with the study by Ferré et al. (2017). They looked at the effects of language status, task type, and word concreteness on the emotional content processing by Catalan-Spanish bilinguals who were early bilinguals highly fluent in both languages but dominant in Catalan. The task was either explicit (affective decision task, Exp.1) or automatic (LDT, Exp.2). Results showed effects of valence and concreteness for both the explicit and implicit tasks (Experiments 1 and 2). In the LDT, negative words took longer to process and elicited more errors than positive, and this effect held true regardless of the language of target stimuli. Hence, despite the participants' dominance in Catalan, the fact that they were all highly proficient in both languages seemed to contribute the most to their performance. Along similar lines, regardless of our bilingual participants' varying dominance, their responses were fastest and most accurate for English, their more proficient language than for Spanish.

In a subsequent experiment (Exp.3), Ferré et al. (2017) employed an LDT with a group of Catalan-Spanish bilinguals who were all late learners of English and dominant in Catalan. While the pattern of results was highly comparable for both Spanish and Catalan, the two languages that the bilingual participants grew up speaking and were immersed in, the effects diverged for English, the less proficient language acquired later in life and in a formal setting. Here, again, dominance did not seem to play a role but age/context of acquisition and proficiency were relevant. While our participant population was more varied, in that it included not only early but also late English-Spanish/Spanish-English bilinguals who were dominant either in their L1 or L2, the common characteristic of the participants in both studies was their high proficiency in the language(s) they performed best at, as well as a highly immersive context offering a rich bilingual and bicultural experience.

In addition, we found a robust emotion effect in our behavioral data such that positive targets were recognized significantly faster than the negative and neutral. This emotion effect was modulated by language of the target stimulus and present only in Spanish. Accordingly, positive Spanish targets were responded to faster compared to negative and neutral targets. In addition, on error rates, a significant valence effect was found with positive targets eliciting significantly fewer errors than either negative or neutral targets. This effect was again constrained by language, such that positive Spanish targets provoked the smallest, and negative/neutral the greatest, number of errors.

Our RT data replicate the facilitatory effect widely reported in LDT studies where RTs are faster for positively valenced over neutral words (e.g., Chen et al. 2015; Conrad et al. 2011; Hofmann et al. 2009; Kanske and Kotz 2007; Kousta et al. 2009; Kuchinke et al. 2005; Mueller and Kuchinke 2016; Recio et al. 2014) and over negative words (Briesemeister et al. 2011; Kanske and Kotz 2007; Kuchinke et al. 2005). Our accuracy data showing that positive words elicited significantly fewer errors than negative or neutral words are again consistent with the data reported in previous studies that employed the LDT

(Briesemeister et al. 2011; Chen et al. 2015; Conrad et al. 2011; Ferré et al. 2017; Kousta et al. 2009; Kuchinke and Lux 2012).

Results from the behavioral data showing an enhanced emotion effect for Spanish, as opposed to English suggest that L1 and L2 emotion word processing might diverge for immersed bilinguals. This absence of the emotion effect for English in our Spanish-English bilinguals might be viewed as indicative of the reduced L2 emotional resonance discussed earlier. For example, in an LDT with Chinese-English bilinguals, Chen et al. (2015) showed a diminished emotional impact of L2 as compared to L1 valenced words in both early and late ERP components. In their study, positive words elicited a larger EPN than neutral words and a smaller LPC than both neutral and negative words, but this emotion effect was only present for L1. Notably, our participants were predominantly L1 Spanish bilinguals dominant in English, so they differed substantially from the bilingual group in Chen et al.'s (2015) study, which employed Chinese-English bilinguals dominant in Chinese and residing in their L1 environment.

Increased RT to negative Spanish but not negative English targets present in our data is also compatible with studies showing a selected attenuated response to L2 negative stimuli (e.g., Jończyk et al. 2016; Wu and Thierry 2012). Wu and Thierry presented Chinese native speakers fluent in English with pairs of English words, some of which had a concealed sound repetition if translated into Chinese. Participants were asked to decide if the word pairs were related in meaning. While sound repetition priming elicited the expected effects for positive and neutral words, English words with a negative valence failed to automatically activate their Chinese translations, suggesting an inhibitory mechanism whereby a negative emotional content in L2 might be suppressed. Attenuated processing of L2 negative emotion words was further corroborated in Jończyk et al.'s (2016) experiment employing a context richer than single words. Late fluent Polish-English bilinguals residing in the UK read English and Polish sentences and indicated whether each sentence, which ended with either a semantically and affectively congruent or incongruent adjective, made sense. Results showed an increased N400 response to L1 Polish emotionally-valenced sentences and a reduction in the N400 amplitude for English sentences ending with negatively-valenced words, independent of semantic congruity.

The fact that our bilingual participants would display this reduced response to emotion words in English, the language in which they became more proficient, would seem to indicate that one's native language continues to be intrinsically more emotional even for highly immersed bilinguals. Those results are inconsistent with the suggestion that an increase in L2 proficiency might lead to a similar emotional sensitivity in bilingual's two languages (e.g., Costa et al. 2014; Jończyk et al. 2019). In their study adopting the "trolley dilemma" (Thomson 1985), Costa et al. (2014) found that the more bilinguals became proficient in their L2, the more likely their performance resembled that of L1 when making moral decisions, as opposed to less proficient bilinguals who would tend to be more moral in their L1 and more utilitarian in their L2.

However, as discussed earlier, emotion effects in L1 and L2 are likely affected not just by proficiency but by a complex interplay of such bilingual characteristics as AoA, the context of acquisition, frequency of daily usage, length of residence in the L2 speaking country, and possibly many other factors. In fact, some researchers have suggested that in order for bilinguals to have comparable emotional responses in L1 and L2, they need to be early AoA learners in addition to being highly proficient (see Harris et al. 2006). Indeed, the study by Harris (2004) has shown that L1 and L2 reprimands and taboo words elicited a comparable GSR in early but not sequential bilinguals, pointing to the possibility that AoA might be crucial in modulating the affective response of the autonomic nervous system (but see Ponari et al. 2015).

Overall, the behavioral analysis showed faster responses in bilinguals' more proficient (English) language, regardless of their dominance. However, in the accuracy analysis, dominance did appear significant, with Spanish-dominant and balanced bilinguals obtaining higher accuracy for Spanish than English targets and English-dominant bilinguals showing higher accuracy for English than for Spanish stimuli. The emotion effect was only observed for Spanish targets. Given that the majority of the bilingual participants were early learners of Spanish, either learning Spanish as L1 or simultaneously with English; this result is consistent with the idea that the first learned language might still evoke a stronger emotion effect than an L2, even if a bilingual person becomes more proficient or more dominant in their L2. As noted earlier, the study by Harris et al. (2003) with highly proficient Turkish speakers of English showed that while reactions to taboo words were identical in both L1 and L2, certain words (childhood reprimands) evoked a larger skin conductivity response in L1 only as compared to L2, suggesting that the language status might override proficiency in certain contexts.

4.2. Electrophysiological Results

4.2.1. L1 vs. L2 Emotion Word Processing

In the electrophysiological data, we found a significant effect of language in both early and late EPN time windows, such that English words elicited a larger negativity than Spanish. No emotion effects were present for either Spanish or English targets in either early or late EPN time windows. In the LPC 500–700 ms time window, a robust main effect of language was again present, manifesting a reverse pattern than that recorded for the EPN. Here, Spanish targets evoked a larger LPC positivity than English targets. A significant effect of language found on both early and late components in our study is generally compatible with findings from Vélez-Uribe and Rosselli (2021), although the pattern of our data diverges from theirs. Participants in Vélez-Uribe and Rosselli's (2021) study were Spanish-English balanced and unbalanced bilinguals highly comparable to our bilingual population, i.e., living immersed in a bicultural environment and receiving their education primarily in English. The EPN amplitude was found to be larger for Spanish than English targets across all valence categories, regardless of the participants' dominance. Vélez-Uribe and Rosselli (2021) suggest that an enhanced EPN in response to Spanish, as opposed to English words might be reflective of the overall higher proficiency of the bilingual participants in English compared to Spanish, thus evoking a larger negativity in the less proficient language. However, in later time windows, their study showed larger LPC amplitudes for English than Spanish words, regardless of the bilingual group.

Overall, discrepancies between our results and those of Vélez-Uribe and Rosselli (2021) might potentially be attributed to task-related demands. While the explicit, valence rating task employed by Vélez-Uribe and Rosselli (2021) might have favored the more proficient language by encouraging a deeper semantic processing in early time windows, an implicit lexical decision task used in the present study might be merely indicative of the automatic attention capture that the EPN typically reflects, without necessarily coinciding with early availability of the emotional content. More globally, lack of emotion effects in English despite an overall enhanced EPN response to English stimuli might possibly be related to our experimental design. Specifically, we used a fully randomized, mixed experimental design likely to have further weakened the strength of L1 and L2 valence effects. Typically, bilingual ERP studies into L1 and L2 emotion effects employ a blocked design for each language (e.g., Conrad et al. 2011; Jończyk et al. 2016; Kissler and Bromberek-Dyzman 2021; Optiz and Degner 2012). Since our bilingual participants were habitual codeswitchers who routinely engage in conversations where lexical items from Spanish and English are used interchangeably, for the sake of ecological validity we purposefully designed a study with a mixed design. Presence of both language stimuli was announced in the instructions and emphasized throughout the experimental set-up and practice block, with participants specifically told that they would see both Spanish and English words/nonwords.

Such a design, nevertheless, might have inadvertently led to brain responses related to an enhanced cognitive control which is called for when bilinguals have to process mixed language stimuli. Several ERP components have been identified as sensitive to codeswitching in contexts where bilingual participants are exposed to mixed language stimuli (see, Van Hell et al. 2018 for an overview). One component of interest to our study is an early frontal positivity (200–300 ms), which has been linked to attention shifts from the expected to unexpected language as well as from a narrow to a broad focus of attention (Beatty-Martínez and Dussias 2017). In a series of ERP experiments, Beatty-Martínez and Dussias (2017) examined whether bilinguals' codeswitching experience would have a modulating effect on the processing of codeswitched stimuli. To that effect, two groups of Spanish-English bilinguals were recruited: the first group routinely exposed to codeswitched speech by virtue of being immersed in a dual-language context and the second group consisting of bilinguals living in a single-language context devoid of the codeswitching experience. Participants were presented with preamble-target sentence pairs, with the first sentence providing supporting context and the second containing a target codeswitch involving an English noun with a Spanish determiner. While the two bilingual groups differed in their sensitivity to the switched targets, only non-codeswitchers manifested an early positivity for switched vs. non-switched conditions.

Beatty-Martínez and Dussias (2017) interpret these results as supporting Green and Wei's (2014) *Control Process* (CP) model which links language users' codeswitching behaviors to distinct control states in bilinguals. Whereas bilinguals in unilingual and bilingual contexts experience a competitive relationship between their languages on account of having to actively select one language only, this is not the case for bilinguals in dense codeswitching contexts where a cooperative relationship between their languages is present. Since the early positive component is an index of attentional control, Beatty-Martínez and Dussias (2017) propose that codeswitches trigger a shift of attention from a narrow, typical of a competitive control state, to a broad focus characterizing a cooperative control state. The early positivity can hence be viewed as an index of control, such that in appropriate contexts encouraging the activation and selection of both languages, attention would already be broad and no shift in attention from focused would be necessitated (Kaan et al. 2020).

Crucially, Beatty-Martínez and Dussias (2017) acknowledge that the early positivity effect could also reflect the overlapping N2 and P3 waves, which are present in this time window (200–300 ms) and suggest that the early frontal positivity might be a combination of P2-N2 and P3 components. Because of this overlap, lack of emotion effects in our EPN data might be attributed to the contamination from the competing codeswitch effects present in the mixed design trials, especially since the P3 component has also been associated with evaluation of the affective valence (see Zhang et al. 2014). Of interest is the question whether our bilinguals, who are habitual codeswitchers and should display a cooperative relationship between their languages, would still experience a codeswitch cost.

In a more recent ERP study relevant to this question, Kaan et al. (2020) examined whether a pro-active selection of both languages primed by the bilingual context would attenuate the early frontal positivity response for a codeswitch vs. no-switch control. They presented Spanish-English bilinguals with English sentences that were English only or contained a codeswitch from English to Spanish. While for one half of the study participants read the sentences together with an English monolingual who accompanied them, in the other half they did so with another Spanish-English bilingual. Consistent with the codeswitching literature, switches elicited an enhanced fronto-central positivity; however, the effect was attenuated in the bilingual condition where a Spanish-English bilingual accompanied a participant. These findings suggest that bilinguals expecting to operate in the bilingual context can accommodate codeswitches, in line with the dynamic control model of language processing.

However, the experimental setup in Kaan et al.'s (2020) study was very elaborate, including the presence of another monolingual/bilingual person and a joint reading task to ensure a strongly priming bilingual context. In their Experiment 1, which included bilinguals with a self-reported regular exposure to codeswitching but which lacked the manipulation of the bilingual context, Kaan et al. did find an enhanced frontal positivity in switch vs. non-switch trials. Kaan et al. (2020) suggest that, despite their participants' codeswitching experience, the use of switches in a written isolated context might not have been strong enough to engage the broad attentional focus that would eliminate the switch cost effect. Along the same lines, while our bilinguals came from the dense switching environment, their codeswitching practices are primarily executed in the spoken language mode and carried out in everyday conversations with Spanish-speaking family members and bilingual peers. The written language they are predominantly exposed to by virtue of their academic career is switch-free English.

In a bilingual ERP study directly relevant to our experimental setup, Christoffels et al. (2007) had German-Dutch bilinguals name pictures in their L1 and L2 in either blocked or mixed language conditions. The bilinguals were dominant in their L1 and switched languages routinely in their everyday lives. Switching costs manifested in the 275–375 ms and 375–475 ms time-windows. In the first time window (275–375 ms), an increased negativity was found for non-switch trials relative to blocked ones for both L1 and L2. The second time window affected mainly participants' L1 and resulted in more enhanced ERP modulations for blocked vs. mixed language conditions. Noteworthy is the fact that Christoffels et al.'s time windows overlap with those selected for our EPN measures (200–300 ms, 300–400 ms). Given that the lowest amplitudes in Christoffels et al.'s data were found for switch and the highest for blocked trials, our attenuated EPN responses might have been modulated by the mixed language condition where switch trials were predominant.

Another component of interest to our results, which has been consistently reported for codeswitched vs. control words, is the LPC (Moreno et al. 2002; Ng et al. 2014; Van Der Meij et al. 2011). LPC modulation in response to a mixed language design was found in Kaan et al.'s (2020) study with Spanish-English bilinguals described above, with switch trials eliciting a larger positivity than non-switch ones. Importantly, LPC switch effects have been found to be particularly prominent in higher proficiency bilinguals (Van Der Meij et al. 2011) and more robust for switches into the non-dominant language (Litcofsky and Van Hell 2017). Litcofsky and Van Hell (2017) used a self-paced reading paradigm with intrasentential codeswitches in both language directions with highly proficient Spanish-English bilinguals who were habitual codeswitchers. The participants were asked to read sentences which switched from L1 to L2 or in the opposite direction. Switched words elicited higher positivities than non-switched ones in the 500–900 ms LPC time window. While no significant differences were found between switched and non-switched sentences for switches into the dominant language, switches into the weaker language elicited a large posterior positivity. According to Litcofsky and Van Hell (2017), this switching cost asymmetry might relate to the fact that, when switching into the nondominant language, bilinguals would need to exercise more cognitive effort to activate their weaker language (cf. Green 1998). Consistent with those findings, enhanced LPC responses to Spanish vs. English targets found in our data might be partially attributed to the switch costs reported in the literature for the nondominant language. Since English was the more proficient language for our bilingual participants and the majority of them reported being dominant in English, a mixed design condition might have contributed to higher LPC amplitudes for Spanish, the weaker language.

Overall, since language control modulates the amplitude of the early P2, P3, and N2 and late LPC components and has been recorded in time windows overlapping with those we measured, our mixed design might have resulted in diminishing emotion effects due to the competing codeswitch effects in the data. While this is certainly a limitation in our study, it offers a valuable insight to take into account when planning future L1 and L2 studies with highly immersed bilinguals.

Crucial in addressing absence of the valence effect in our data are findings from Delaney-Busch et al. (2016), who examined how stimulus characteristics, such as valence and arousal as well as experimental task demands, affect the LPC response. Two ERP experiments were conducted, each with a different group of participants but with the identical set of stimuli. In Experiment 1, a semantic-monitoring task was used where neither valence nor arousal of the stimulus words were relevant for its successful completion. In this task, participants were asked to press a button if a word presented on the screen belonged to the category of animals. While judging the word category membership encourages deep semantic processing in that it requires participants to access semantic features of the target, the dimensions of valence and arousal are task-irrelevant. Results showed no effect of valence but a significant effect of arousal, such that high-arousal words elicited a larger LPC amplitude than low-arousal ones. On the other hand, in Experiment 2, where participants were instructed to make an explicit judgment regarding the valence of each stimulus word, the LPC showed a significant effect of valence with negative words eliciting the largest response, but no effect of arousal.

According to Delaney-Busch et al. (2016), these results can explain inconsistencies in the ERP literature where no effect of valence might be present on the LPC if stimuli are matched on arousal, as opposed to the strong LPC effect that would be recorded for valenced words which are high in arousal and hence likely to differ from low-arousal neutral ones (see also Recio et al. 2014). Importantly, as discussed earlier (Section 2.2), both positive and negative targets in our study were matched on arousal, with their ratings ranging from medium (4.5) to medium-high (7.5) on a 1–9 scale and the average arousal rating of $M = 5.8$ across all Spanish and English stimuli. Neutral words were medium arousal, ranging from 3.0–5.5, with an averaged mean for Spanish and English stimuli, $M = 4.4$. While arousal ratings for valenced vs. neutral stimuli differed significantly in the statistical analysis, the arousal level difference between the emotion and neutral words in our study was substantially smaller than that reported in the bilingual literature. For example, in Chen et al.'s (2015) study, ratings of arousal for positive ($M = 5.40$) and negative ($M = 5.41$) words were over twice higher than those for neutral ones ($M = 2.65$). Differences in arousal ratings coupled with task demands might hence explain absence of emotion effects in our ERPs data.

To sum up, with regard to our first research question, the present study revealed significant differences in behavioral and electrophysiological responses to English and Spanish words, with the emotion effect only present for Spanish in the RT data, the overall attenuation of the effect in the ERP data, and the divergent pattern of results for Spanish and English in early vs. late time windows.

4.2.2. Bilingual Characteristics Modulating Emotion Word Processing

Our second research question looked at the potential influence of bilingual participant characteristics, such as dominance and AoA, on emotion word processing. Given unequal numbers of participants in the early/late AoA groups, this variable was not explored. In turn, dominance failed to show a significant effect in either early or late time windows, with the EPN and LPC responses varying solely as a function of the target language. Accordingly, while participants' more proficient language (English) evoked more pronounced EPN amplitudes than the less proficient (Spanish), the reverse effect was found for the LPC time window.

Hence, similar to Vélez-Uribe and Rosselli's (2021) study, the results reported here seem to point to a crucial role that proficiency plays in bilingual emotion word processing. Interestingly, while both proficiency and dominance appeared significant in Vélez-Uribe and Rosselli's study, in our case dominance did not emerge as important. To further explore possible causes of the absence of the dominance effect in our data, we ran an *a posteriori* correlation analysis between participants' Spanish/English proficiency ratings and their dominance score. Ideally, we should expect proficiency ratings to be highly correlated with the dominance score in each of the participants' languages. As per the BDS coding, the higher the value on the scale, the more dominance in English it indicated. Conversely, the lower (more negative) the value on the BDS, the higher the dominance in Spanish. Scores of approximately 0 indicated a balanced bilingual. We found a significant negative correlation between proficiency in Spanish and participants' dominance score [$r(26) = -0.66$, $p < 0.001$], suggesting that higher proficiency in Spanish was also associated with Spanish dominance. In contrast, analysis with English proficiency failed to yield significant results [$r(26) = 0.26$, $p = 0.098$], implying that regardless of their dominance participants were all comparably highly proficient in English. A follow-up ANOVA run with dominance as a grouping variable and proficiency in English/Spanish as dependent variables, confirmed these results. English proficiency failed to reach significance in the analysis, with Spanish proficiency only marginally significant, $F(2,23) = 4.15$, $p = 0.01$. Post-hoc comparisons of English proficiency ratings between Spanish-dominant ($M = 6.45$, 95% CI [5.88, 7.02]), English-dominant ($M = 6.68$, 95% CI [6.35, 7.02]), and balanced bilinguals ($M = 6.29$, 95% CI [5.67, 6.92]) were all insignificant, whereas comparisons for Spanish yielded a significant difference only between English-dominant ($M = 4.33$, 95% CI [3.42, 5.25]) and Spanish-dominant bilinguals ($M = 6.50$, 95% CI [5.82, 7.18]; t_{Tukey} (23) = −2.76, $p < 0.05$), such that English-dominant bilinguals were significantly less proficient in Spanish than those who were Spanish-dominant. Overall, these results indicate that language proficiency might be a better predictor of L1/L2 emotion word processing than dominance.

Indeed, a LDT study by Ponari et al. (2015) seems to point to the superiority of proficiency over other participant characteristics, such as dominance, AoA or L1/L2 status in affective processing of L2 words. Bilingual participants in the study were recruited from diverse L1 families, including sign and non-Latin-script languages with a varying degree of typological distance from English. All bilinguals were highly proficient L2 speakers of English. Results of a LDT on negative, positive, and neutral words showed a comparable emotion facilitation effect for bilingual participants and native speakers of English, regardless of the bilinguals' varied L1 backgrounds, AoA, the degree of immersion, or the frequency and domain of L2 use. Likewise, based on the review of the functional neuroimaging studies using PET and fMRI to explore cerebral language organization in bilinguals during comprehension and production tasks, Abutalebi et al. (2001) suggest that proficiency is the most important factor affecting the bilingual language system, much more so than age of acquisition.

Despite its critical importance in bilingual studies, language proficiency has been notoriously difficult to objectively measure and conceptualize. Generally defined as the ability to use a language fluently, proficiency is viewed as a multidimensional construct subsuming linguistic components, such as phonology, orthography, morphology, syntax, and lexicon, in addition to pragmatic, sociolinguistic, and discourse-level features (De Souza and Silva 2015).

Weak correlation between our participants' dominance and proficiency points to a larger question identified in the bilingual literature, namely, the employment of self-assessment proficiency measures. In their review of 140 empirical papers published in the journal *Bilingualism: Language and Cognition* between 1998–2011, Hulstijn (2012) notices that over half of them included self-assessment of language proficiency (LP) rather than an objective LP test as an independent variable; yet, participants' LP scores were seldom applied to explaining variance obtained in the dependent variables (see also De Souza and Silva 2015).

Despite these criticisms, self-ratings have been extensively used to assess bilingual language proficiency, and multiple studies have shown highly robust correlations between self-ratings and such objective proficiency measures as reading/auditory comprehension, reading fluency, grammaticality judgment speed/accuracy, picture naming, receptive vocabulary, and sound awareness (see Marian et al. 2007). Indeed, the Language History Questionnaire (LHQ; Li et al. 2006, 2019) employed in the present study has been widely used in the bilingual literature to examine language proficiency and the background of bi/multilingual language users, and its scores have been validated with objective measures of proficiency, such as, for example, verbal fluency (Li et al. 2019). Apart from the self-assessment module, where participants rate their proficiency in reading, writing, speaking, and listening, the LHQ examines participant's AoA, language of instructed education, length of using the languages, the frequency of daily language use and language mixing, the current country of residence, as well as language preference and cultural identity.

More nuanced than the reliability of self-assessment ratings, however, is the issue of the correlation between self-ratings of proficiency and dominance, as well as bilingual participants' ability to classify themselves into dominance groups. In the study more directly relevant to our results, Gollan et al. (2012) looked at the usefulness of proficiency self-ratings for establishing participants' spoken language dominance. In order to obtain objective measures of spoken proficiency, 52 young and 20 aging Spanish-English bilinguals were interviewed in each language using a structured oral proficiency interview and completed a picture naming test in each language. In addition, participants self-rated their language proficiency using a 10-point scale ranging from (1) novice low to (10) superior. Based on participants' performance on each of the measures, Gollan et al. (2012) calculated an index score that reflected the degree of balanced bilingualism. This index was obtained by dividing the lower score obtained in whichever language by the higher one for each measure. For example, a participant who rated themselves as superior (10) in English and intermediate high (6) in Spanish, would be classified as 60% bilingual. Thus, the index scores reflected the degree to which knowledge of each language was similar, regardless of the direction of dominance.

Results revealed a significant correlation between self-reported proficiency in English and objective measures, such as the oral proficiency and naming tests. Similarly, correlations between self-reported level of proficiency in Spanish, which was the nondominant language for most participants, and the objective measures of proficiency was high. On the other hand, the correlations between self-rated and objective index scores were only marginally significant, suggesting that while bilinguals were fairly accurate in assessing which of their two languages is more dominant than the other, they were much less accurate in estimating the degree of difference between proficiency in each language. For example, bilinguals who rated their proficiency as equal in English and Spanish were later shown to perform better in English on both the interview and naming tasks. Similarly, both Spanish- and English-dominant bilinguals, especially the young participants, tended to overestimate their abilities in their dominant language. While some amount of overestimation in proficiency self-ratings might have been present in our study, the fact that our bilinguals consistently self-rated as more proficient in English and responded consistently faster to English than Spanish targets seems to indicate a high degree of overlap between subjective (the LHQ) and objective (LDT) measures of language proficiency. In turn, low correlations between proficiency ratings and dominance scores might be partially a product of unequal comparison groups, with the majority of our bilinguals (16) reporting dominance in English, and only 5 in Spanish.

More generally, our results that do not fit neatly into the existing L2 emotion processing literature can be attributed to the uniqueness of our bilingual population consisting of habitual codeswitchers whose L1 has ceased to be their dominant language. It has been suggested that neurocognitive mechanisms of language use and control in such bilinguals might differ qualitatively from those in non-habitual codeswitchers (e.g., Green 2011; Green and Wei 2014). In line with this assumption, Pliatsikas et al. (2017) suggest that the major mechanism shaping cortical regions in the bilingual brain is a continuous L2 usage in an immersive environment. Pliatsikas et al. (2017) acquired brain scan images from 20 sequential (late) learners of English varying in their L1 backgrounds and residing in the English-speaking country for an average of almost 11 years. Significant subcortical reshaping of the basal ganglia and thalamus was visible, mirroring the data obtained earlier in simultaneous (early) bilinguals. Since the participants in Pliatsikas et al.'s (2017) study were all late bilinguals, the authors suggest that structural changes in the bilingual brain are primarily modulated by the amount of L2 immersion. In line with this suggestion, the time spent in the UK turned out to be a significant predictor for the expansion of the right globus pallidus, a nucleus in the basal ganglia. Analyses with proficiency and AoA showed no significant effects, indicating that brain restructuring in bilinguals depends on the active and continuous usage of L2 in the immersive context. Thus, immersion emerges as a crucial factor to consider when comparing L1 and L2 emotion effects in bilingual participants.

Finally, regardless of an individual bilingual's dominance, L1/L2 learning history, proficiency, AoA, or immersive experience, bilinguals' two languages might differ in their sensitivity to emotional content depending on the context of use, as would be the case when one language is primarily used for professional purposes and another at home in a less formal and more emotion-laden context. In this view, words develop emotional resonances depending on the intensity of the emotional context in which they were first learned and subsequently used throughout the bilingual's life experiences. The importance of context-dependent emotional learning is captured by the *emotional context of learning hypothesis* (Caldwell-Harris 2014). Briefly, the hypothesis postulates that emotional resonance is affected by the context in which language is learned and that language will be perceived as more emotional if it has been acquired and used in emotional settings. This is akin to the *contextual-learning hypothesis* (Barrett et al. 2007) emphasizing the interplay between learning and experience. Individual's personal experiences will shape emotional processing in each language and modulate the vividness of emotional reaction. For non-immersed bilinguals, who are typically unbalanced, more proficient in their L1 and who often learned L2 in a more formal, emotionally neutral context, L1 and L2 emotional resonance might substantially differ. Immersed bilinguals recruited in our study have likely experienced both languages in emotionally-grounded contexts at various stages of their linguistic and cognitive development where linguistic information was strongly linked with emotional experiences. Such varying individual experiences in each language might thus be another factor constraining the strength of the emotion effect as measured by behavioral and electrophysiological data.

An important limitation in our study is a lack of equal groups of AoA early/late English/Spanish bilinguals, which prevented us from assessing the relevance of AoA in a full analysis. To adequately assess the effect of each of the bilingual characteristics on emotion processing in L1 and L2, we would need a fully-crossed design with multiple groups of participants varying in terms of their AoA, L1/L2 status, dominance, proficiency, the degree of their immersive experience, and possibly other relevant factors, such as frequency of codeswitching or the emotional context of language usage. The border town from which our Spanish-English/English-Spanish participants were recruited is an essentially rich bicultural and bilingual community where both cultures and languages are tightly interwoven and both languages spoken interchangeably on a regular basis. Comparing the results from such a population against those reported for L1-dominant bilinguals traditionally employed in emotion processing studies is hence challenging

5. Conclusions

In conclusion, the present study examined behavioral and electrophysiological correlates of L1 and L2 emotion processing in immersed highly proficient Spanish-English/English-Spanish bilinguals residing in the bilingual community characterized by dense codeswitching practices. We wanted to see whether there would be any qualitative or quantitative differences between L1 and L2 emotion word processing and whether bilingual participant characteristics, such as dominance or AoA would constrain the L1/L2 emotion effects. Behavioral data showed faster and more accurate responses to English than Spanish targets, reflecting the fact that all of the bilinguals participating in the study were more proficient in English than in Spanish. However, the emotion effect was only present for Spanish, which was the first language for the overwhelming majority of our participants. Electrophysiological data showed a significant effect of language, such that early and late EPN responses were more pronounced for English than Spanish, with the reverse effect found on the LPC component, where Spanish targets elicited a higher positivity than those that were English. Dominance did not turn out to be a significant predictor of bilingual performance. Overall, emotion word processing in highly proficient immersed bilinguals might reflect a complex interaction of a number of participant factors, such as proficiency, AoA, the length of the immersive experience, or individual histories with each of the languages and how they were grounded in the emotional context. Further research with more diverse bilingual populations and a wider range of tasks is needed to more accurately assess the dynamic interaction of the various participant characteristics in the course of L1/L2 emotion word processing.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/languages8010042/s1. Experimental Stimuli. ACC_LGTARGETXVALENCEXDSCORE analysis. RT_LGTARGETXVALENCEXDSCORE analysis.

Author Contributions: Conceptualization, A.B.C.; data curation, A.B.C. and B.L.G.; formal analysis, A.B.C. and B.L.G.; investigation, A.B.C. and B.L.G.; methodology, A.B.C.; project administration, A.B.C. and B.L.G.; software, A.B.C. and B.L.G.; supervision, A.B.C.; validation, A.B.C. and B.L.G.; visualization, A.B.C.; writing—original draft, A.B.C.; writing—review and editing, A.B.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the TAMIU Advancing Research and Curriculum Initiative (TAMIU ARC) awarded by the US Department of Education Developing Hispanic Serving Institutions Program (Award # P031S190304) and by the NSF BCS Division Of Behavioral and Cognitive Science. MRI: Acquisition of a Biosemi Event Related Potentials Active Two Acquisition System to Enhance Research and Training at Texas A&M International University (Award # 1229123).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Texas A&M International University (protocol code 2019-02-13; date of approval: 11 March 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available from the corresponding author upon request.

Acknowledgments: We want to thank the following TAMIU students for their help in administering the experiment: Alexandra Reyes, Ruby Salas, Rebeca Salazar, Devon Nuñez, Mariella Soto Ruiz, Jessica Garza, Alexandra Rodriguez, and Samantha Andrade.

Conflicts of Interest: The authors declare no conflict of interest.

References

Abutalebi, Jubin, Stefano F. Cappa, and Daniela Perani. 2001. The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition* 4: 179–90. [CrossRef]

Altarriba, Jeanette, and Dana M. Basnight-Brown. 2010. The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism* 15: 310–28. [CrossRef]

Altarriba, Jeanette, and Lisa M. Bauer. 2004. The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *The American Journal of Psychology* 117: 389–410. [CrossRef] [PubMed]

Altarriba, Jeanette, and Tina M. Canary. 2004. The influence of emotional arousal on affective priming in monolingual and bilingual speakers. *Journal of Multilingual and Multicultural Development* 25: 248–65. [CrossRef]

Altarriba, Jeanette. 2006. Cognitive approaches to the study of emotion-laden and emotion words in monolingual and bilingual memory. *Bilingual Education and Bilingualism* 56: 232.

Anooshian, Linda J., and Paula T. Hertel. 1994. Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion* 8: 503–14. [CrossRef]

Ayçiçeği, Ayşe, and Catherine Harris. 2004. BRIEF REPORT Bilinguals' recall and recognition of emotion words. *Cognition and Emotion* 18: 977–87. [CrossRef]

Ayçiçeği-Dinn, Ayşe, and Catherine L. Caldwell-Harris. 2009. Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition* 12: 291303. [CrossRef]

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–78. [CrossRef]

Barrett, Lisa F., Kristen A. Lindquist, and Maria Gendron. 2007. Language as context for the perception of emotion. *Trends in Cognitive Science* 11: 327–32. [CrossRef]

Bayer, Mareike, Werner Sommer, and Annekathrin Schacht. 2010. Reading emotional words within sentences: The impact of arousal and valence on event-related potentials. *International Journal of Psychophysiology* 78: 299–307. [CrossRef] [PubMed]

Beatty-Martínez, Anne L., and Paola E. Dussias. 2017. Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language* 95: 173–89. [CrossRef]

Bernat, Edward, Scott Bunce, and Howard Shevrin. 2001. Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *International Journal of Psychophysiology* 42: 11–34. [CrossRef]

Bond, Michael H., and Tat-ming Lai. 1986. Embarrassment and code-switching into second language. *Journal of Social Psychology* 126: 179–86.

Bradley, Margaret M., and Peter J. Lang. 1999. *Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings*. Technical Report C-2. Gainesville: University of Florida.

Briesemeister, Benny B., Lars Kuchinke, and Arthur M. Jacobs. 2011. Discrete emotion norms for nouns: Berlin affective word list (DENN–BAWL). *Behavior Research Methods* 43: 441–48. [CrossRef] [PubMed]
Brysbaert, Marc, Amy B. Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46: 904–11. [CrossRef] [PubMed]
Brysbaert, Marc, Boris New, and Emmanuel Keuleers. 2012. Adding Part-of-Speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods* 44: 991–97. [CrossRef] [PubMed]
Caldwell-Harris, Catherine L. 2014. Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology* 5: 1055. [CrossRef]
Caldwell-Harris, Catherine L. 2015. Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science* 24: 214–19. [CrossRef]
Chen, Peiyao, Jie Lin, Bingle Chen, Chunming Lu, and Taomei Guo. 2015. Processing emotional words in two languages with one brain: ERP and fMRI evidence from Chinese-English bilinguals. *Cortex* 71: 34–48. [CrossRef]
Christoffels, Ingrid K., Christine Firk, and Niels O. Schiller. 2007. Bilingual language control: An event-related brain potential study. *Brain Research* 1147: 192–208. [CrossRef] [PubMed]
Citron, Francesca M. 2012. Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language* 122: 211–26. [CrossRef] [PubMed]
Citron, Francesca M., Brendan S. Weeks, and Evelyn C. Ferstl. 2011. Neural correlates of emotion word processing: The interaction between emotional valence and arousal. *Psychophysiology* 48: S3–S3.
Citron, Francesca M., Brendan S. Weeks, and Evelyn C. Ferstl. 2013. Effects of valence and arousal on written word recognition: Time course and ERP correlates. *Neuroscience Letters* 533: 90–95. [CrossRef] [PubMed]
Colbeck, Katie L., and Jeffrey S. Bowers. 2012. Blinded by taboo words in L1 but not L2. *Emotion* 12: 217–22. [CrossRef] [PubMed]
Conrad, Markus, Gillermo Recio, and Arthur M. Jacobs. 2011. The time course of emotion effects in first and second language processing: A cross cultural ERP study with German-Spanish bilinguals. *Frontiers in Psychology* 2: 351. [CrossRef]
Costa, Albert, Alice Foucart, Sayuri Hayakawa, Melina Aparici, Jose Apesteguia, Joy Heafner, and Keysar Boaz. 2014. Your morals depend on language. *PLoS ONE* 9: e94842. [CrossRef] [PubMed]
Cuetos, Fernando, Maria Glez-Nosti, Analía Barbón, and Marc Brysbaert. 2011. SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicologica* 33: 133–43. Available online: https://www.redalyc.org/articulo.oa?id=16923102001 (accessed on 1 February 2019).
Cuthbert, Bruce N., Harald T. Schupp, Margaret M. Bradley, Niels Birbaumer, and Peter J. Lang. 2000. Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology* 52: 95–111. [CrossRef]
De Souza, Ricardo, and Jesiel Soares Silva. 2015. Exploring the measurement of vocabulary size to differentiate Brazilian Portuguese-English bilinguals' access to grammatical knowledge in the L2. *Revista Linguística* 11: 187–214. [CrossRef]
Delaney-Busch, Nathaniel, Gianna Wilkie, and Gina Kuperberg. 2016. Vivid: How valence and arousal influence word processing under different task demands. *Cognitive, Affective, and Behavioral Neuroscience* 16: 415–32. [CrossRef] [PubMed]
Delorme, Arnaud, and Scott Makeig. 2004. EEGLAB: An open-source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134: 9–21. [CrossRef]
Dunn, Alexandra L., and Jean E. Fox Tree. 2009. A quick, gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition* 12: 273–89. [CrossRef]
Eilola, Tiina M., and Jelena Havelka. 2011. Behavioral and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *International Journal of Bilingualism* 15: 353–69. [CrossRef]
Eilola, Tiina M., Jelena Havelka, and Dinkar Sharma. 2007. Emotional activation in the first and second language. *Cognition and Emotion* 21: 1064–76. [CrossRef]
Estes, Zachary, and James S. Adelman. 2008. Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006). *Emotion* 8: 441–44. [CrossRef] [PubMed]
Ferré, Pilar, Manuel Anglada-Tort, and Marc Guash. 2017. Processing of emotional words in bilinguals: Testing the effects of word concreteness, task type and language status. *Second Language Research* 34: 371–94. [CrossRef]
Ferré, Pilar, Teófilo García, Isbel Fraga, Rosa Sánchez-Casas, and Margarita Molero. 2010. Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and second language? *Cognition and Emotion* 24: 760–85. [CrossRef]
Fischler, Ira, and Margaret Bradley. 2006. Event-related potential studies of language and emotion: Words, phrases, and task effects. *Progress in Brain Research* 156: 185–203. [CrossRef]
Gibbons, Henning. 2009. Evaluative priming from subliminal emotional words: Insights from event-related potentials and individual differences related to anxiety. *Consciousness and Cognition* 18: 383–400. [CrossRef]
Gollan, Tamaer H., Gali H. Weissberger, Elin Runnqvist, Rosa I. Montoya, and Cynthia M. Cera. 2012. Self-ratings of spoken language dominance: A Multi-Lingual Naming Test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism (Cambridge, England)* 15: 594–615. [CrossRef]
Gonzales-Reigosa, Fernando. 1976. The anxiety-arousing effect of taboo words in bilinguals. In *Cross-Cultural Anxiety*. Edited by Charles D. Spielberger and R. Diaz-Guerrero. London: Taylor & Francis, pp. 89–105.
Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1: 67–81. [CrossRef]

Green, David W. 2011. Language control in different contexts: The behavioral ecology of bilingual speakers. *Frontiers in Psychology* 2: 103. [CrossRef] [PubMed]

Green, David W., and Li Wei. 2014. A control process model of code-switching. *Language, Cognition and Neuroscience* 29: 499–511. [CrossRef]

Harris, Catherine L. 2004. Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development* 25: 223–47. [CrossRef]

Harris, Catherine L., Ayşe Ayçiçeği, and Jean B. Gleason. 2003. Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics* 24: 561–79. [CrossRef]

Harris, Catherine L., Ayşe Ayçiçeği, and Jean B. Gleason. 2006. When is a first language more emotional? Psychophysiological evidence from bilingual speakers. *Bilingual Education and Bilingualism* 56: 257–83.

Harris, Scott R. 2015. *An Invitation to the Sociology of Emotions*. London: Routledge.

Herbert, Cornelia, Johanna Kissler, Markus Junghöfer, Peter Peyk, and Brigitte Rockstroh. 2006. Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology* 43: 197–206. [CrossRef]

Herbert, Cornelia, Markus Junghofer, and Johanna Kissler. 2008. Event related potentials to emotional adjectives during reading. *Psychophysiology* 45: 487–98. [CrossRef]

Heredia, Roberto R., and Anna B. Cieślicka. 2014. Bilingual memory storage: Compound-coordinate and derivatives. In *Foundations of Bilingual Memory*. Edited by Roberto R. Heredia and Jeanette Altarriba. Berlin and Heidelberg: Springer, pp. 11–39. [CrossRef]

Hinojosa, José A., Constantino Méndez-Bértolo, and Miguel A. Pozo. 2010. Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology* 47: 748–57. [CrossRef]

Hinojosa, José A., Jacobo Albert, Sara López-Martín, and Luis Carretié. 2014. Temporospatial analysis of explicit and implicit processing of negative content during word comprehension. *Brain and Cognition* 87: 109–21. [CrossRef]

Hinojosa, José A., Natalia Martínez-García, Cristina Villalba-García, Uxia Fernández-Folgueiras, Alberto Sánchez-Carmona, Miguel A. Pozo, and Pedro R. Montoro. 2016. Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavioral Research Methods* 48: 272–84. [CrossRef]

Hofmann, Markus J., Lars Kuchinke, Sascha Tamm, Melissa L. H. Võ, and Arthur M. Jacobs. 2009. Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, and Behavioral Neuroscience* 9: 389–97. [CrossRef]

Holt, Daphne J., Spencer K. Lynn, D. Kreher, Scott Rauch, Phillip J. Holcomb, and Gina R. Kuperberg. 2005. Neurophysiological correlates of reading emotional sciences. Poster presented at 12th Annual Cognitive Neuroscience Society Meeting, New York, NY, USA, April 12.

Hulstijn, Jan H. 2012. The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition* 15: 422–33. [CrossRef]

Iacozza, Sara, Albert Costa, and Jon A. Duñabeitia. 2017. What do your eyes reveal about your foreign language? Reading emotional sentences in a native and foreign language. *PLoS ONE* 12: e0186027. [CrossRef] [PubMed]

Imbir, Kamil K., Tomasz Spustez, and Jarosław Żygierewicz. 2016. Effects of valence and origin of emotions in word processing evidenced by event related potential correlates in a lexical decision task. *Frontiers in Psychology* 7: 271. [CrossRef] [PubMed]

Jankowiak, Katarzyna, and Pawel Korpal. 2017. On modality effects in bilingual emotional language processing: Evidence from galvanic skin response. *Journal of Psycholinguistic Research* 47: 663–67. [CrossRef]

Jończyk, Rafał, Bastien Boutonnet, Kamil Musiał, Katie Hoemann, and Guillaume Thierry. 2016. The bilingual brain turns a blind eye to negative statements in the second language. *Cognitive, Affective, and Behavioral Neuroscience* 16: 527–40. [CrossRef]

Jończyk, Rafał, Inga Korolczuk, Evangelia Balatsou, and Guillaume Thierry. 2019. Keep calm and carry on: Electrophysiological evaluation of emotional anticipation in the second language. *Social Cognitive and Affective Neuroscience* 14: 885–98. [CrossRef]

Junghöfer, Markus, Margaret M. Bradley, Thomas R. Elbert, and Peter J. Lang. 2001. Fleeting images: A new look at early emotion discrimination. *Psychophysiology* 38: 175–78. [CrossRef]

Kaan, Edith, Souad Kheder, Ann Kreidler, Aleksandra Tomić, and Jorge R. Valdés Kroff. 2020. Processing code-switches in the presence of others: An ERP study. *Frontiers in Psychology* 11: 1288. [CrossRef]

Kanske, Philipp, and Sonja A. Kotz. 2007. Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research* 1148: 138–48. [CrossRef] [PubMed]

Kazanas, Stephanie A., and Jeanette Altarriba. 2016. Emotion word processing: Effects of word type and valence in Spanish–English bilinguals. *Journal of Psycholinguistic Research* 45: 395–406. [CrossRef] [PubMed]

Kim, Myung-Sun. 1993. The Effects of Emotional Connotations of English Words on Event-Related Potentials (ERPs) in Korean-English Bilinguals and English-Speaking Monolinguals. Doctoral dissertation, University of Georgia, Athens, GA, USA.

Kissler, Johanna, and Cornelia Herbert. 2013. Emotion, etmnooi, or emitoon?: Faster lexical access to emotional than to neutral words during reading. *Biological Psychology* 92: 464–79. [CrossRef] [PubMed]

Kissler, Johanna, and Katarzyna Bromberek-Dyzman. 2021. Mood induction differently affects early neural correlates of evaluative word processing in L1 and L2. *Frontiers in Psychology* 11: 3809. [CrossRef]

Kissler, Johanna, Cornelia Herbert, Irene Winkler, and Markus Junghöfer. 2009. Emotion and attention in visual word processing: An ERP study. *Biological Psychology* 80: 75–83. [CrossRef]

Kissler, Johanna, Cornelia Herbert, Peter Peyk, and Markus Junghöfer. 2007. Buzzwords—Early cortical responses to emotional words during reading. *Psychological Science* 18: 475–80. [CrossRef] [PubMed]
Kissler, Johanna, Ramin Assadollahi, and Cornelia Herbert. 2006. Emotional and semantic networks in visual word processing: Insights from ERP studies. *Progress in Brain Research* 156: 147–83. [PubMed]
Kousta, Stavroula-Thaleia, David P. Vinson, and Gabriella Vigliocco. 2009. Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112: 473–81. [CrossRef]
Kuchinke, Lar, Aurthur M. Jacobs, Claudia Grubich, Melissa L.-H. Võ, Markus Conrad, and Manfred Herrmann. 2005. Incidental effects of emotional valence in single word processing: An fMRI study. *Neuroimage* 28: 1022–32. [CrossRef] [PubMed]
Kuchinke, Lars, and Vanessa Lux. 2012. Caffeine improves left hemisphere processing of positive words. *PLoS ONE* 7: e48487. [CrossRef]
Kuchinke, Lars, Melissa L.-H. Võ, Markus Hofmann, and Arthur M. Jacobs. 2007. Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology* 65: 132–40. [CrossRef] [PubMed]
Lang, Peter J., Margaret M. Bradley, and Bruce N. Cuthert. 1997. Motivated attention: Affect, activation, and action. *Attention and Orienting: Sensory and Motivational Processes* 97: 135.
Larsen, Randy J., Kimberly A. Mercer, and David A. Balota. 2006. Lexical characteristics of words used in emotional Stroop experiments. *Emotion* 6: 62–72. [CrossRef] [PubMed]
Li, Ping, Fan Zhang, Anya Yu, and Xiaowei Zhao. 2019. Language history questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition* 23: 938–44. [CrossRef]
Li, Ping, Sara Sepanski, and Xiaowei Zhao. 2006. Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods* 38: 202–10. [CrossRef]
Litcofsky, Kaitlyn A., and Janet G. Van Hell. 2017. Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97: 112–39. [CrossRef]
Lopez-Calderon, Javier, and Steven J. Luck. 2014. ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience* 8: 1–14. [CrossRef]
Makeig, Scott, Anthony Bell, Tzyy-Ping Jung, and Terrence J. Sejnowski. 1995. Independent component analysis of electroencepha lographic data. *Advances in Neural Information Processing Systems* 8: 145–51.
Marian, Viorica, Henrike K. Blumenfeld, and Margarita Kaushanskaya. 2007. The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50: 940–67. [CrossRef]
Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen, and Douglas Bates. 2017. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94: 305–15. [CrossRef]
Moreno, Eva M., Kara D. Federmeier, and Marta Kutas. 2002. Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain Language* 80: 188–207. [CrossRef] [PubMed]
Mueller, Christina J., and Lars Kuchinke. 2016. Individual differences in emotion word processing: A diffusion model analysis. *Cognitive, Affective, and Behavioral Neuroscience* 16: 489–501. [CrossRef] [PubMed]
Ng, Shukhan, Christian Gonzalez, and Nicole Y. Wicha. 2014. The fox and the cabra: An ERP analysis of reading code switched nouns and verbs in bilingual short stories. *Brain Research* 1557: 127–40. [CrossRef] [PubMed]
Oldfield, Richard C. 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97–113. [CrossRef] [PubMed]
Optiz, Bertram, and Juliane Degner. 2012. Emotionality in a second language: It's a matter of time. *Neuropsychologia* 50: 1961–67. [CrossRef]
Palazova, Marina, Katharina Mantwill, Werner Sommer, and Annekathrin Schact. 2011. Are effects of emotion in single words non-lexical? *Evidence from event-related brain potentials. Neuropsychologia* 49: 2766–75. [CrossRef]
Palazova, Marina, Katharina Mantwill, Werner Sommer, and Annekathrin Schact. 2013. Interplay of emotional valence and concreteness in word processing: An event-related potential study with verbs. *Brain and Language* 125: 264–71. [CrossRef]
Pavlenko, Aneta. 2012. Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology* 47: 405–28. [CrossRef]
Pliatsikas, Christos, Vincent DeLuca, Elisavet Moschopoulou, and James D. Saddy. 2017. Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function* 222: 1785–95. [CrossRef]
Ponari, Marta, Sara Rodríguez-Cuadrado, David Vinson, Neil Fox, Albert Costa, and Gabriella Vigliocco. 2015. Processing advantage for emotional words in bilingual speakers. *Emotion* 15: 644–52. [CrossRef] [PubMed]
R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. (Version 4.0. R Packages Retrieved from MRAN Snapshot 1 April 2021) [Computer Software]. Available online: https://cran.r-project.org (accessed on 1 April 2021).
Recio, Guillermo, Markus Conrad, Laura B. Hansen, and Arthur M. Jacobs. 2014. On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language* 134: 34–43. [CrossRef] [PubMed]
Redondo, Jaime, Isabel Fraga, Isabel Padrón, and Montserrat Comesaña. 2007. The Spanish adaptation of ANEW (Affective Norms for English Words). *Behavior Research Methods* 39: 600–5. [CrossRef] [PubMed]
Rellecke, Julian, Marina Palazova, Werner Sommer, and Annekathrin Schacht. 2011. On the automaticity of emotion processing in words and faces: Event-related brain potentials evidence from a superficial task. *Brain and Cognition* 77: 23–32. [CrossRef]

Rubin, David C., and Michael Friendly. 1986. Predicting which words get recalled: Measures of free recall, availability, goodness, emotionality, and pronounceability for 925 nouns. *Memory and Cognition* 14: 79–94. [CrossRef]
Schacht, Annekathrin, and Werner Sommer. 2009a. Emotions in word and face processing: Early and late cortical responses. *Brain and Cognition* 69: 538–50. [CrossRef]
Schacht, Annekathrin, and Werner Sommer. 2009b. Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, and Behavioral Neuroscience* 9: 28–43. [CrossRef]
Schapkin, Sergei A., Alexei N. Gusev, and Julius Kuhl. 2000. Categorization of unilaterally presented emotional words: An ERP analysis. *Acta Neurobiologiae Experimentalis* 60: 17–28.
Schneider, Walter, Amy Eschman, and Anthony Zuccolotto. 2002. *E-Prime: User's Guide*. Sharpsburg: Psychology Software Incorporated.
Schupp, Harald T., Markus Junghöfer, Almut I. Weike, and Alfons O. Hamm. 2003. Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. *Neuroreport* 14: 1107–10. [CrossRef] [PubMed]
Scott, Graham G., Patrick ODonnell ', Hartmut Leuthold, and Sara C. Sereno. 2009. Early emotion word processing: Evidence from event-related potentials. *Biological Psychology* 80: 95–104. [CrossRef] [PubMed]
Sheikh, Naveed A., and Debra Titone. 2016. The embodiment of emotional words in a second language: An eye movement study. *Cognition and Emotion* 30: 488–500. [CrossRef]
Sutton, Tina M., Jeanette Altarriba, Jennifer L. Gianico, and Dana. M. Basnight-Brown. 2007. The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion* 21: 1077–90. [CrossRef]
The Jamovi Project. 2021. *Jamovi*. (Version 1.8) [Computer Software]. Available online: https://www.jamovi.org (accessed on 1 February 2022).
Thomson, Judith Jarvis. 1985. The Trolley Problem. *The Yale Law Journal* 94: 1395–415. [CrossRef]
Toivo, Wilhelmina, and Christoph Scheepers. 2019. Pupillary responses to affective words in bilinguals' first versus second language. *PLoS ONE* 14: e0210450. [CrossRef] [PubMed]
Van Der Meij, Maartje, Fernando Cuetos, Manuel Carreiras, and Horacio A. Barber. 2011. Electrophysiological correlates of language switching in second language learners. *Psychophysiology* 48: 44–54. [CrossRef]
Van Hell, Janet G., Carla B. Fernandez, Gerrit Jan Kootstra, Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2018. Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism* 8: 134–61. [CrossRef]
Vélez-Uribe, Idaly, and Mónica Rosselli. 2021. Electrophysiological correlates of emotion word processing in Spanish–English bilinguals. *Bilingualism: Language and Cognition* 24: 31–55. [CrossRef]
Voeten, Cesko. 2020. Using 'Buildmer' to Automatically Find and Compare Maximal (Mixed) Models. Available online: https://cran.rproject.org/web/packages/buildmer/vignettes/buildmer.html (accessed on 1 February 2022).
Wu, Yan J., and Guillaume Thierry. 2012. How reading in a second language protects your heart. *The Journal of Neuroscience* 32: 6485–89. Available online: https://www.jneurosci.org/content/32/19/6485 (accessed on 15 January 2022). [CrossRef]
Yao, Zhao, Deshui Yu, Lili Wang, Xiagru Zhu, Jingjing Guo, and Zhenhong Wang. 2016. Effects of valence and arousal on emotional word processing are modulated by concreteness: Behavioral and ERP evidence from a lexical decision task. *International Journal of Psychophysiology* 110: 231–42. [CrossRef]
Zhang, Dandan, Wei-Qi He, Ting Wang, Wenbo Luo, Xiang-Ru Zhu, Ruolei Gu, Hong Li, and Yue-Jia Luo. 2014. Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience* 9: 1897–903. [CrossRef]

languages MDPI

Article

Emotion Processing in a Highly Proficient Multilingual Sub-Saharan African Population: A Quantitative and Qualitative Investigation

Dana M. Basnight-Brown [1], Asiya Ayoob [1] and Jeanette Altarriba [2,*]

[1] School of Humanities and Social Sciences, United States International University Africa, P.O. Box 14634, Nairobi 00800, Kenya
[2] Department of Psychology, University at Albany, State University of New York, Albany, NY 12222, USA
* Correspondence: jaltarriba@albany.edu

Abstract: Research using traditional experimental paradigms (e.g., Priming, Stroop and Simon tasks), narratives and interview type data have revealed that bilingual speakers process and express emotion differently in their two languages. In the current study, both a qualitative and quantitative approach were taken to investigate how individuals who know and regularly use several languages process emotion in each of their languages. In Experiment 1, emotional stimuli in the L2 and L3 was quantitatively investigated using an Affective Simon Task. Participants consisted of Sub-Saharan African multilinguals who had acquired Kiswahili (L2) after their mother tongue (L1), followed by English (L3). The results revealed no difference in the way emotion and emotion-laden words were processed in the two languages (Kiswahili and English). However, significant Affective Simon Effects emerged for positive emotion and emotion-laden words, suggesting that these multilinguals largely process positive emotions in their L2 and L3. In Experiment 2, narrative data generated by multilinguals was used to determine how language selection was influenced by context and type of emotional situation. Themes that emerged within the qualitative analysis revealed that one's L1 was the more emotional language when expressing negative emotions, while the L2 and L3 were reported to be used more frequently when expressing positive emotions, or when discussing more sensitive or embarrassing topics.

Keywords: Bilingualism; Multilingualism; emotion; emotional language

Citation: Basnight-Brown, Dana M., Asiya Ayoob, and Jeanette Altarriba. 2022. Emotion Processing in a Highly Proficient Multilingual Sub-Saharan African Population: A Quantitative and Qualitative Investigation. *Languages* 7: 280. https://doi.org/10.3390/languages7040280

Academic Editor: John W. Schwieter

Received: 7 June 2022
Accepted: 11 October 2022
Published: 2 November 2022

1. Introduction

Over the years, research using traditional cognitive paradigms (e.g., Priming, Stroop, and Simon tasks), narratives, and interview type data has revealed that bilingual speakers process and express emotion differently in their two languages (Altarriba and Basnight-Brown 2010; Dewaele 2008; Harris 2004; Marian and Kaushanskaya 2004; Pavlenko 2005). This finding has implications not only for issues surrounding processing and acquisition in a foreign language, but also for those working in clinical settings, where bilinguals will often switch into their second language if they want to distance themselves emotionally from an event, as the less dominant language has been described as the emotionally distant language in many cultures (Harris 2004; Santiago-Rivera and Altarriba 2002). Early research exploring the intersection of language selection and emotional expression, using open-ended responses, observed that many multilinguals report the emotional weight of the phrase 'I love you' to be strongest in their first language (L1) (Dewaele 2008). In addition, multilinguals also report feeling different when they switch into a second (L2) or third (L3) language, describing that often they feel less logical, less serious, or even less emotional (Dewaele and Nakano 2012). Meanwhile, early research aimed at exploring physiological markers of emotional processing used skin conductance responses (SCRs) to measure the level of arousal from various stimuli (often words and pictures). As one example, Harris et al. (2003) reported that SCRs were particularly strong in the L1 for negative stimuli (e.g., cancer, kill, death), taboo items (e.g., curse words, sexual terms), and reprimands ('Don't do that'; 'Shame on you') (see also Harris 2004, for a review).

1.1. Valence Differences in Emotion Processing in Stroop and Affective Simon Tasks

Valence differences in language processing (i.e., differences between positive and negative stimuli) have been widely reported within the monolingual language processing literature. For example, studies using the Emotion Stroop paradigm often reveal that negative stimuli produce larger Stroop interference effects (i.e., indicating that negative emotional items interfere more with cognitive processing and are activated to a greater degree in early stages of processing) (McKenna and Sharma 1995). Similar findings have been reported with other paradigms, such as rapid serial visual presentation (RSVP) tasks, where it was found that repetition blindness effects were larger for emotional items, particularly that recall was improved for negative items. This was interpreted as occurring because those items capture attention more readily (Knickerbocker and Altarriba 2013).

Another paradigm that has been used to examine a level of emotion processing, at early stages of activation, is the Affective Simon Task (AST) (De Houwer and Eelen 1998). In this original work, participants were presented with nouns and adjectives that had either a positive, negative, or a neutral affective connotation (e.g., baby, murder, happy, cruel). Participants were instructed to respond to the grammatical aspect of the words, with half of the participants responding with 'positive' to nouns and 'negative' to adjectives, while the other half responded 'positive' to adjectives and 'negative' to nouns, irrespective of the valence of the word. The results suggested that even though the valence of the stimulus was not relevant and was to be ignored, participants responded faster to the words which had the same valence as the response.

In 2003, De Houwer adapted the AST, in terms of how participants were to respond. For example, participants were first presented with words in white ink and were instructed to press P on the keyboard for positive words (e.g., kind) and Q for negative words (e.g., hostile). Within the experiment, they were also presented with the same words appearing in blue or green ink colors. For these trials, participants were instructed to respond by pressing P on the keyboard for all words appearing in green, and to respond by pressing Q for all words that appeared in blue. Using this methodology, De Houwer observed that participants responded faster to positively valenced words that appeared in green (i.e., congruency effect). The results from this study demonstrated that key responses were facilitated or inhibited by the valence of the words. The responses given by the participants were based on color discrimination, and yet, valence seemed to play an important role in processing. Interestingly, and of relevance to the discussion on valence differences, the size of the effect was significantly greater for negative stimuli as compared to positive stimuli.

Not surprisingly, the use of these cognitive paradigms has been extended to the study of emotion processing in bilingual populations. The Emotion Stroop task, mentioned earlier, has been used to explore whether emotion words are processed differently in the L1 versus L2. Sutton et al. (2007) investigated Emotional Stroop effects in Spanish-English bilinguals. They examined response latencies to negative and neutral color words that appeared in both Spanish and English. Significant interference effects (i.e., slower response times to emotion words) were observed in both languages. Specifically, the size of the interference effects did not differ across languages, indicating that emotion words are capable of capturing attention in both of a bilinguals' languages (i.e., at least for early bilinguals who were highly proficient in both languages). In a second demonstration of emotional Stroop processes in bilinguals, emotion word processing was tested in Finnish-English bilinguals, who also exhibited high levels of proficiency in their two languages (Eilola et al. 2007). Eilola et al. (2007), like Sutton et al. (2007), observed significant interference effects in both L1 and L2.

Like the Stroop task, the AST has also been used to examined emotion processing in the L1 and L2. Altarriba and Basnight-Brown (2010) used the AST to explore positive and negative emotion word representation in English-speaking monolinguals and Spanish-

English bilinguals. In additional to the bilingual component, another novel aspect of this study was that emotional stimuli were separated based on whether they were emotion or emotion-laden. In the past, items of this nature have typically been combined in various experimental paradigms aimed as examining emotion processing, yet differences in how they are activated do appear to exist. Emotion words are typically those that characterize a particular emotional state (e.g., happy, sad, angry), while emotion-laden words often indirectly activate a positive or negative emotion (e.g., cancer, grave, gold). In both the monolingual and bilingual emotion processing literature, differences have been consistently reported between emotion and emotion-laden words for both behavioral and neurological data (Altarriba and Basnight-Brown 2010; Knickerbocker and Altarriba 2013; Zhang et al. 2017).

For this reason, Altarriba and Basnight-Brown (2010) sought to utilize the AST in a bilingual sample, while also making a distinction between emotion and emotion-laden items in their design. In Experiments 1 and 2, with monolingual participants, significant Affective Simon Effects were observed for positive and negative emotion-laden words (Experiment 1), as well as for negative emotion words (Experiment 2). Positive emotion words did not produce the congruency effects that are characteristic of the AST. According to Altarriba and Basnight-Brown (2010), a possible explanation for the difference in the two experiments could be that emotion-laden words are often objects or situations that are concrete and more imageable than emotion words. Thus, it is possible that the Affective Simon Effect obtained by previous studies (in which emotion and emotion-laden words were mixed) could be due to the greater influence and the presence of emotion-laden words in the stimuli lists. In addition, when the task was applied to the bilingual sample, significant Affective Simon Effects were observed for positive and negative emotion-laden words, in both Spanish and English (Experiment 3). For emotion words, significant effects emerged in English, for positive and negative items, but in Spanish, significant effects were produced for negative items only. In summary, this study revealed that the AST can be effectively used to examine emotion processing in bilinguals, while also demonstrating the importance of distinguishing between emotion and emotion-laden words in a paradigm that is different from that previously used to examine these discrepancies. In the current study, this same paradigm will be extended to the processing of emotion and emotion-laden stimuli in one's L2 and L3, in an effort to expand this original work by Altarriba and Basnight-Brown (2010) to that of a multilingual population.

1.2. Translation Differences in Multilingual Populations

In addition to processing differences that exist for emotional stimuli in the L1 and L2, bilingual language representation is often characterized by issues related to translation, or more specifically, how well an item translates across languages. It is well established, through both linguistic and psychological research, that words in one language do not always have a direct translation across languages. This linguistic challenge is often presented as one of the reasons why bilinguals code-switch, to better convey the full meaning of what they desire to express (Heredia and Altarriba 2001). When it comes to emotion processing, anecdotal evidence often suggests that emotion words are some of the most difficult to translate for those who know and use more than one language. As Pavlenko (2005) points out, much of this is likely due to the fact that languages differ in the size of their emotional lexicon. For example, she describes that the English language has an estimated 2000 emotion words, with Dutch estimated to contain roughly 1500, meanwhile, the Indonesian language only has a couple hundred words to describe emotion. This is important, as naturally it will affect how one's lexicon develops, the manner in which they express emotion, as well as the language they may choose to use when discussing an emotional event.

For many years, despite acknowledging that translation ambiguities exist, few studies using traditional cognitive paradigms systematically examined whether this variable was influencing outcomes and how it may be considered in models of bilingual memory. In 2002, Tokowicz, Kroll, de Groot, and van Hell (Tokowicz et al. 2002) conducted an analysis of stimuli use in previously published bilingual studies and observed that roughly a quarter of items used in these well-known studies had multiple translations. Since then, translation ambiguity has been explored in various paradigms and been manipulated in studies focused on language acquisition (Basnight-Brown and Altarriba 2016; Bracken et al. 2016; Degani et al. 2016; Tokowicz and Kroll 2007; Tseng et al. 2014).

Despite bilinguals reporting specific challenges in translating emotional stimuli, quantitative evidence describing the extent of this phenomenon was limited. In order to empirically determine the degree of translation ambiguity for emotion words, Basnight-Brown and Altarriba (2014) examined the number of translations generated for concrete, abstract, and emotion words in two bilingual populations (Spanish-English and Chinese-English bilinguals). It was consistently observed, across both language directions (L1–L2 and L2–L1), that emotion words produced the largest number of translations across languages, for both bilingual groups. Interestingly, Chinese emotion words specifically elicited a very high number of translations, indicating that a direct translation for many emotions is even more challenging in that language. This was important, as it provided translation norms for words that describe a specific emotion, for both bilinguals that share an alphabetical script across their languages, as well as for those who do not.

1.3. Current Study

In recent years, the number of multilinguals has grown globally, due to changes in political, educational, and environmental situations. However, to date, there is a limited understanding in how multilinguals express and process emotion in their languages, given that most studies have focused on bilingual samples. The focus of the current study is to explore emotion processing in a multilingual population from Sub-Saharan Africa, an area known for its linguistic depth and cultural diversity. In addition, an even more limited understanding of the cognitive processes surrounding multilinguals in Africa exists, given that so few studies in this domain have been conducted on the continent. For example, a recent systematic review of papers published in top psychology journals discovered that less than 1% of first authors and participant samples come from Africa, despite the continent comprising almost 20% of the global population (Thalmayer et al. 2021). For that reason, it is our goal that the current study can begin to shed some light on the processing of language and emotion in this linguistically rich population, and that our understanding of language processes will be enhanced through the study of populations that are under-represented in the research landscape.

The focus of the current study was to examine how a highly proficient multilingual population processes emotion in their L2 and L3, using a sequential mixed methods approach (Quantitative → Qualitative). The quantitative portion (Experiment 1) examined the processing of emotion in the L2 and L3, using the Affective Simon task, analogous to the Altarriba and Basnight-Brown (2010) study. As mentioned, the East African region is characterized by tremendous linguistic diversity and depth. In Kenya specifically, many individuals refer to their ethnic language (referred to by many as the mother tongue) as their L1. Given that there are more than 40 ethnic groups represented in Kenya, many different mother tongues exist (e.g., Kikuyu, Kamba, and Luo are some of the most widely spoken for the sample in this study). In addition, Kiswahili serves as the National Language of the country, while English serves as the Official Language of the country. As a result, the average Kenyan knows and uses a minimum of three languages daily, with many knowing several more.

Therefore, the main aim of Experiment 1 was to determine whether multilingual individuals process emotions in their L2 and L3 differently, or whether all learned languages apart from the mother tongue have a similar representation. In addition, a secondary goal was to determine whether emotion and emotion-laden words are processed differently in the L2 and L3, given that they have been found to be different in both monolingual and bilingual populations. Since emotion words often do not translate directly across languages and may result in multiple translations, a norming study was piloted first, in order to determine which Kiswahili translations were the best representation, for this specific population. For the quantitative measures, we chose to focus on L2 and L3 processing, primarily because this task involved a timed reading task. Prior experience with this population revealed that they are highly skilled readers in their L2 and L3, having conducted much of their education in those languages. The L1 varies tremendously across the population, due to the diversity in ethnic groups that exist within Kenya. Furthermore, self-reports from this multilingual population indicate that many do not read in their L1, as that is a language used primarily for spoken communication with family members. As a result, it was important to focus on those languages (L2 & L3) in which the participants read, given the nature of the experimental task. For the qualitative portion (Experiment 2), multilingual participants generated a written narrative essay, a reflection to a study they learned about in a group setting. This allowed them to freely discuss how they use each of their known languages to express emotion in various situations. Key themes from these responses were identified as dominant patterns in emotional selection for the multilinguals and are discussed in terms of how they supplement the quantitative findings from the Affective Simon Task.

2. Methods

2.1. Experiment 1 (Part A—Norming Study)

2.1.1. Participants

Fifty participants from the United States International University—Africa (USIU—Africa) completed the norming study. These students were from Kenya, Tanzania, and Rwanda, and all participation was voluntary. The study was approved by the USIU-Africa ethical review board.

All participants were multilingual and knew English and Kiswahili as two of their languages. The L1 (mother tongue) varied across participants, which is common in the East African region, given that there are dozens of mother tongue languages represented. Self-reported data from the Language History Questionnaire (LHQ, adapted from Marian et al. 2007) revealed that the average AoA of Kiswahili and English was 3.6 and 4.5 years old, respectively (see Table 1). Despite English and Kiswahili both being acquired at young ages (and in close succession), English skills were rated higher and English was reported to be used for a larger percentage during each day. This is likely because the participants were attending university at the time of investigation, and university instruction across Kenya is conducted in English. Therefore, verbal communication, as well as reading and writing in English, likely composed a large portion of their daily language behavior. The multilinguals also reported using “other” languages for a smaller percentage of the day, referring to the use of their mother tongue, and/or to a fourth language for some (often those are reported to be languages studied at university).

2.1.2. Stimuli

All participants were presented with a list of 20 emotion words and 20 emotion-laden words, in English, taken from Altarriba and Basnight-Brown (2010). For each category of emotion words, 10 were positively valenced and 10 were negatively valenced. All items were matched on word frequency and length, and all items were high in arousal, with positive and negative items being matched on this characteristic (according to Bradley and Lang 1999). As expected, valence significantly differed across the categories, with positive words being high in valence and negative words being low in valence (see Altarriba and Basnight-Brown 2010).

Table 1. Language History Questionnaire Data for Norming and Affective Simon task: Means and standard deviation (in parentheses).

	Norming	**Affective Simon**
Kiswahili (L2)		
Age of acquisition (years)		
Speaking	3.6 (2.2)	4.7 (2.8)
Reading	5.6 (2.4)	6.4 (2.7)
Self ratings of language skills		
Spoken comprehension	5.5 (1.4)	6.0 (1.0)
Speaking	5.2 (1.5)	5.6 (1.2)
Reading	5.2 (1.6)	5.7 (1.3)
Writing	4.9 (1.7)	5.3 (1.5)
English (L3)		
Age of acquisition (years)		
Speaking	4.5 (2.8)	6.3 (3.2)
Reading	5.3 (2.5)	6.8 (3.1)
Self rating of language skills		
Spoken comprehension	6.3 (0.97)	6.1 (0.9)
Speaking	6.0 (1.1)	5.8 (0.7)
Reading	6.3 (1.0)	6.2 (0.8)
Writing	6.1 (1.1)	5.9 (0.9)
Estimated daily language usage		
Kiswahili	34%	38%
English	54%	45%
Other	12%	17%

2.1.3. Procedure

Participants were provided with the list of English words in written format and were instructed to translate the words into Kiswahili without looking at a dictionary and without referring to anyone. They were told that they could leave a blank space for any words they were unable to translate. After they were done with the translation, they were asked to fill in the LHQ.

2.2. Experiment 1 (Part B—Affective Simon Task)

2.2.1. Participants

Forty-two students from USIU-Africa (36 females/6 males; mean age = 27.01), who did not participate in the norming task, completed the Affective Simon task (AST). For this task, one requirement was that the multilinguals had been exposed to Kiswahili (L2) before they were exposed to English (L3). LHQ data revealed that the participants had an average AoA of Kiswahili and English of 4.7 and 6.3 years old, respectively (see Table 1). All participants had a native language (L1) that was different from both Kiswahili and English. Native languages consisted of Arabic, Amharic, German, Kikamba, Kikuyu, Luhya, Luo, Munyo Yaya, Kimeru, Somali, and Pokot.

All participants had normal or normal-to-corrected vision and were tested for color blindness. Ishihara's (1939) test of color blindness was used to confirm whether the participants were able to distinguish between the colors used in the AST. All participants (except for 1), were able to correctly identify the plates on Ishihara's color blindness test. However, the participant who did not pass the test for all colors was able to accurately distinguish between the blue and green colors of the Simon Task and was therefore still included.

2.2.2. Stimuli

As mentioned earlier, the English stimuli were originally derived from Altarriba and Basnight-Brown's study (2010), then normed in this study. All Kiswahili words were derived from the norming study in Experiment 1a. This resulted in a set of 16 emotion-laden words and 14 emotion words, for use in the AST (see Appendix A, Tables A1 and A2 for full list). For the English words, all items were matched on length, frequency, and rated high in arousal. Valence significantly differed across items, with positive words having higher valence than negative words (Bradley and Lang 1999). The average length of words in Kiswahili and English were matched for each condition (see Table 2).

Table 2. Attributes for the word stimuli used in the Affective Simon Task.

	EMOTION-LADEN WORDS		**EMOTIONAL WORDS**	
	Positive	**Negative**	**Positive**	**Negative**
English	*Friend*	*Death*	*Happy*	*Depressed*
Frequency	65.88	51.25	34.71	43.57
Arousal	5.50	5.91	5.75	6.15
Valence	7.34	2.49	7.61	2.63
Length	5.50	5.13	7.43	6.14
Kiswahili	*Rafiki*	*Kifo*	*Furaha*	*Huzuni*
Length	5.75	5.13	6.43	6.14

2.2.3. Procedure

Participants were first told to read and sign the consent. After making sure their phone devices were silent or switched off, Ishihara's (1939) color blindness test was given. The experiment started with written instructions on the screen. Participants were instructed to read carefully and then explain what they understood. If the participants misunderstood any of the instructions, they were corrected. All participants were reminded to be as accurate and as fast as possible when responding to the words.

The experiment was carried out on an HP W1972a desktop computer with an 18.5 inch LCD monitor. The computer software used was E-prime, version 2.0 (Schneider et al. 2002). All the words were displayed on the center of the screen on a black background. All words were presented in Courier New font with a font size of 18 and were in bold. Before the experimental trials, participants were given 20 practice trials. The first block (10 trials) had only white words, and the second practice block (10 trials) had blue and green words. Before each of the practice and test blocks, specific sets of instructions were given. For the white words, they were instructed to respond to the meaning of the word (e.g., positive/negative), and for the colored words, they were in-structed to respond to the color (e.g., blue/green). Participants responded by pressing 'P' and 'Q' keys with their index fingers. They pressed 'P' for positive white words and for blue words, and they pressed 'Q' for negative white words and for green words.

Each word appeared once, in random order, in three different colors; white, blue and green. The blue and green shades of colors were the same as that used by De Houwer (2003). Thus, every test block had 90 trials (30 emotional stimuli in each of the 3 colors). The same procedure was repeated for the second language, counterbalanced across participants, such that some participants received the Kiswahili block first and others received the English block first. The total number of test trials that each participant responded to was 180. All the words appeared on the center of the screen for 3000 ms, which was preceded by a fixation (+) that appeared on the screen for 500 ms. The inter-trial interval was for 1500 ms.

If the participant made an error, a red '+' sign appeared on the center of the screen. Once the participant completed the AST, they were asked to fill out the LHQ.

2.3. Experiment 2 (Qualitative Component)

In the qualitative portion of this mixed method study, narrative data from multilingual participants, who represented various Sub-Saharan African countries, was ex-amined in order to determine how the language selected in a given situation determined the level of emotional expression. Using this methodology, we sought to allow the multilinguals to freely discuss, through writing, how they select and use their languages in an emotional context.

2.3.1. Participants

Thirty-two participants from USIU-Africa participated in Experiment 2. They were from the same population as those who participated in Experiment 1. Like that sample, the L1 varied across participants, with Swahili and English, serving as their L2 and L3. The AoA for Swahili and English was 5.2 and 4.5 years old, respectively.

2.3.2. Procedure/Task

In a group setting, participants were exposed to the ideas presented by Chen et al. (2012, see review), which emphasize that how parents express emotion with their children, and the language in which they choose to do so, can affect emotional development and communication patterns. The participants were then asked, individually, to write a brief essay, reflecting on their own language experiences, particularly in terms of how and when they use their multiple languages. The session lasted approximately 50 min.

3. Results

3.1. Experiment 1 (Part A—Norming Study)

The translations given by the participants were evaluated and the translation that appeared most frequently was picked for use in the Affective Simon Task. Overall, 4 emotion-laden words (2 positive and 2 negative), and 6 emotion words (3 positive and 3 negative) were removed from the original stimuli list. This analysis revealed that some items had mixed valences (e.g., puppy was characterized as being "positive" in previous studies, but in Kiswahili, 'mbwa' was determined to have a "negative" connotation for many in the culture studied). In addition, due to the nature of Kiswahili having fewer emotion words as compared to English, several of the English words resulted in the same Kiswahili translation. In these cases, participants reported the items having the same meaning as another word in the document (e.g., anxious and nervous both translated into the same word in Kiswahili [wasiwasi]; frustration did not appear to have any direct translation in Kiswahili, producing 18 different translations in total). This was expected, and therefore, was one of the reasons why the norming study was important to pilot prior to conducting the Affective Simon Task. Since duplicate Swahili words could not be used in the Affective Simon Task, this resulted in a slightly smaller subset of stimuli for use in that task, as compared to that used in Altarriba and Basnight-Brown (2010).

3.2. Experiment 1 (Part B—Affective Simon Task)

Analyses of variance (ANOVAs) were conducted using a 2 (Language: Kiswahili and English) $\times$ 2 (Word Type: emotion and emotion-laden) $\times$ 2 (Valence: positive and negative) $\times$ 2 (Congruency: congruent and incongruent) model for all four variables. Data were trimmed for each participant such that any value that went above or below 2.5 standard deviations from the mean (for that individual participant) was replaced by 2.5 $\pm$ standard deviations from the mean. Overall, accuracy rates were very high, but a threshold was set where any participant with an accuracy rate below 70% was removed from the analyses (this applied to one participant only).

The results revealed a main effect for congruency, F (1, 40) = 18.68, $p < 0.001$, showing that participants were significantly faster to respond to congruent trials as compared to incongruent trials. The analysis also showed a marginal main effect for word type, F (1, 40) = 2.95, $p = 0.09$, indicating that participants were slower to respond to emotion words (851 ms) as compared to emotion-laden words (841 ms). Planned comparisons revealed that positive emotion and emotion-laden words, in English, produced significant Affective Simon Effects, t (40) = 5.31, $p < 0.001$, and t (40) = 2.19, $p < 0.05$. For the negative emotion words in English, there was no significant effect, t (40) = 1.54, $p = 0.13$. Similar results were found for Kiswahili items, though the effect was more robust for positive emotion-laden words, t (40) = 3.97, $p < 0.001$ (positive emotional words, t (40) = 1.65, $p = 0.106$). None of the negative words showed a significant Affective Simon effect in Kiswahili (see Figures 1 and 2).

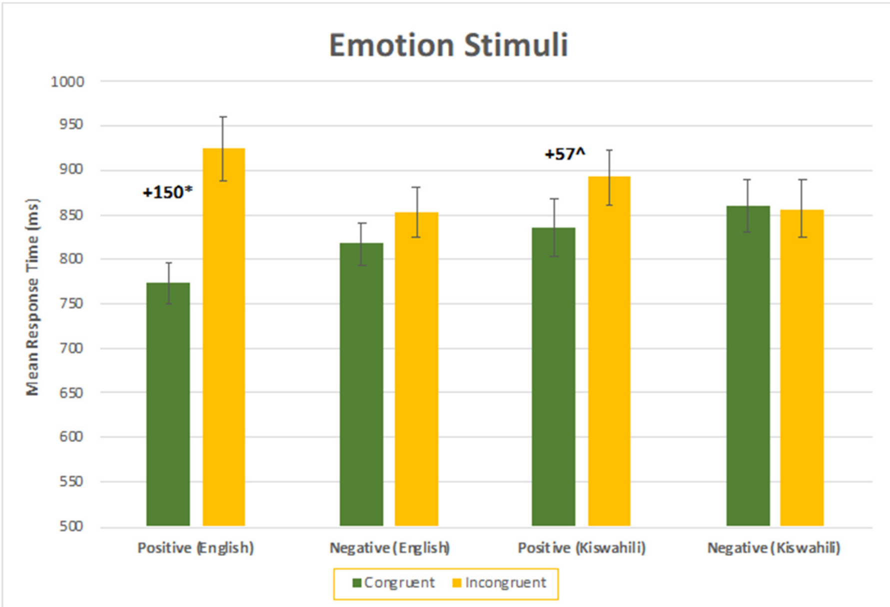

Figure 1. Mean Response times for emotion words. * Indicates $p < 0.05$, ^ indicates $p = 0.10$.

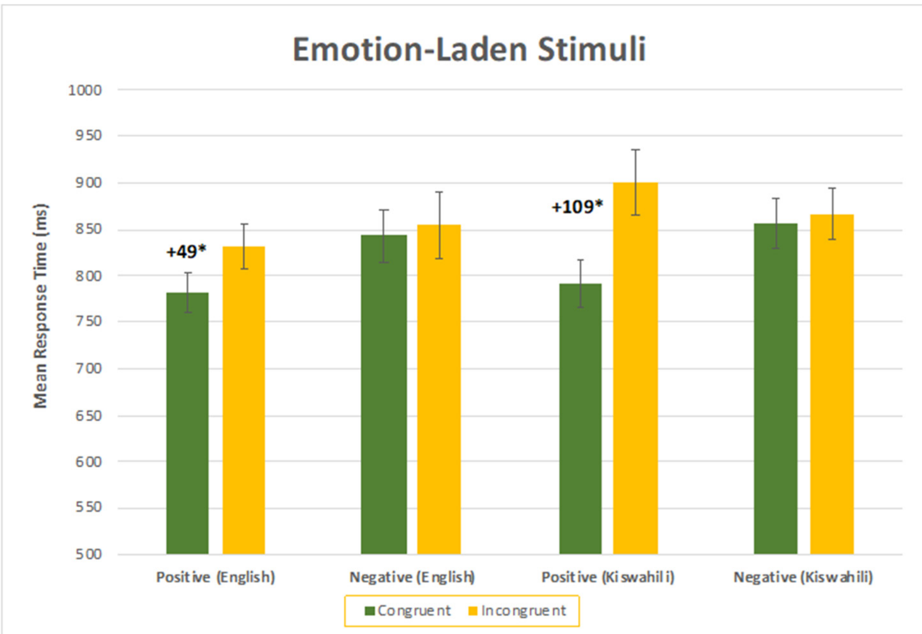

Figure 2. Mean response times for emotion-laden words. * Indicates $p < 0.05$.

In addition, the analysis further revealed a significant interaction between word type and valence, $F (1, 40) = 4.72$, $p < 0.05$, as well as between congruency and valence, $F (1, 40) = 9.99$, $p = 0.002$. The interaction between valence and congruency indicates that significant Affective Simon effects are largest for positively valenced emotional stimuli. The three-way interaction between language, word type, and congruency was also significant, $F (1, 40) = 11.62$, $p = 0.003$, suggesting that while both types of positive emotional stimuli produced significant effects, the effects were not of the same magnitude across languages (larger effects for positive emotion words in English, but in Kiswahili, for positive emotion-laden items). Overall, the results from the quantitative portion of this study revealed significant effects only for positive emotion items, suggesting that stimuli which elicit a negative emotion are not as strongly encoded in one's L2 and L3, for this population at least. As discussed earlier, previously published work suggests that negative emotion words are often more strongly encoded in one's L1, as compared to their L2. It is possible that this is the case for these multilinguals as well, although we acknowledge that L1 Affective Simon processing was not examined due to the limitations mentioned, and the desire to focus on L2 and L3 processing. Despite that, the current findings do show that negative emotion is not captured as strongly in the L2 and L3, as compared to positive emotional items. As a result, the aim of Experiment 2 was to examine how multilinguals from this population observe and use emotional language in the L1, L2, and L3.

3.3. Experiment 2 (Qualitative Component)

The narrative responses were then anonymized, compiled, and analyzed manually using the six-stage process developed by Braun and Clarke (2006). Three main outcome themes emerged from the analysis.

3.3.1. Difficulty Finding a Direct Translation

Participants frequently described that it was often difficult to find a direct translation for an emotional expression in some of their languages, making it difficult for them to accurately express the emotion they wanted in that language. For example, one individual stated:

> "some words when translated don't carry the same weight as before, losing the whole essence of the word." (PN03)

This challenge was also expressed in the following ways:

> "It may be difficult to find words, which specify the actual feeling term which describes anger whereas in English there are different words for describing levels of anger ranging from irritation, annoyance, anger, rage to fury, therefore it is important to name the feelings correctly (e.g., 'I'm annoyed' is different from 'I'm angry'), however, we use one word to express all these different levels of anger." (PN06)

> "It feels like some words don't exist in Kiswahili, so I say them in Kikuyu because then the sentence I'm trying to construct will sound more meaningful, or the story I'm sharing will sound much funnier, sad . . . ." (PN28)

Meanwhile, others touched on cultural factors that may impact translation difficulties, primarily, when semantics are lost in translation due to differences in what is acceptable within a culture.

> "In my tribe it is not okay to talk about your emotions to just anybody and therefore there aren't many words to explain a certain emotion or feeling without sugar coating if it was negative and downplaying it if it was positive; the same goes for Swahili." (PN07)

> "Meaning may differ or even completely get lost in the translation. There is a lot of emphasis on the western traits with this approach. Some of the traits used to describe personality are relevant only to the western culture." (PN19)

3.3.2. L1 Was the Language Selected for Use When Expressing Negative Emotions

A second major theme that emerged, was that one's L1 was often the dominant language used to express negative emotions. Many participants reflected on situations and examples where their parents or themselves were prone to switch into their L1 when experiencing anger or grief. Statements detailing this type of language selection included:

> "My mother always switches from English to Fulani when she is upset or when she is trying to threaten us because there are some Fulani words that when translated to English, do not sound as scary as they should be." (PN25)

> "For my mother, when she wants to express her negative emotion, she will use her L1, which is Kikuyu. Emotions such as anger and frustration are very easy and acceptable to express and discuss in our culture. As for the "difficult" emotions such as affection and love, she will use her L2 languages, Swahili or English, to express them." (PN04)

> [spoken by a participant describing when she learned about the death of a family member] "When I arrived at the hospital and got the tragic news, I switched from English to Kamba, to the amazement of my children, who though this being a somber time found it funny that I was speaking Kamba, something I rarely do . . . it felt more real and emotional to speak Kamba." (PN03)

> "This is also the case with my parents who speak Kikuyu (which is their mother tongue), Kiswahili, and English fluently. Whenever they want to express their praise, appreciation and love for me they always use English . . . when expressing anger however, especially with my parents they stick to using Kikuyu because they feel that expressing it in English will not give it the same weight it would have in Kikuyu." (PN13)

3.3.3. L2 or L3 Was Used More Often When Expressing Positive Emotions or When Discussing More Sensitive or Embarrassing Topics

Finally, another important theme that emerged from the data emphasized that multilinguals from this population often prefer to use their L2 or L3 when expressing positive emotions or when discussing topics that one might consider to be embarrassing or uncomfortable. Many of the examples given, pertained to instances where a family member desired to express their love, but felt it would not be appropriate in their mother tongue language. Statements included:

> "As a Kenyan who is a Kikuyu and I know both Kiswahili and English; when expressing affection to my mother by saying I love you, I don't say it in Kikuyu I say it in English my second language since its more expressive and meaningful." (PN14)

> "As for the 'difficult' emotions, such as affection and love, she [my mother] will use Swahili (L2) and English (L3)" (PN08)

> "I come from a strong African background where emphasis is placed on how one should behave and in society and what should be said in a particular situation. Expressions or emotion especially emotions of love are very rare especially from the parents to their children. However expressions of emotions like anger are very well expressed by the parents. Times however are changing and gone are the days where parents would never show affection or tell their children that they love them. The use of the English language has made all this possible, African parents can now express their love to their children by saying "I love you." (PN09)

> "We mostly speak Kirundi at home, but it is very common for my mother to switch to French (L2) for endearing expressions such as "ma cherie" which can be translate to "my darling" or "sweetheart"; this is because our Burundian culture does not give a lot of room for such expressions" (PN16)

> "I realise in my family, my mother has never told us that she loves us in Kikuyu. I have never heard any members of my extended family say it either. Any time she has told us that she does, it has been in English . . . my mother will use Kikuyu when she is expressing shock and anger a lot of the time. She would shout at us in Kikuyu when we were younger." (PN22)

Finally, in line with this theme, several participants also reported that discussing more taboo or embarrassing topics was easier to do when one was using their L2 or L3. This is illustrated below:

> "I remember my mother discussing sex with me, in English (L3), yet I was proficient in Luo (L1). She would first start speaking in Luo (L1) and noted that it was embarrassing for her speaking about the topic and quickly switched to English (L3). This was also the case when we were discussing HIV and the use to condoms she would speak in English." (PN08)

Overall, these three themes emerged as the most robust within the qualitative analyses, providing some supplemental insight into how emotion is conceptualized and expressed in this multilingual population. One important thing to note, due to the open-ended nature of the generated responses, was that many of the multilinguals provided examples of emotional conversations and comments made by their parents. Therefore, some of the interactions are based on observations of how emotion was used in their community and familial systems, perhaps not attributed to the participant directly. This is insightful and suggests that future research on this topic could implement experimental designs and questions that tap into the participant's personal language use.

4. Discussion

The current work provides a novel contribution to the study of emotion, by ex-amining language usage in a population that is not just bilingual, but in those who know and use multiple languages on a daily basis. Furthermore, it explores the role of multi-lingualism and emotional expression in those on the African continent, an area that is often underrepresented in much of the psychological literature.

In summary, results from the quantitative data revealed significant effects for positive emotion items only, suggesting that stimuli which elicit a negative emotion are not as strongly encoded in one's L2 and L3, for this population at least. This supports previously published work that indicates that negative stimuli (e.g., negative emotion words, taboo words, reprimands) are more strongly encoded in one's L1. Although the L1 was not examined in this study, due to the lack of consistency in participants having a shared L1, the absence of these effects in L2 and L3 processing are indicative of stronger negative emotion representation in the L1 (especially when the qualitative data are considered). Specifically, results from the qualitative data revealed that the language selected for emotional expression was highly dependent upon the topic being discussed. For example, the majority of multilinguals reported using their L1 when expressing negative emotions (e.g., predominantly frustration or anger), but then switching to their L2 or L3 when discussing love or other positive emotions. This finding supports the Affective Simon Task results where significant effects were observed in the L2 and L3 for positive stimuli only.

As noted earlier, Dewaele (2008) originally reported that the phrase 'I love you' was perceived with more emotional weight (felt strongest in) the bilingual's L1 as compared to their L2. Interestingly, with this group of multilinguals, many reported (themselves or family members) choosing to express that level of emotion in the L2 or L3. This is likely influenced by cultural factors, and not solely driven by language per se, as many discussed that saying I love you in their mother tongue would not be as socially acceptable. In a similar vein, the data also revealed that some individuals reported switching to a less dominant language (often the L2 or L3) when discussing highly emotional or embarrassing topics, as a way of distancing themselves emotionally from the situation. For example, one participant noted this experience particularly well in reflecting on how their mother switched to her L3 when discussing the topic of sex. This supports some of the research within the clinical domain, where it has been reported that bilinguals will switch to a less dominant language as way of emotionally distancing themselves from an event (Ladegaard 2018; Santiago-Rivera and Altarriba 2002). As noted earlier, the open-ended nature of the generated responses resulted in many multilinguals providing examples of emotional expression and interaction involving their parents. As a result, it is important to note that the L2 and L3 (often Swahili and English) are less dominant languages for the parents, according to the responses from the multilinguals. However, it is important to note that for the participants, Swahili and English were likely languages that were more dominant for them, due to being educated primarily in Swahili and English, as well as in attending an English-speaking university. This "dominance shift", often observed in many language users, is likely caused by the differing educational backgrounds between parents and children, differences in the use of the mother tongue across generations, and the age of the multilinguals.

Next, focusing on the quantitative data in more depth, allows us to examine two interesting questions (1) does the processing of emotion in the L2 and L3 differ for this multilingual population, and (2) is there a difference in emotion and emotion-laden processing across languages? For the first question, the results reveal that out of the 8 word categories (i.e., postive and negative emotion words in L3, positive and negative emotion words in L2, positive and negative emotion-laden words in L3, positive and negative emotion-laden words in L2), only 3 of them (positive emotion words in L3, positive emotion-laden words in L2 and L3) produced a significant Affective Simon Effect. This suggests that the processing and representation of emotion in the L3 is not identical to that in the L2. According to the results, there seems to be a pattern, in that, only positively valenced words appeared to produce the Affective Simon Effects in both L2 and L3. This finding could be attributed to proficiency differences across the L2 and L3. As Table 1 reveals, the multilinguals reported being more proficient in English (L3), as compared to Kiswahili (L2). These proficiency differences may be responsible for the larger effects observed for English emotion words. At the same time, the effects for positive emotion-laden words were larger in Kiswahili, a class of words characterized more often by nouns. The norming process conducted on emotion stimuli revealed that Kiswahili has fewer emotion words in the lexicon as compared to English, suggesting that emotion words are used more frequently in English and therefore, produced stronger effects. In turn, emotion-laden words had more consistency, due to being simple nouns, indicating that activation for those items could be heightened in Swahili if they are used more often than emotion items in that language.

In Altarriba and Basnight-Brown's (2010) study, the Spanish bilinguals produced the Affective Simon Effects in both positive and negative emotion words, as well as in positive and negative emotion-laden words in the L2. However, this study replicated the Altarriba and Basnight-Brown (2010) findings only for positively valenced items. However, when one takes the self-reported reflections of the qualitative responses into consideration, it becomes evident that due to linguistic and cultural factors, there is a strong tendency for positive emotions to be more robust in the L2 and L3, as compared to negative emotions, which were consistently reported to be elicited to a stronger degree in the L1. For the second question presented, significant Affective Simon Effects emerged for positive emotion-laden words in both languages, while for emotional words, the effect was only observed for positive emotion words in English. However, overall, the effects for emotional words were larger as compared to emotion-laden, suggesting that processing differences do exist. Overall, the findings from the current study reveal that English emotional stimuli are processed differently depending on whether they are presented in the L2 (in Altarriba and Basnight-Brown 2010) or in the L3 (current study). More importantly, the findings also suggest that positive emotions are processed strongly in both the L2 and L3, for a highly proficient multilingual population. This is consistent with studies that suggest that people prefer to express negative emotions in their first language. Thus, it is possible that the reason behind the null effects observed for negative emotion and emotion-laden words in the L2 and L3 is because these emotions are largely processed in one's L1, which was not the focus of the Affective Simon Task in the current experiment.

Applied Research Domains

The importance of knowing which languages more richly code emotional language (L1, L2, L3, etc.) cannot be overemphasized. Published works have indicated that the strength and intensity of emotional language varies depending upon the context in which that language is acquired (Altarriba 2003, 2008). Native or first-learned languages often reserve the most richly deep and diverse emotional concepts and corresponding labels, quite possibly because that is the language in which emotional concepts were first tied to actual language labels, and, where physiological reactions were also associated with and derived from the use of those words (e.g., reprimands, taboo words, use of the word "no", etc.).

Knowing that L1 might capture the broadest range of emotional language in terms of valence and intensity, researchers have written about how the switching and mixing between languages might be used strategically in order to elicit fuller memories and reports of past experiences in situations such as those involving counseling and therapy, responding as eyewitnesses, or merely telling autobiographical stories in testimonials and other like situations (see, e.g., Marmolejo et al. 2009; Santiago-Rivera et al. 2009). For example, Santiago-Rivera and colleagues note that clients who were in therapy settings often switched from their native language into a second language when discussing particularly negative, distressing events in their lives. They aimed towards distancing themselves from those negative emotions, and instead, they preferred to use a second or even third language so that they could more easily discuss events without the distress caused by the use of negative langauges. Over time, those who are administering treatment can work through ways of having the client switch, when ready to do so, to work through those more emotional memories, when they deem the time is right, within a therapeutic session. By knowing that emotional language is often processed differently across languages in bilingual and multilingual speakers, strategies can be developed that can help to elicit information in a more appropriate manner. These kinds of switching techniques can be developed and formalized in any situation in which a bilingual or multilingual speaker is being asked to recount events from memory whether recent, or from the distant past.

5. Conclusions

From this study it could be concluded that the way Spanish bilinguals process their second language may be different from the way Kiswahili multilinguals process their second language. Additionally, it may be that Kenyans reserve their expression of negative emotions for their first language only, a language that was not investigated in the current study. This difference could be attributed to culture and lifestyle. Furthermore, the language people use can reflect the culture they came from since language and culture are so closely intertwined. Thus, when English was introduced to the Kenyan population during British colonialism, it did not just bring a language, it brought a culture. It brought a lifestyle where emotions may be expressed in a different way than the Kenyan cultures do. It can be suggested that the English language may have altered the way Kenyans express their emotions particularly when using language. It may have enhanced the use of positive emotional language.

In summary, the current study examined emotional processing and expression in East African multilinguals, with a goal of conducting exploratory research on a population that is not well represented in the emotion processing literature. Due to the exploratory nature of the design, there are several limitations that exist. As noted earlier, we were not able to examine L1 emotion processing in the quantitative portion of the study, due to the varied L1s that exist within an urban East African sample. Furthermore, as mentioned, a homogenous L1 Kenyan sample would not completely address this issue, given that many Kenyans from this age group report not reading in their mother tongue (preferring the L2 and L3 for reading, given that the education system within the country focuses on those languages). Despite those challenges, it is possible that future research endeavors focused on quantitative data could make use of experimental designs that do not rely as heavily on literacy skill of the mother tongue and/or which focus on a group of multilinguals who may be from a different age group and who read more regularly in their L1. Second, the qualitative design emphasized open-ended responses, allowing participants to freely discuss and generate responses based on their experiences and linguistic behaviors and observations surrounding emotional expression. In the future, qualitative data for this population could be obtained using direct questions, which are narrower in nature, allowing for more pointed responses (such as that implemented by Dewaele and colleagues). Finally, the current study examined emotion use and expression in a more familial context, yet outcomes may be different when contexts change (e.g., emotion selection when communicating with a friend, stranger, work colleague, etc.). This is another important domain for future study of emotional expression within multilingual communities. It is our goal that the exploratory findings of the current study can serve as a foundation and motivation for others interested in exploring emotion in less represented, yet linguistically rich, multilingual populations.

Author Contributions: Conceptualization, D.M.B.-B., A.A., and J.A.; methodology, D.M.B.-B., A.A., and J.A.; software, D.M.B.-B. and A.A.; validation, D.M.B.-B. and A.A; formal analysis, D.M.B.-B. and A.A.; investigation, D.M.B.-B. and A.A.; resources, D.M.B.-B. and J.A.; data curation, D.M.B.-B. and A.A.; writing—original draft preparation, D.M.B.-B., A.A., and J.A.; writing—review and editing D.M.B.-B., A.A., and J.A.; visualization, A.A.; supervision, D.M.B.-B.; project administration, D.M.B.-B., and A.A.; funding acquisition, J.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Approved by the USIU-Africa IRB.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors delcare no conflicts of interest.

Appendix A

Table A1. Emotion words used in AST.

KISWAHILI	ENGLISH
Uoga	Fear
Hasira	Angry
Huzuni	Depressed
Wasiwasi	Anxious
Umia	Hurt
Upweke	Lonely
Kusumbuka	Troubled
Furaha	Happy
Mkweli	Honest
Tumaini	Hopeful
Tajiri	Wealthy
Mshindi	Triumphant
Nguvu	Powerful
Shukrani	Thankful

Table A2. Emotion-laden words used in AST.

KISWAHILI	ENGLISH
Kifo	Death
Kosa	Fault
Mafuriko	Flood
Mazishi	Funeral
Mwizi	Thief
Papa	Shark
Nyoka	Snake
Baka	Rape
Ndoto	Dream
Maua	Flower
Rafiki	Friend
Zawadi	Gift
Mbinguni	Heaven
Afya	Health
Tamani	Wish
Miujiza	Miracle

References

Altarriba, Jeanette. 2003. Does cariño equal "liking"? A theoretical approach to conceptual nonequivalence between languages. *International Journal of Bilingualism* 7: 305–22. [CrossRef]

Altarriba, Jeanette. 2008. Expressions of emotion as mediated by context. *Bilingualism: Language and Cognition* 11: 165–67. [CrossRef]

Altarriba, Jeanette, and Dana M. Basnight-Brown. 2010. The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism* 15: 2310–328. [CrossRef]

Basnight-Brown, Dana M., and Jeanette Altarriba. 2014. Number of translation differences in Spanish and Chinese bilinguals: The difficulty in finding a direct translation for emotion words. In *Psychology Serving Humanity: Vol. II*. Edited by Saths Cooper and Kopano Ratele. New York: Taylor & Francis, pp. 240–51.

Basnight-Brown, Dana M., and Jeanette Altarriba. 2016. Multiple translations in bilingual memory: Processing differences across con-crete, abstract, and emotion words. *Journal of Psycholinguistic Research* 45: 1219–45. [CrossRef]

Bracken, Jennifer, Tamar Degani, Chelsea Eddington, and Natasha Tokowicz. 2016. Translation semantic variability: How semantic relatedness affects learning of translation-ambiguous words. *Bilingualism: Language and Cognition* 20: 783–94. [CrossRef]

Bradley, Margaret M., and Peter J. Lang. 1999. *Affective Norms for English Words (ANEW)*. Gainesville: The NIMH Center for the Study of Emotion and Attention, University of Florida.

Braun, Virginia, and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3: 77–101. [CrossRef]

Chen, Stephen H., Morgan Kennedy, and Qing Zhou. 2012. Parents' expression and discussion of emotion in the multilingual family: Does language matter? *Perspectives in Psychological Science* 7: 365–83. [CrossRef]

De Houwer, Jan. 2003. The extrinsic Affective Simon Task. *Experimental Psychology* 50: 77–85. [CrossRef] [PubMed]

De Houwer, Jan, and Paul Eelen. 1998. An affective variant of the Simon paradigm. *Cognition & Emotion* 12: 45–61.

Degani, Tamar, Anat Prior, Chelsea M. Eddington, Ana B. Arêas da Luz Fontes, and Natasha Tokowicz. 2016. Determinant of translation am-biguity. *Linguistic Approaches to Bilingualism* 6: 290–307. [CrossRef]

Dewaele, Jean-Marc. 2008. The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics* 40: 1753–80. [CrossRef]

Dewaele, Jean-Marc, and Seiji Nakano. 2012. Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development* 1: 1–14. [CrossRef]

Eilola, Tiina M., Jelena Havelka, and Dinkar Sharma. 2007. Emotional activation in the first and second language. *Cognition and Emotion* 21: 1064–76. [CrossRef]

Harris, Catherine L. 2004. Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multi-lingual and Multicultural Development* 25: 223–47. [CrossRef]

Harris, Catherine L., Ayşe Ayçíçeğí, and Jean Berko Gleason. 2003. Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics* 24: 561–79. [CrossRef]

Heredia, Roberto R., and Jeanette Altarriba. 2001. Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science* 10: 164–68. [CrossRef]

Ishihara, Shinobu. 1939. *Tests for Colour-Blindness*. Tokyo: Kanehara & Co., Ltd.

Knickerbocker, Hugh, and Jeanette Altarriba. 2013. Repetition blindness with emotion and emotion-laden word targets. *Visual Cognition* 21: 599–627. [CrossRef]

Ladegaard, Hans J. 2018. Codeswitching and emotional alignment: Talking about abuse in domestic migrant-worker returnee narratives. *Language in Society* 47: 693–714. [CrossRef]

Marian, Viorica, and Margarita Kaushanskaya. 2004. Self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language* 51: 190–201. [CrossRef]

Marian, Viorica, Henrike K. Blumenfeld, and Margarita Kaushanskaya. 2007. The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50: 940–67. [CrossRef]

Marmolejo, Gloria, Kristen A. Diliberto-Macaluso, and Jeanette Altarriba. 2009. False memory in bilinguals: Does switching languages increase false memories? *The American Journal of Psychology* 122: 1–16.

McKenna, Frank P., and Dinkar Sharma. 1995. Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology Learning, Memory, and Cognition* 21: 1595–607. [CrossRef]

Pavlenko, Aneta. 2005. *Emotions and Multilingualism*. New York: Cambridge University Press.

Santiago-Rivera, Azara L., and Jeanette Altarriba. 2002. The role of language in therapy with the Spanish-English bilingual client. *Professional Psychology: Research and Practice* 33: 30–38. [CrossRef]

Santiago-Rivera, Azara L., Jeanette Altarriba, Norma Poll, Normaris Gonzalez-Miller, and Carrie Cragun. 2009. Therapists' views on working with bilingual Spanish-English speaking clients: A qualitative investigation. *Professional Psychology: Research and Practice* 40: 436–43. [CrossRef]

Schneider, Walter, Amy Eschman, and Anthony Zuccolotto. 2002. *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Sutton, Tina M., Jeanette Altarriba, Jennifer L. Gianico, and Dana M. Basnight-Brown. 2007. The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion* 21: 1077–90. [CrossRef]

Thalmayer, Amber Gayle, Cecilia Toscanelli, and Jeffrey Jensen Arnett. 2021. The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist* 76: 116–29. [CrossRef]

Tokowicz, Natasha, and Judith F. Kroll. 2007. Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes* 22: 727–79. [CrossRef]

Tokowicz, Natasha, Judith F. Kroll, Annette De Groot, and Janet G. Van Hell. 2002. Number of translation norms for Dutch-English transla-tion pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers* 34: 435–51.

Tseng, Alison M., Li-Yun Chang, and Natasha Tokowicz. 2014. Translation ambiguity between English and Mandarin Chinese: The role of proficiency. In *The Development of Translation Competence: Theories and Methodologies from Psycholinguistics and Cognitive Science*. Edited by Aline Ferreira and John W. Schwieter. Cambridge: Cambridge Scholars Publishing, pp. 107–65.

Zhang, Juan, Chenggang Wu, Yaxuan Meng, and Zhen Yuan. 2017. Different neural correlates of emotion-label words and emotion-laden words: An ERP study. *Frontiers in Human Neuroscience* 11: 1–9.

languages

MDPI

Review

A Selective Review of Event-Related Potential Investigations in Second and Third Language Acquisition of Syntax

Tanja Angelovska [1,*] and Dietmar Roehm [2]

[1] Department of English and American Studies, University of Salzburg, 5020 Salzburg, Austria
[2] Department of Linguistics, University of Salzburg, 5020 Salzburg, Austria; dietmar.roehm@plus.ac.at
* Correspondence: tanja.angelovska@plus.ac.at

Abstract: The aim of this contribution is to highlight the role and relevance of neurolinguistics accounts for second and third language syntactic acquisition/processing. This chapter begins with a brief historical overview of the field of experimental psychology and the birth of the EEG methodology. We then provide a general introduction of the ERP methodology and the language-related ERP components, explaining what they show and how they are to be interpreted. A special focus is given on the clear distinction between behavioral measurements in contrast to real-time measures and the leading role of ERPs is elaborated on. We then provide a selective narrative review of existing L2 and L3 syntax acquisition studies with the EEG methodology within the domain of syntax that we consider relevant for deriving implications for language instructed settings. We discuss results from EEG studies on second and third language syntactic acquisition/processing and finally, highlight several conclusions important for the field.

Keywords: EEG; neurolinguistics; second language acquisition; syntactic processing; third language acquisition

Citation: Angelovska, Tanja, and Dietmar Roehm. 2023. A Selective Review of Event-Related Potential Investigations in Second and Third Language Acquisition of Syntax. *Languages* 8: 90. https://doi.org/10.3390/languages8010090

Academic Editor: John W. Schwieter

Received: 16 May 2022
Revised: 7 February 2023
Accepted: 10 March 2023
Published: 22 March 2023

1. Introduction

From a layperson's perspective, language comprehension seems to be an imperceptible and apparently effortless process. However, the human language processing system is constantly confronted with unexpected, meaning-related conflicting events that must be resolved if comprehension is to proceed successfully. Successful language comprehension depends not only on the involvement of different domain-specific linguistic processes, but also on their respective time-course. A large part of the recent work in psycho- and neurolinguistics has focused on trying to determine which processes play a role and how these processes interact in time. The following section begins with a brief overview of the field followed by an overview of the ERP methodology and the language-related ERP components, highlighting the importance of real-time measurements in contrast to behavioral measures.

2. A Brief Overview of the EEG Methodology

The initial beginnings of experimental psychology lead us back to Wilhelm Wundt. With his paradigm-shifting approach, psychology was established as an empirical discipline, whose aim is to analyze the consciousness processes precisely, to measure elementary perceptions, to break down associated conscious processes and complex interrelationships and find out the regularities behind such relationships. However, the first successful EEG results about the human brain origin from Berger (1929) who, as a result, received the recognition as founder of the human electroencephalogram. He not only expanded earlier research with animals to the human brain, but he also delivered a thorough description of the conditions, under which these rhythmic EEG processes were observed. His biggest achievement is his first description of an objective EEG correlate of the mental state. In other words, cognitive changes are reflected through the observed rhythmic activation in EEG. For instance, an alpha rhythm (a 10 Hertz oscillation) is shown in adults when they are relaxed and awaken with their eyes closed. The so-called Berger-effect (alpha-blockade) emerges when the person opens the eyes, and the mental state is challenged (i.e., alpha waves vanish or are reduced during problem solving). This effect was the decisive starting point for the area of psychophysiological EEG research (Altenmüller and Gerloff 1999). The EEG began to be used for detecting electrical potentials in the well-structured form of brain waves. The new discipline focused on the relationship between the frequencies of these brain waves and human behavior.

In 1951, the electromechanically based summative method as a means of EEG research technique was introduced (Dawson 1954). With the development of the digital computers in the early sixties several labs managed to apply this technique successfully to electrophysiological data. This milestone represented the basis for an improvement in the signal-to-noise ratio in the EEG analysis and led to the discovery of small endogenic event-related potentials (ERPs), which are bound to sensory and cognitive stimulus (e.g., the presentation of a word). After the importance of these potentials was recognized in the sixties, the EEG-based research focused exclusively on the research of the ERP-components (Altenmüller and Gerloff 1999). This tendency was strengthened through the development of digital computer techniques, which allowed that ERPs are calculated with special computer programs and represented as time-dependent functions (Barlow 1957; Brazier 1960). It was shown that ERPs vary systematically with stimulus properties (e.g., tone pitch, color, intensity) with respect to several measurable parameters (amplitude strength, latency, polarity, topography). However, it took another two decades to achieve a breakthrough in EEG and language.

Kutas and Hillyard (1980) reported in a science publication for the first time a clear language-related ERP correlate in a ground-breaking study investigating the processing of English sentences in three different conditions. The conditions were created based on the sentence's final word which was either syntactically and semantically well-formed (1a), included a plausibility violation (1b), or a mere physical deviation (letters size; 1c):

(1a) It was his first day at work;
(1b) He spread the warm bread with socks;
(1c) She put on her high heeled SHOES.

(Kutas and Hillyard 1980)

The activity at the critical sentence's final word was measured. While the physical deviation led to a late positive potential, the semantic violation resulted in a late negative component with an amplitude maximum at around 400 ms after the beginning of the presentation of the critical word (=N400). An N400 is, however, not only a correlate of semantic plausibility, but it is also influenced by several other factors such as word frequency, sentence context, repetition in relation to lexical status (open/close class words) and the degree of unexpectedness of a word, as well as a plethora of other lexical-semantic and contextual factors (for an overview see Kutas and Federmeier 2011). Furthermore, it was found that the N400 is part of the normal brain activity during the perception and processing of words and other meaningful (or potentially meaningful) stimuli (e.g., signs in sign language, numbers, pictures, faces, sounds and odors).

Despite the vivid rise in EEG research and related publications on sentence processing following Kutas and Hillyard (1980), the search for a syntactic ERP correlate (as a counterpart of the lexical-semantic N400) took another decade. It was not later than the 1990s when Osterhout and Holcomb (1992) described a late positive component as a correlate of a syntactic violation, which emerged with a wide parietal distribution and a latency of 600 ms after the stimulus onset. They labelled this component as the P600. Indeed, they found that this component not only shows up when the sentence was clearly ungrammatical (as in example 2), but also in so-called garden-path sentences (example 3), in which a temporal structural ambiguity (here, the verb "persuaded" with an ambiguity between a finite and a non-finite reading) leads to a wrong initial syntactic analysis, which can be revised later on the basis of the apparently incompatible information (here, the infinitive marker *to*).

(2) * The broker hoped to sell the stock was sent to jail;
(3) The broker persuaded to sell the stock was sent to jail.

Hence, it was suggested that the P600 is an index of syntactic reanalysis and repair processes. In addition, the P600 was found in sentences with a more complex structure (e.g., wh-questions) and was discussed as a correlate of syntactic integration costs (Kaan et al. 2000). For example, Burkhardt (2005, 2006) reported a P600 as a correlate of the establishment of new discourse referents in a mental model. Because the P600 was also found in non-linguistic domains (e.g., for violations of mathematical rules; Lelekov et al. 2000) it was suggested that the P600 always emerges when a stimulus is hard to be integrated in the structure of the preceding context. Since then, many other ERP effects as correlates of language processing processes in different domains were found (e.g., for congruency violations, scrambling, subject preference, etc.; for an overview see Bornkessel-Schlesewsky and Schlesewsky 2009; Kaan 2007).

Prior to discussing the specifics about ERPs (cf. Roehm 2004), we begin with an example of a language ambiguous sentence, which leads to comprehension difficulties. Such an example was given in (3) and is repeated as (4) (from Osterhout and Holcomb 1992, 1993):

(4) The broker persuaded to sell the stock was sent to jail.

When sentence (1) is processed sequentially, the verb *persuaded* is initially analyzed as the finite main verb (as in "The broker persuaded the manager to sell the stock."). This decision must be revised when the reader/listener reaches "to" in the sentence as it becomes evident that persuaded is a non-finite verb in a reduced relative clause. Hence, the comprehension difficulty is a result of an ambiguity, as the reader/listener is prone to misanalysis pertaining to properties of syntactic structure. The comprehension difficulty is evident as a processing cost is imposed on the reader/listener. This type of enhanced processing cost gives insights into the architecture and mechanisms of the language processing system (Kimball 1973; Fodor et al. 1974; Frazier 1987; Clifton et al. 1994). Similarly, as in the investigations of other cognitive domains (e.g., memory and attention), the simplest account of this "processing difficulty" is realized through behavioral measures (for example, reaction times measurements). However, using these measures as a means of characterizing underlying mechanisms of linguistic analysis presupposes that the locus of the processing problem can be straightforwardly established. In this way, for an implausible sentence such as in (5), the reader/listener will certainly have longer reaction times in the critical region and lower acceptability ratings in contrast to a plausible sentence that will end with *butter.*

(5) He spread the warm bread with socks.

In contrast to example (4) where the processing cost was due to a structural ambiguity (due to the ambiguity of the verb *persuaded*), in (5) the processing costs are clearly based on the semantic oddness and thus, can be attributed to the lexico-semantic processing domain. Evidently, behavioral measures present pure quantitative measures of processing difficulty and cannot disentangle underlyingly different linguistic domains from one another. They merely provide rather unspecific global measures of processing difficulty representing only the result of the comprehension process locally (i.e., on the word level) and/or globally (i.e., on the sentence level). We can conclude that behavioral measures (such as, e.g., self-paced-reading, speeded-grammaticality judgments, or lexical decisions) do not allow conclusions about the precise time course of underlying processing mechanisms (Schütze 1996).

While chronometric procedures allow only for indirect measures of psychological processes in the form of reaction times (RTs) regarding a given stimulus, the measurements of brain waves with electroencephalography directly reveal the psychophysiological processes in the brain. In other words, the RTs are a measure of the outcome of cognitive processes and the brain waves measure the cognitive processes as they unfold in real time. To fully represent the linguistic processing domains including its temporal processing characteristics, we need comprehension measures that provide not only quantitative estimations but allow for qualitative characterizations and an ongoing assessment of the comprehension process (or rather, its underlying processing characteristics). The comprehension process is the cornerstone of both first and second language (L2) processing. While behavioral measures only assess the overall performance of learnt L2 knowledge (i.e., final product based on one's displayed competence), real-time measures (such as EEG) track the online processing, internalization, and proceduralization of L2 knowledge in a moment-by-moment manner. Comparing the EEGs by second language learners with those from native speakers enables us to gain insights into the similarities and differences in the process itself and to observe how far the EEGs of L2 learners are native-like.

3. Neurolinguistic Measures for Syntactic Processing: Event-Related Brain Potentials (ERPs) and Language-Related ERP Components

The recording of the human electroencephalogram (EEG) is one of the methods that provides a direct reflection of underlying brain processes with the advantage to supply an excellent time resolution in the millisecond range. Using EEG, we obtain a direct (although damped) reflection of the summed electrophysiological activity of the brain by means of electrodes applied to the surface of the human scalp. Throughout the presentation of words/sentences, the event-related brain potential (ERPs) technique provides a continuous record of language comprehension processes as they unfold in real time. As mentioned above, ERPs provide a multi-dimensional characterization on processing difficulties during language comprehension, in which various language-related components are identified based on parameters such as polarity, amplitude, latency and topography. The high temporal resolution of EEG signals is a property which is of primary importance if the very rapid and complex processes that make up language comprehension and production are to be captured adequately (Bornkessel-Schlesewsky and Schlesewsky 2009). However, one must keep in mind that ERPs represent the average of EEG samples obtained in several trials (typically, between 30 and 40) belonging to the same experimental condition. Thus, the averaged waveforms represent estimates of the time-locked neural activity engendered by the presentation of stimuli belonging to different experimental conditions. Differences between ERP waveforms derived from different conditions therefore represent differences in the neural activity engaged by the items belonging to each condition (cf. Roehm 2004).

Only through the application of the ERP methodology, the syntactic processing difficulty in (4) and the semantic violation in (5) indeed elicit distinct ERP components, namely a parietal positivity (Osterhout and Holcomb 1992, 1993) with a maximum at approx. 600 ms (P600 or 'syntactic positive shift', SPS) vs. a centro-parietal negativity with a maximum at approx. 400 ms post critical word onset (N400; Kutas and Hillyard 1980). Based on such findings, the N400 was initially seen as an unambiguous general marker of lexical-semantic processes (Kutas and Federmeier 2000) used as a 'diagnostic tool' in cases where the nature of the observed processing difficulty could not be established straightforwardly. However, later ERP findings revealed that precisely the N400 can be found in several areas which are clearly independent of the lexical-semantic domain (e.g., Bornkessel et al. 2002; Osterhout 1997). The fact that the N400 component cannot be attributed to a single specific language processing domain therefore shows that the desired one-to-one mapping between ERP components and linguistic processes is at least problematic (see below).

Most of the EEG-based language research includes ERPs. However, complementary analyses of stimulus-related oscillatory EEG activity (rhythmic activity in the EEG) have only recently come to be used. Oscillatory activity is an inherent property of neurons and neuronal assemblies. It is generally agreed upon that the timing of oscillatory dynamics encodes information (Buzsaki and Draguhn 2004). Oscillations in the theta, gamma, and beta frequency range have been linked to mental effort and have recently been related to language processing (Lewis et al. 2016), whereas theta and gamma oscillations (theta: 3 to

7; gamma: 25 to 100 Hz (typically 40 Hz)) are associated to the encoding and retrieval of declarative knowledge and beta oscillations (13 to 30 Hz) have been treated as a potential index of syntactic integration, and more recently as a reflection of predictive processing during language comprehension. The next section includes a discussion about the language-related ERP components.

Several language-related ERP components have been identified and viewed as correlates of linguistic domains (e.g., semantic, and syntactic processing). As discussed above, the N400 is a negative ERP-deflection peaking around 400 ms after the onset of a potentially meaningful stimulus. In language comprehension, it has been interpreted as reflecting lexical–semantic processing (e.g., the integration of word-associated information into a meaningful context). Thus, the N400 is considered a reflection of 'contextual integration'. However, it is not only a reflection of the relative ease of semantic integration, but the N400 amplitudes also vary as a function of frequency (Van Petten and Kutas 1990) and repetition (Rugg 1990; Van Petten et al. 1991). The P600 is a late positive component with a broad parietal distribution and a typical latency between 600 and 900 ms post critical word onset (Osterhout and Holcomb 1992). It was first observed as a correlate of outright syntactic violations (following an early left anterior negativity, the so-called ELAN) and interpreted as an index of repair processes in sentences such as (6) (from Osterhout and Holcomb 1992).

(6) * The broker hoped to sell the stock was sent to jail.

The ELAN (Friederici 1995), an early left-anterior negativity between 150–200 ms, has been observed in several ERP studies in situations where the brain is confronted with phrase structure violations due to outright word category violations (such as the word "to" in example 6). It is typically interpreted as a highly automatic correlate of initial structure-building processes (first pass parsing processes responsible for local phrase structure building; Friederici 1995, 1999; Hahne and Friederici 1999). A slightly later left-anterior negativity labeled as LAN with a maximum between approx. 300 and 500 ms after onset of the critical item has been reported for agreement violations with legal words (e.g., subject-verb agreement or wrong pronoun case) (cf. Coulson et al. 1998; Kutas and Hillyard 1983; Osterhout and Mobley 1995; Gunter et al. 1997), as well as with morphologically marked pseudowords (from Münte et al. 1997). It has been suggested (e.g., Münte et al. 1997) that the LAN specifically reflects the actual detection of a morphosyntactic mismatch. However, because a LAN can also be found in grammatically correct sentences, others have claimed that it indexes some aspect of working memory usage (Kluender and Kutas 1993a, 1993b; King and Kutas 1995; Rösler et al. 1998).

Following the declarative/procedural model (Ullman 2001), the declarative memory system underlies the mental lexicon (storing rote-learned linguistic mappings) and the procedural memory system underlies the mental grammar (necessary for the acquisition and use of rule-governed computations in language). Indeed, it has been shown that declarative memory processes are likely to evoke N400 responses (cf. Morgan-Short et al. 2010), while procedural memory processes (especially with respect to morphosyntactic relations) primarily show ELAN/LAN responses. Moreover, it has been argued that ERPs can reveal implicit processing of (morpho-)syntactic knowledge that may be distinct from explicit (morpho-)syntactic processing as evidenced in grammaticality judgment tasks (GJT), and thus may allow differentiating cross-linguistic differences and similarities (Tokowicz and MacWhinney 2005). With respect to classroom-like instructed subjects, N400 and P600 effects have been found with L2 learners even when they lack immersion exposure (Morgan-Short et al. 2010) that is regarded essential for implicit learning and the consolidation of knowledge.

Findings that language-related ERP components can be used as diagnostic indicators of linguistic or cognitive domains, such as the N400 as a correlate of lexical–semantic processing or declarative memory contributions to language, versus the LAN/P600 as correlates of morphosyntactic processing or procedural memory, seem highly appealing. However, this possibility has been recently challenged by a continually increasing number of findings (Bornkessel-Schlesewsky and Schlesewsky 2009).

Having reviewed the general EEG methodology and the ERP components, in the next section, we discuss findings from event-related potentials relevant for the second language acquisition process.

4. EEG and Second Language Acquisition of Syntax

The field of neurocognitive SLA has followed two main aims in the past: (1) to compare the neurocognitive representation of L2 processing with that of native speakers' and, subsequently, (2) to explain possible differences or similarities as a function of factors, such as proficiency, or age of onset in the L2, etc. (Angelovska and Roehm 2020). The most comprehensive and informative narrative review of L2 ERP research encompassing studies from 2008 to 2013 (Morgan-Short 2014) has identified several findings which are relevant for language teachers. The reilluminating focus in EEG studies in the field of second language acquisition of syntax falls under the domain of investigating whether the ERP patterns between L2 learners and native speakers differ and which type of training (implicit or explicit) is more beneficial for comprehension and/or production of the target language, as well as which factors moderate such effects. L2 proficiency has been considered as one of the most influential factors. For instance, regarding L2 proficiency and (morpho-)syntactic processing in late L2 learners, Morgan-Short et al. (2010) reviewed a body of ERPs studies indicating that the L2 learners' neural activity in the brain during processing (morpho-)syntactic violations can be modulated by several inter-related factors, such as the similarity or dissimilarity of syntactic structures in L2 and L1 and the L2 learners' level of L2 proficiency—a finding which is relevant for second language educators interested in deriving ERP-based implications for teaching. We build upon this review focusing on newer results and including studies of relevance for instructed settings. We divide this review of relevant studies according to whether the target language is a natural or an artificial language.

4.1. EEG Studies with an Artificial Language as a Target L2

Grey et al. (2017) examined the role of learners' language experience in an EEG study which included target language training and behavioral and neural correlates as measures. They included Mandarin–English bilinguals and English monolinguals following an explicit grammar instruction session of only 13 minutes using an artificial language as a target language. Learners were trained in comprehension and production and received the task to judge sentences grammaticality. Despite the finding that monolinguals and bilinguals did not differ on the behavioral measures, they differed regarding the ERPs whereby proficiency proved to matter considerably. More specifically, at low-proficiency level, only bilinguals showed an expected P600 effect, whereas at high proficiency both monolinguals and bilinguals showed a P600. An anterior positivity was found for monolinguals only, which is not typically found in native speakers. The general conclusion is that when acquiring an additional target language and even after a minimal amount of instruction, bilinguals show ERP patterns similarly to native speakers.

SLA researchers still debate the question of which type of instruction (explicit or implicit) is more effective for learners. Morgan-Short et al. (2010) tested low- and high-proficiency learners who were instructed with explicit (classroom-like) and implicit (immersion-like) training when acquiring an artificial language as a target language. Regarding the behavioral measures, no group differences were found. However, group differences were evident regarding the ERP patterns. The implicit group showed significant native-like patterns on noun–adjective and noun–article agreement and were considerably more successful than the explicit group, which showed gains only for noun-article agreement. Morgan-Short et al. (2010) interpret these results in line with the declarative/procedural model maintaining the view that with an increasing L2 experience and proficiency, learners will become more dependent on L1 neurocognitive mechanisms, whereas if learners are less proficient, they would rely more on the declarative memory for lexical and semantic processing.

In a subsequent longitudinal EEG study, Morgan-Short et al. (2012) confirmed the results from Morgan-Short et al. (2010) regarding whether explicit and implicit training would differentially affect neural and behavioral performance measures of syntactic processing. The confirmatory findings highlight the benefits of the implicit training. To be precise, the implicitly trained group showed native-like patterns (an anterior negativity followed by a P600 accompanied by a late anterior negativity) at a high proficiency level and only an N400 at a low proficiency level. The explicitly trained group demonstrated only an anterior positivity followed by a P600 at high proficiency and no significant effect at low proficiency. The results clearly speak for language learners' reliance on native-like mechanisms, as evident through electrophysiological signatures, only when they are exposed to implicit instruction. An important conclusion relevant for second language pedagogy is that language teaching practitioners should focus on increasing the amount of exposure to implicit training conditions for their learners, thereby maximizing the immersion-like and communicative experience for their learners.

(Pili-Moss et al. 2019) investigated whether learners' declarative and procedural abilities predict comprehension and production accuracy and automatization when doing L2 practice. Their results point to positive associations between comprehension accuracy and declarative learning ability, which did not change across practice. Furthermore, they found that with increasing procedural learning ability, the automatization increased as well. Likewise, learners with a higher declarative learning ability showed stronger automatization. A more recent synthesis of studies with neurocognitive measures within the domain of second language research with artificial languages can be found in Morgan-Short (2020).

We can derive two conclusions based on the reviewed EEG studies with an artificial language as a target language: (1) in both studies (Grey et al. 2017) increasing language experience and proficiency proved to be the main factors accounting for more native-like ERP signatures; and (2) implicit instruction seems to be inevitable and more beneficial for the second language learners to obtain native-like mechanisms as automatization is increased, as found in Morgan-Short et al. (2012) and Pili-Moss et al. (2019).

4.2. EEG Studies with a Natural Language as a Target L2

With respect to the EEG components in L2 acquisition, P600 is often considered a measure of native-like processing, while smaller or delayed N400 and P600 effects were found for L2 learners compared to native speakers (Hahne 2001; Sabourin and Stowe 2008). However, these findings depend on several factors which are implemented in the analyses and shown to moderate these effects. In what follows, we review selected ERPs studies about syntactic acquisition with relevance to classroom implications. We begin by reviewing studies that consider the effects of the linguistic factors' effects (e.g., similarity, prosody, etc.) on the outcomes of L2 processing, then we focus on considering the role of proficiency and working memory as variables in ERP studies and finally we review studies which focused on the training type and context, which is decisive for deriving a conclusion of relevance for instructional contexts.

Kotz et al. (2008) investigated how native and non-native readers of English respond to syntactic ambiguity and anomaly. The learners they focused on were highly proficient L1 Spanish and L2 English individuals who had acquired English informally around the age of five years. Their results point to the fact that early acquisition of L2 syntactic knowledge leads to comparable online sensitivity towards temporal syntactic ambiguity and syntactic anomaly in early and highly proficient non-native speakers. Guo et al. (2009) aimed at finding out which kind of information (semantic or syntactic) is more important in sentence processing when verb sub-categorization violations are included. Using ERPs, they investigated strategies employed by 17 native English speakers and 28 L2 English learners with Chinese as L1. They found a P600 effect to verb sub-categorization violations for the native speaker group and an N400 effect for the L2 learners. Even though they interpret these findings as electrophysiological evidence for different strategies used by native speakers and L2 learners in sentence processing (i.e., shallower processing), one must be aware that the proficiency measure they used in their study is not robust (self-rating scores) without any stay-abroad experience. Both studies Kotz et al. (2008) and Guo et al. (2009) offer relevant pieces of evidence for considering the additional type of information when interpreting ERP results.

Nickels and Steinhauer (2018) considered the role of prosodic information available in syntactic processing which is a rather not addressed aspect in L2 instruction with neurocognitive studies on prosody–syntax interactions being rare. They analyzed the ERPs of Chinese and German learners of English as L2 to those of native English speakers (control group). Their findings showed that the L1 background and L2 proficiency influence the online processing whereby the garden-path effects were triggered by prosody. They used linear mixed effect models treating the L2 proficiency factor as a continuous variable. Their main finding refers to the fact that both L1 background and proficiency are predictors for the syntactic processing. Similarly, they (2020) investigated the ERP signatures of beginning, intermediate and advanced learners when processing L1–L2 word order conflicts. They found N400-like signatures for beginners. In contrast, the intermediate and the advanced group showed native-like P600 signatures for the violated sentences without conflict between L1 and L2, and delayed P600 signatures for the violated sentences with L1-L2 conflicts, although behaviorally both groups were on a native level. It was shown that processing L2 word-order that conflicts with L1 is an obstacle even at an advanced proficiency level, and that progression through developmental stages depends not only on proficiency, but also on how similar L1 and L2 are of the learners.[1]

Tokowicz and MacWhinney (2005) investigated explicit and implicit processes during L2 sentence comprehension using ERPs and a grammaticality judgment task among 20 L2 Spanish learners with L1 English. Three syntactic phenomena were tested: tense-marking (which exhibits similarity to the L1), determiner-number agreement (which exhibits differences to L1) and determiner-gender agreement (unique to the L2). The experimental sentences varied the form of three different syntactic constructions: (a) tense-marking, which is formed similarly in the first language (L1) and the L2; (b) determiner-number agreement, which is formed differently in the L1 and the L2; and (c) determiner-gender agreement, which is unique to the L2. They found P600 sensitivity for tense-marking violations in L2, but this effect was not found for determiner-number agreement and for determiner-gender agreement. These findings show that L2 learners' processing depends on the similarity between the L1 and the L2.

We can conclude that Nickels and Steinhauer (2018), Mickan and Lemhöfer (2020) and Tokowicz and MacWhinney (2005) agree in bringing forth one main conclusion that includes not only proficiency but also similarity between the L1 and the L2 in order to predict how L2 learners process syntactic information in an L2.

We now turn to elaborating on the role of proficiency and working memory in L2 sentence processing. Dallas et al. (2013) examined whether young adult L2 English learners with L1 Chinese would show differential results in relation to their proficiency level and working memory capacity. Using ERPs, they investigated L2 English learners' real-time processing of sentences containing filler-gap dependencies. Their finding revealed that with rising L2 proficiency, the sensitivity to plausibility variations increases likewise, regardless of learners' working memory capacity. However, working memory proved to be a strong predictor of learning in more explicit settings with behavioral measures (e.g., Linck and Weiss 2015). Another study with important findings for the learning context and working memory is Faretta-Stutenberg and Morgan-Short (2018). They focused on examining the role of the acquisitional context. In two longitudinal studies, they tested young adult (at university level) native speakers of English studying L2 Spanish in an "at-home" ($n = 29$) and "study-abroad" context ($n = 20$) at the intermediate level. They used measures of learners' working memory capacity, declarative and procedural learning ability, and EEG data, which was collected while learners completed a syntactic grammaticality judgment task. Results revealed gains for at-home intermediate learners in their ability to detect phrase structure violations and N400 and P600 effects at the end of the semester. However, the study-abroad learners showed larger behavioral gains, which were accounted for by their gained procedural learning ability and working memory capacity. It seems that more immersed-like learning conditions are more beneficial when acquiring a target language.

Bowden et al. (2013) included young-adult learners without immersion experience and at a low proficiency level. Bowden and Steinhauer et al. (2013) tested two groups of late-learned L2 Spanish speakers (low vs. advanced proficiency) and a control group of L1 Spanish speakers. Regarding semantic processing, N400s were found in all three groups, whereas LAN/P600 responses were evident for syntactic word-order violations only in the native speakers and in the L2 advanced group, with a statistically indistinguishable outcome. The authors concluded that L2 semantic processing relies on similar neural process as in L1, whereas the factors proficiency or exposure do not moderate these effects. On the other hand, L2 syntactic processing differs from L1 syntactic processing at the low proficiency levels; however, with increasing proficiency and exposure (immersion experience), it can become native-like. In other words, the language learning process qualitatively changes the syntactic processing of L2 learners.

Both studies (Faretta-Stutenberg and Morgan-Short 2018 and Bowden et al. 2013) bring evidence about the benefits of immersion settings for the acquisition of syntax, which has influential outcomes for how learning conditions should be planned by educators.

We now analyze the effects of the differential training types in L2 acquisition. Within a pre- and post-test design, Davidson and Indefrey (2009) examined L1 Dutch learners' development of L2 German (adjective declension and both article–noun and adjective–noun gender agreement) during only one instructional session, which consisted of explicit training on German. Learners were instructed and given grammatical information about adjective declension and gender and given short feedback (correct/incorrect). The post-test was delivered one week after the instruction. The control group with L1 German speakers completed the final test only. The findings reveal that the L2 learners improved, showing a P600 response only for declension. Despite the result suggesting that even a relatively short training session may have resulted in instructional benefits and a P600 response development in L2 learners for adjectival declension, we do not know whether the result will be confirmed if learners underwent a different type of instruction. We also lack information on the amount of instruction needed for the P600 response to appear for all cases.

Batterink and Neville (2013) examined whether second language learners recruit some of the same language-processing mechanisms as those employed by native speakers of the same language. They investigated L2 syntactic processing on article-noun agreement, subject-verb agreement, and word order with 24 French native speakers and 67 L1 English speakers learning French as L2 who were assigned two differential trainings (implicit vs. explicit). The implicitly trained group ($n = 44$) did not receive any explicit rule provision, while the explicitly trained group ($n = 23$) received formal instruction on the underlying grammatical rules. Their findings show that both instructed groups did not show any differences in the comprehension scores; however, the native speaker group achieved significantly higher comprehension scores. On the grammaticality judgment task, the explicitly trained group performed better than the implicitly trained group. Regarding the ERP results, the native speaker group showed a biphasic response, consisting of an early negativity effect followed by a later P600 effect, and the early negativity effect did not prove significant in the noun agreement condition. Similarly, both implicitly and explicitly trained high proficiency groups showed P600 effects to all three agreement conditions but the low proficiency implicitly trained group did not elicit P600 in any of the conditions. The only condition in which the training predicted the improvement effects was the word order violation condition where the implicit training predicted the P600 amplitude. The authors concluded that even a short implicit training alone is predictive of a P600 effect for L2 learners, suggesting that the implicit training results in L2 learners recruiting the same neural mechanisms as those employed by native speakers even at early acquisitional stages of syntactic processing and regardless of the qualitative differences.

We can conclude that the results regarding the beneficial effects of explicit versus implicit training still are inconclusive. Generally, some studies (e.g., Guo et al. 2009; Morgan-Short et al. 2012) found N400 instead of the expected P600 for late bilinguals in their early stages of L2 acquisition. In the case when late learners demonstrate native-like processing, it is mainly dependent on proficiency, although the problem of disentangling age of onset from proficiency has already been acknowledged (Friederici et al. 2002). Lower proficiency learners have demonstrated weaker P600 in comparison to higher proficiency learners and native speakers (e.g., Bowden et al. 2013). Early and highly proficient bilinguals' syntactic processing resembles native processing (P600) when both L1 and L2 share the syntactic property. In regard to the learning conditions, we can conclude that ERP studies on natural L2 training under differential input conditions in instructed classroom settings is limited to non-existent. These types of studies are very extensive and difficult to conduct in such "noisy" settings as natural language classrooms and often carry several confounding factors that are hard to control for. For example, classroom foreign language learners are at different ages and proficiency levels and differ in other individual variables of both an affective and cognitive nature. Certainly, the instructed SLA domain would profit from ERP results which test language acquisition in designs where the target language is a natural language not biasing the ongoing explicit-instruction trend. However, it remains a challenge to create such experimental designs that would meet the criteria of both lab and instructional conditions.

5. EEG and L3 Acquisition of Syntax

The research on L3 morphosyntactic processing using EEG is rather scarce with a clear focus on transfer. Eye-tracking results in L3 studies (Abbas et al. 2021) revealing that during L3 processing, L3 learners have access to both L1 and L2, which can be the source of transfer in L3 but with different time-courses, have been corroborated by EEG results. The first study to compare bilinguals' processing to monolinguals' processing of an additional language using ERPs was by Grey et al. (2017). L1 Mandarin-L2 English bilinguals and L1 English monolinguals learned an artificial language, Brocanto2. The behavioral data showed no group differences on comprehension and production accuracy or RTs. There was an improvement of the GJT accuracy after a second training compared to the first session, but no significant differences between groups. The bilingual group exhibited a P600 response to ungrammatical sentences in the first session, whereas after the second training both bilinguals and monolinguals showed a P600 response. Moreover, monolinguals exhibited an anterior positivity in the 400–700 ms time window that was absent in bilinguals, who had a visually apparent but not significant anterior negativity instead. These findings prove that even without bilingual/monolingual behavioral differences, bilinguals show distinct ERP patterns which are more like native speakers.

Andersson et al. (2019) examined how native Swedish speakers (controls) and German (+V2 in L1) and English (−V2 in the L1) learners of Swedish produce, judge, and process grammatical V2 and ungrammatical verb-third (V3) word orders in Swedish sentences with sentence initial adverbials. Behaviorally, the learners differed from the native speakers only on judgements, but all were highly proficient. Neurocognitively, all groups showed a similar increased posterior negativity, as well as posterior P600 ERP-effect, but German learners displayed more native-like anterior ERP-effects than English learners. Interestingly, learners and native speakers did not differ in P600 amplitude. German learners had a similar anterior negative effect as Swedish learners, whereas English learners had a larger frontal positivity. This could suggest that they used more attentional resources, while German learners used a less demanding processing strategy because their L1 has a similar structure to Swedish. The results suggest crosslinguistic influence in that the presence of a similar word order in the L1 can facilitate online processing in an L2, even if no offline behavioral effects are discerned. Despite the information that the AoA of Swedish for L1 German learners was 21;5 (SD = 2;5) and exposure 3;4 (SD = 2;10), whether English was the learners' L2 has not been specified. Most German schools teach English from the 5th grade, which corresponds to AoA of English at about 11 years old, making it likely that L1 German speakers acquired English before Swedish. Therefore, it is not clear whether this was the first study of L3 Swedish word order processing.

Gonzalez Alonso et al. (2020) used artificial languages (Als), Mini-English and Mini-Spanish, to distinguish several transfer models using ERPs. One group of L1 Spanish-L2 English participants learned Mini-English as L3 and the other Mini-Spanish, both at lower intermediate to advanced proficiency in English. The mini grammars used English and Spanish lexicons, respectively, while the morphemes were the same in both ALs. Grammatical gender, only present in Spanish, was featured in both ALs. There were no differences between conditions (grammatical, gender violation and number violation) in accuracy and RTs. There was an early fronto-lateral negativity for gender violations in Mini-English, as well as a broad positivity over all regions for gender violations in Mini-Spanish. The early positivity could be interpreted as P300, since gender violations might have been attended to as target stimuli and detected more easily due to this grammatical feature being present in Spanish.

What these L3 acquisition studies of syntax have in common is the fact that both found typology effects as attested with EEG measures accounting for the fact that syntactic processing resembles native processing if the languages under observation share typological similarity.

6. Conclusions

We began with a brief historical overview of the field of experimental psychology and the birth of the EEG methodology. We then provided a full description of the ERP methodology and the language-related ERP components explaining what they show and how they are interpreted. A special focus was given to the clear distinction between real-time measurements in contrast to behavioral measures, and the leading role of ERPs was elaborated on.

We provided a narrative review of classroom-relevant EEG studies about the acquisition of syntax in L2 and L3 by reviewing selected studies. Based on our review, we can conclude that an adult second-language learner's brain is highly dynamic, even during the earliest stages of L2 learning (Osterhout et al. 2008). Furthermore, proficiency as a factor matters in explaining the variation in EEG patterns in such a way that only a higher proficiency is associated with P600 effects typical for monolingual native speakers. Further, less proficient learners employ declarative knowledge and more proficient bilinguals rely on native-like mechanisms, when exposed to implicit instruction. Late bilinguals do not show native-like neurocognitive signatures in the early stage of L2 acquisition. Early and highly proficient bilinguals' syntactic processing resembles native processing if the L1 and L2 are typologically similar. Other factors that showed to be predictors are working memory, timing, and instructional setting. Early acquisition of L2 syntactic knowledge leads to a higher mastery of syntax. Thus, for language policy makers, it is relevant to plan an early introduction of syntactic instruction in foreign languages in the respective national curricula. The results from the L3 studies using EEG methodology confirm findings from SLA.

Despite the fact that, on one hand, we evidenced very few EEG studies having examined ecologically valid learning situations that can tell us about second language education itself, the findings of the so far conducted EEG studies in lab settings offer to second-language educators important implications that can be derived about the nature of second/third language knowledge, its consolidation and proceduralization, as well as about the mechanisms involved in these processes. Based on such findings, second language educators can learn more about how particular language domains are processed and acquired and what teaching techniques they should employ to foster that particular type of knowledge generation and consolidation.

Having in mind the importance of these findings for the actual language teaching, we hope that teachers will deepen their understanding of the neurocognitive processing in L2 and make use of research-informed implications. We allow ourselves a note of caution in this context, as sometimes deriving concrete pedagogical implications is rather challenging, especially when the reported results show differential ERP effects across studies which also differ in the amount of input exposure and controlled individual differences, which could account for language learner variation. Moreover, the number of EEG studies comparing various instructional interventions is limited to non-existent. This makes it likewise impossible to derive any conclusions about the effectiveness of such interventions as neurocognitive assessment measures are lacking. L2 instructional interventions focus on providing learners with meaningful input. This is relevant from an acquisitional point of view. However, as pointed out by Angelovska and Roehm (2020), it is not feasible to suggest implications of practical relevance for teaching if EEG results about the effectiveness of the different grammar intervention types are lacking. In addition, what further remains unclear is: (a) whether and to what extent classroom instructional grammar interventions can compensate for immersion-context acquisition opportunities; (b) what duration should those interventions have; (c) what type of feedback should be given and to what extent would it be sufficient if we want our learners to generate more native-like neurocognitive responses. They called for longitudinal L2 EEG studies.

Finally, with continued interdisciplinary approaches and sophisticated research designs, the potential of ERP results to inform central questions of second and third language acquisition needs to be further explored.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Note

[1] According to an anonymous reviewer, with similar behavioral outcomes for both groups, it is unclear whether the neural differences between groups suggest an "obstacle" per se or simply a different neurocognitive mechanism of accomplishing the same outcome.

References

Abbas, Nawras, Tamar Degani, and Anat Prior. 2021. Equal Opportunity Interference: Both L1 and L2 Influence L3 Morpho-Syntactic Processing. *Frontiers in Psychology* 12: 673535. [CrossRef]

Altenmüller, Eckart O., and Christian Gerloff. 1999. Psychophysiology and the EEG. In *Electroenecphalography: Basic Principles, Clinical Applications, and Related Fields*, 4th ed. Edited by Niedermeyer Ernst and Lopes da Silva Fernando. Charm: Baltimore.

Angelovska, Tanja, and Dietmar Roehm. 2020. Neurolinguistic Implications for L2 Learning and Teaching. In *The TESOL Encyclopedia of English Language Teaching*. Edited by John I. Liontas, TESOL International Association and M. DelliCarpini. Hoboken: Wiley Online Library. [CrossRef]

Barlow, John S. 1957. An electronic method for detecting evoked responses of the brain and for reproducing their average waveforms. *Electroencephalography and Clinical Neurophysiology* 9: 340–43. [CrossRef]

Batterink, Laura, and Helen Neville. 2013. Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience* 25: 936–51. [CrossRef]

Berger, Hans. 1929. Über das Elektrenkephalogramm des Menschen. 1. *Mitteilung. Archiv für Psychiatrie* 87: 527–70. [CrossRef]

Bornkessel, Ina, Matthias Schlesewsky, and Angela D. Friederici. 2002. Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition* 85: B21–B30. [CrossRef]

Bornkessel-Schlesewsky, Ina, and Matthias Schlesewsky. 2009. *Processing Syntax and Morphology: A Neurocognitive Perspective*. Oxford: OUP.

Bowden, Harriet Wood, Karsten Steinhauer, Cristina Sanz, and Michael T. Ullman. 2013. Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia* 51: 2492–511. [CrossRef]

Brazier, Mary A. B. 1960. Some uses of computers in experimental neurology. *Experimental Neurology* 2: 123–42. [CrossRef]

Burkhardt, Petra. 2005. *The Syntax-Discourse Interface: Representing and Interpreting Dependency*. Amsterdam: John Benjamins.

Burkhardt, Petra. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language* 98: 159–68. [CrossRef]

Buzsaki, Gyorgy, and Andreas Draguhn. 2004. Neuronal Oscillations in Cortical Networks. *Science* 304: 1926–29. [CrossRef]

Clifton, Charles, Jr., Lyn Frazier, Keith Rayner, and Charles Clifton. 1994. *Perspectives on Sentence Processing*. Hillsdale: Lawrence Erlbaum Associates.

Coulson, Seana, Jonathan W. King, and Marta Kutas. 1998. Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes* 13: 21–58. [CrossRef]

Dallas, Andrea, Gayle DeDe, and Janet Nicol. 2013. An Event-Related Potential (ERP) Investigation of Filler-Gap Processing in Native and Second Language Speakers. *Language Learning* 63: 766–99. [CrossRef]

Davidson, Douglas J., and Peter Indefrey. 2009. An Event-related Potential Study on Changes of Violation and Error Responses during Morphosyntactic Learning. *Journal of Cognitive Neuroscience* 21: 433–46. [CrossRef]

Dawson, George D. 1954. A summation technique for the detection of small evoked potentials. *Electroencephalography and Clinical Neurophysiology* 6: 65–84. [CrossRef]

Faretta-Stutenberg, Mandy, and Kara Morgan-Short. 2018. The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research* 34: 67–101. [CrossRef]

Fodor, Jerry A., Thomas G. Bever, and Merril F. Garrett, eds. 1974. *The Psychology of Language: An Introduction to Psycholinguistics in Generative Grammar*. New York: McGraw-Hill.

Frazier, Lyn. 1987. Sentence processing. A tutorial review. In *Attention and Performance XII*. Edited by Max Coltheart. Hove: Lawrence Erlbaum, pp. 559–86.

Friederici, Anegla D., Steinhauer Karsten, and Pfeifer Erdmut. 2002. Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences* 99: 529–34. [CrossRef]

Friederici, Angela D. 1995. The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language* 50: 259–81. [CrossRef]

Friederici, Angela D. 1999. The Neurobiology of Language Comprehension. In *Language Comprehension: A Biological Perspective*, 2nd ed. Edited by Angela D. Friederici. Berlin/Heidelberg and New York: Springer, pp. 265–304.

Grey, Sarah, Cristina Sanz, Kara Morgan-Short, and Michael T. Ullman. 2017. Bilingual and monolingual adults learning an additional language: ERPs reveal differences in syntactic processing. *Bilingualism: Language and Cognition* 21: 970–94. [CrossRef]

Gunter, Thomas C., Laurie A. Stowe, and Gusbertus Mulder. 1997. When syntax meets semantics. *Psychophysiology* 34: 660–76. [CrossRef]

Hahne, Anja, and Angela D. Friederici. 1999. Electrophysiological Evidence for Two Steps in Syntactic Analysis: Early Automatic and Late Controlled Processes. *Journal of Cognitive Neuroscience* 11: 194–205. [CrossRef]

Kaan, Edith, Anthony Harris, Edward Gibson, and Phillip Holcomb. 2000. The P600 As an Index of Syntactic Integration Difficulty. *Language and Cognitive Processes* 15: 159–201. [CrossRef]

Kaan, Edith. 2007. Event-Related Potentials and Language Processing: A Brief Overview. *Language and Linguistics Compass* 1: 571–91. [CrossRef]

Kimball, John. 1973. Seven Principles of Surface Structure Parsing in Natural Language. *Cognition* 2: 15–47. [CrossRef]

King, Jonathan W., and Marta Kutas. 1995. Who Did What and When? Using Word- And Clause-Level ERPs to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience* 7: 376–95. [CrossRef]

Kluender, Robert, and Marta Kutas. 1993a. Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. *Journal of Cognitive Neuroscience* 5: 196–214. [CrossRef]

Kluender, Robert, and Marta Kutas. 1993b. Subjacency as a Processing Phenomenon. *Language and Cognitive Processes* 8: 573–633. [CrossRef]

Kotz, Sonja A., Phillip J. Holcomb, and Lee Osterhout. 2008. ERPs Reveal Comparable Syntactic Sentence Processing in Native and Non-Native Readers of English. *Acta Psychologica* 128: 514–27. [CrossRef]

Kutas, Marta, and Kara D. Federmeier. 2000. Electrophysiology Reveals Semantic Memory Use in Language Comprehension. *Trends in Cognitive Sciences* 4: 463–70. [CrossRef]

Kutas, Marta, and Kara D. Federmeier. 2011. Thirty Years and Counting: Finding Meaning in The N400 Component of The Event-Related Brain Potential (ERP). *Annual Review of Psychology* 62: 621–47. [CrossRef]

Kutas, Marta, and Steven A. Hillyard. 1980. Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science* 207: 203–5. [CrossRef]

Kutas, Marta, and Steven A. Hillyard. 1983. Event-Related Brain Potentials to Grammatical Errors and Semantic Anomalies. *Memory & Cognition* 11: 539–50. [CrossRef]

Lelekov, Taïssia, Peter F. Dominey, and Luis Garcia-Larrea. 2000. Dissociable ERP Profiles for Processing Rules vs. Instances in A Cognitive Sequencing Task. *Neuroreport* 11: 1129–32. [CrossRef] [PubMed]

Lewis, Ashley Glen, Kristin Lemhöfer, Jan-Mathijs Schoffelen, and Herbert Schriefers. 2016. Gender Agreement Violations Modulate Beta Oscillatory Dynamics During Sentence Comprehension: A Comparison of Second Language Learners and Native Speakers. *Neuropsychologia* 89: 254–72. [CrossRef] [PubMed]

Linck, Jared A., and Daniel J. Weiss. 2015. Can working memory and inhibitory control predict second language learning in the classroom? *SAGE Open* 5: 1–11. [CrossRef]

Morgan-Short, Kara, Cristina Sanz, Karsten Steinhauer, and Michael T. Ullman. 2010. Second Language Acquisition of Gender Agreement in Explicit and Implicit Training Conditions: An Event-Related Potential Study. *Language Learning* 60: 154–93. [CrossRef] [PubMed]

Morgan-Short, Kara, Karsten Steinhauer, Cristina Sanz, and Michael T. Ullman. 2012. Explicit and Implicit Second Language Training Differentially Affect the Achievement of Native-Like Brain Activation Patterns. *Journal of Cognitive Neuroscience* 24: 933–47. [CrossRef] [PubMed]

Morgan-Short, Kara. 2014. Electrophysiological Approaches to Understanding Second Language Acquisition: A Field Reaching Its Potential. *Annual Review of Applied Linguistics* 34: 15–36. [CrossRef]

Morgan-Short, Kara. 2020. Insights into the neural mechanisms of becoming bilingual: A brief synthesis of second language research with artificial linguistic systems. *Bilingualism: Language and Cognition* 23: 87–91. [CrossRef]

Münte, Thomas F., Mike Matzke, and Sönke Johannes. 1997. Brain Activity Associated with Syntactic Incongruencies in Words and Pseudo-Words. *Journal of Cognitive Neuroscience* 9: 318–29. [CrossRef]

Nickels, Stefanie, and Karsten Steinhauer. 2018. Prosody–Syntax Integration in a Second Language: Contrasting Event-Related Potentials from German and Chinese Learners of English Using Linear Mixed Effect Models. *Second Language Research* 34: 9–37. [CrossRef]

Osterhout, Lee, and Linda A. Mobley. 1995. Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language* 34: 739–73. [CrossRef]

Osterhout, Lee, and Phillip J. Holcomb. 1992. Event-Related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language* 31: 785–806. [CrossRef]

Osterhout, Lee, and Phillip J. Holcomb. 1993. Event-Related Potentials and Syntactic Anomaly: Evidence of Anomaly Detection During the Perception of Continuous Speech. *Language and Cognitive Processes* 8: 413–37. [CrossRef]

Osterhout, Lee, Andrew Poliakov, Kayo Inoue, Judith McLaughlin, Geoffrey Valentine, Ilona Pitkanen, Cheryl Frenck-Mestre, and Julia Hirschensohn. 2008. Second-language learning and changes in the brain. *Journal of Neurolinguistics* 21: 509–21. [CrossRef] [PubMed]

Osterhout, Lee. 1997. On the Brain Response to Syntactic Anomalies: Manipulations of Word Position and Word Class Reveal Individual Differences. *Brain and Language* 59: 494–522. [CrossRef] [PubMed]

Pili-Moss, Diana, Katherine A. Brill-Schuetz, Mandy Faretta-Stutenberg, and Kara Morgan-Short. 2019. Contributions of Declarative and Procedural Memory to Accuracy and Automatization During Second Language Practice. *Bilingualism: Language and Cognition* 23: 639–51. [CrossRef]

Roehm, Dietmar. 2004. Waves and Words: Oscillatory Activity and Language Processing. Unpublished dissertation, Philipps-University Marburg, Marburg, Germany.

Rösler, Frank, Thomas Pechmann, Judith Streb, Brigitte Röder, and Erwin Hennighausen. 1998. Parsing of Sentences in A Language with Varying Word Order: Word-By-Word Variations of Processing Demands Are Revealed by Event-Related Brain Potentials. *Journal of Memory and Language* 38: 150–76. [CrossRef]

Rugg, Michael D. 1990. Event-Related Brain Potentials Dissociate Repetition Effects of High-And Low-Frequency Words. *Memory & Cognition* 18: 367–79. [CrossRef]

Tokowicz, Natasha, and Brian MacWhinney. 2005. Implicit and Explicit Measures of Sensitivity to Violations in Second Language Grammar: An Event-Related Potential Investigation. *Studies in Second Language Acquisition* 27: 173–204. [CrossRef]

Ullman, Michael T. 2001. A Neurocognitive Perspective on Language: The Declarative/Procedural Model. *Nature Reviews Neuroscience* 2: 717–26. [CrossRef]

Van Petten, Cyma, and Marta Kutas. 1990. Interactions Between Sentence Context and Word Frequencyinevent-Related Brainpotentials. *Memory & Cognition* 18: 380–93. [CrossRef]

Van Petten, Cyma, Marta Kutas, Robert Kluender, Mark Mitchiner, and Heather McIsaac. 1991. Fractionating the Word Repetition Effect with Event-Related Potentials. *Journal of Cognitive Neuroscience* 3: 131–50. [CrossRef] [PubMed]

languages

MDPI

Article

Reading Comprehension in French L2/L3 Learners: Does Syntactic Awareness Matter?

Juwairia Sohail [1,*], Kathleen Hipfner-Boucher [1,*], Hélène Deacon [2] and Xi Chen [1]

[1] Ontario Institute for Studies in Education, University of Toronto, Toronto, ON M5S 1V6, Canada
[2] Department of Psychology and Neuroscience, Dalhousie University, Halifax, NS B3H 4R2, Canada
* Correspondence: juwairia.sohail@utoronto.ca (J.S.); khibou59@gmail.com (K.H.-B.)

Abstract: This study examines the contributions of syntactic awareness to reading comprehension, both within and across languages, in third-grade children learning French as a second (L2) or third language (L3). Participants were 72 non-francophone children enrolled in a Canadian French immersion program in which all academic instruction is in French. Children completed measures of reading comprehension, syntactic awareness, word reading, vocabulary, and reading-related control variables in both English and French. Regression analyses examining within-language relations revealed that French syntactic awareness made a significant unique contribution to French reading comprehension after controlling for nonverbal reasoning, language status (French as either L2 or L3), word reading, and vocabulary. Furthermore, French syntactic awareness contributed across languages to English reading comprehension, after accounting for English controls (word reading, vocabulary, syntactic awareness) in addition to nonverbal reasoning and language status. In sharp contrast, measures of English syntactic awareness made no unique contribution to reading comprehension in either English or French after the aforementioned controls. These findings add to theoretical models of reading comprehension by highlighting the importance of syntactic awareness in the language of instruction in supporting bilingual children's reading comprehension.

Keywords: syntactic awareness; reading comprehension; bilinguals

Citation: Sohail, Juwairia, Kathleen Hipfner-Boucher, Hélène Deacon, and Xi Chen. 2022. Reading Comprehension in French L2/L3 Learners: Does Syntactic Awareness Matter? *Languages* 7: 211. https://doi.org/10.3390/languages7030211

Academic Editor: John W. Schwieter

Received: 11 April 2022
Accepted: 13 July 2022
Published: 9 August 2022

1. Introduction

Learning in school hinges on children's ability to understand what they read, making reading comprehension the cornerstone of academic achievement. In order to successfully comprehend text, readers rely on a number of underlying metalinguistic capabilities (Kuo and Anderson 2008). Among these is syntactic awareness, or the capacity to attend to and manipulate the structure of sentences (Durgunoğlu 2002; Nagy 2007). Readers frequently encounter sentences in text that are more syntactically complex than those encountered in oral language (Perfetti et al. 2005); their task is to make sense of them. Syntactic complexity is a key determinant of text difficulty (Stenner and Swartz 2012) and is the feature of text that increases most dramatically across grade levels (Graesser et al. 2011). Accordingly, syntactic awareness impacts the ability to derive meaning from sentences in text (e.g., Tunmer et al. 1987). A growing body of work demonstrates that syntactic awareness plays a role in first language (L1) reading comprehension (e.g., Deacon and Kieffer 2018; Demont and Gombert 1996; Low and Siegel 2005; for a review, see Mackay et al. 2021). Currently, however, there is limited evidence attesting to the role of syntactic awareness in reading comprehension among children who are learning to read in a language that differs from their L1. Our study reports findings that contribute to this evidence base.

It is important to investigate the skills related to reading comprehension in children acquiring literacy in a language other than their L1 given their increasing numbers worldwide (Paradis et al. 2011). Included in this group of learners are the growing number of children enrolled in French immersion programs across Canada. French immersion is a means of promoting proficiency in English and French, Canada's two official languages. French immersion classrooms reflect the cultural and linguistic diversity that is characteristic of the Canadian population. In large urban centres such as the one in which our study was conducted, they are made up of majority English-speaking children, for whom French is their L2, and minority language children from a range of L1 backgrounds, for whom English is their L2 and French their third language (L3).

In early French immersion programs, children receive all literacy instruction in French beginning in senior kindergarten or Grade 1. English language arts are typically introduced in the third or fourth grade; however, the majority of instruction continues to be offered in French (e.g., Ontario Ministry of Education 2013). Exposure to English during the early grades in French immersion occurs in informal contexts, for example, in interactions with peers on the school playground, with members of the broader community outside of school, and at home. Students must therefore acquire early language and literacy skills simultaneously through formal instruction in French, the L2/L3, with very limited, if any, exposure to French beyond the classroom.

A significant body of research reveals the influence that metalinguistic skills—most notably, phonological and morphological awareness—acquired in one language have on learning to read in another language (for a review, see Hipfner-Boucher and Chen 2016). This is known as 'transfer' and is key to explaining French immersion children's rapid progress in learning to read in English following initial instruction in French. Consequently, in examining the role of syntactic awareness on reading comprehension in our sample of French immersion children, we examined both within-language effects (i.e., French syntactic awareness on French reading comprehension, English syntactic awareness on English reading comprehension) and cross-language effects (i.e., French syntactic awareness on English reading comprehension, English syntactic awareness on French reading comprehension). We investigated these effects in Grade 3, a point at which children are beginning to tackle texts that are relatively complex from the perspective of syntax. Understanding the within and cross-language contributions of syntactic awareness to French and English reading comprehension among children in French immersion may serve to inform comprehensive models of reading comprehension, as well as educational policy, teacher training, and instructional practices intended to promote student reading outcomes in immersion settings.

1.1. Syntactic Awareness and Reading Comprehension within L1 and L2

Theories of reading comprehension point to the role of syntax in comprehending text (Kintsch 1998; Perfetti et al. 2005; Perfetti and Stafura 2014; Scarborough 2002). Within the Reading Systems Framework (Perfetti and Stafura 2014), syntax is hypothesized to play multiple roles. It is essential to the process of sentence parsing, a process shown in empirical studies to support reading comprehension (Deacon and Kieffer 2018; Gaux and Gombert 1999; Kuhn and Stahl 2003; RAND Reading Study Group 2002). Indeed, to derive meaning from text at the sentence level, readers may first parse complex sentences into more manageable "chunks" (e.g., noun or verb phrases) for processing purposes before recombining them to reconstitute the whole (Perfetti and Stafura 2014). This process is enabled by awareness of sentence structure (Deacon and Kieffer 2018). Consider, for example, this sentence taken from Chapter 4 of White's (1952, p. 28) children's novel, *Charlotte's Web*: "If there were something that was less than nothing, then nothing would not be nothing, it would be something—even though it's just a very little bit of something". Parsing this sentence into its four constituent clauses allows readers to derive its meaning more efficiently than attempting to process its 29 words as a whole. The reader may then recruit higher-order knowledge and skills (e.g., background knowledge, knowledge of story structure, inferencing skills) to integrate information within and across sentences to form a coherent and cohesive model of the passage in its entirety (Perfetti et al. 2005).

Because syntactic awareness enables the reader to parse a sentence into more 'digestible' chunks, it may also reduce the demands of complex sentences on working memory

(Perfetti et al. 2005). Indeed, studies on children's sentence comprehension and processing have found that increases in syntactic complexity are associated with greater working memory involvement (Caplan and Waters 1999; Montgomery et al. 2008). Going back to the complex sentence above, it is less cognitively burdensome to retain its constituent clauses in memory while processing meaning locally than it is to attempt to retain and process the sentence as a whole. Although they are related, syntactic skills and working memory are distinct, and both have been found to explain unique variances in reading comprehension (Poulsen et al. 2022).

Lastly, syntactic awareness may play a role in comprehension monitoring (Bowey 1986; Gaux and Gombert 1999; Tunmer et al. 1987). Comprehension monitoring enables skilled readers to detect breakdowns in their understanding and to apply reparative strategies, such as rereading, questioning and clarifying, or reading ahead, to restore coherence in their understanding (Wagoner 1983; Wassenburg et al. 2015; Yang 2006). Children's awareness of syntactic structures may facilitate their recognition of comprehension errors, such as syntactic violations. For example, if a reader misreads the word *than* as *thank* in the example above, awareness of syntax could alert them to the syntactic violation that occurred wherein a preposition was replaced by a verb. Recent neuroimaging studies have demonstrated greater neural activation in certain regions of the brain (e.g., Broca's area) when readers are confronted with syntactic ambiguities (e.g., ambiguity about the main verb in a sentence, multiple syntactic interpretations) or conflict between syntactic and semantic information in text (see Baker et al. 2014 for a review). Although neural activation does not necessarily indicate awareness, the findings nevertheless suggest a link between syntactic challenges and the processing demands needed to comprehend a sentence.

Many empirical studies have demonstrated that syntactic awareness is related to L1 reading comprehension among English-speaking children (e.g., Deacon and Kieffer 2018; Low and Siegel 2005; Mokhtari and Neiderhauser 2013; Muter et al. 2004; Nation and Snowling 2000). This relation holds after controlling for a range of variables, including nonverbal reasoning, vocabulary, phonological awareness, decoding and word reading, short-term memory, and working memory. The effect has also been demonstrated to be robust in both concurrent and longitudinal analyses. For example, Deacon and Kieffer (2018) assessed syntactic awareness in English-speaking monolinguals in Grades 3 and 4. After controlling for nonverbal reasoning, phonological awareness, vocabulary, word reading, and morphological awareness, syntactic awareness uniquely predicted reading comprehension in each grade. Additionally, syntactic awareness measured in Grade 3 predicted gains in reading comprehension the following year after controlling for autoregressive effects (Deacon and Kieffer 2018).

Studies have also demonstrated the relation between syntactic awareness and reading comprehension in monolingual French children (Demont and Gombert 1996; Gaux and Gombert 1999; Plaza and Cohen 2003). For example, Plaza and Cohen (2003) assessed the contribution of syntactic awareness to a composite written language measure comprising various reading (and spelling) tasks, including sentence-level reading comprehension among Grade 1 students. They found that syntactic awareness was a significant predictor of the outcome variable when phonological awareness, short-term memory, and rapid naming speed were controlled for. In a study by Gaux and Gombert (1999), sixth-grade French speakers were assessed on syntactic awareness. After controlling for verbal and nonverbal reasoning, short-term memory, vocabulary, and listening comprehension, regression and path analyses showed that syntactic awareness was a direct contributor to reading comprehension.

On the other hand, studies that have examined French L2 children have revealed inconsistent results (Lefrançois and Armand 2003; Simard et al. 2014; Sohail et al. 2022). In a study involving 10-year-old Portuguese L1 children schooled in French in a predominantly French-speaking environment, Simard et al. (2014) found that syntactic awareness explained significant variance in L2 reading comprehension in a regression model that included age, vocabulary, syntactic knowledge, and phonological memory as simultaneous predictors. In contrast, two other studies with French L2 learners found evidence of correlations between syntactic awareness and reading comprehension in French among children aged 6 to 11 years that did not remain after controls (Lefrançois and Armand 2003; Sohail et al. 2022). For example, Sohail et al. (2022) observed that French syntactic awareness was correlated with reading comprehension in French among English-speaking children enrolled in Grade 1 of French immersion, relations that did not remain after cognitive ability, vocabulary, and word reading controls. Children's degree of exposure to French may account for the inconsistency in findings between these studies. Although the Portuguese speakers in Simard et al. (2014) were French language learners, they were nonetheless exposed to French as the majority language since birth. In contrast, the participants in Sohail et al. (2022) and Lefrançois and Armand (2003) were in the early stages of acquisition. It is possible that syntactic awareness only begins to contribute to L2 reading comprehension as children develop greater L2 oral proficiency.

1.2. Cross-Language Transfer of Syntactic Awareness to Reading Comprehension

There is a theoretical basis for hypothesizing the cross-language transfer of syntactic awareness (Cummins 1979, 1980; Geva and Ryan 1993; Koda 2008). Cummins (1979, 1980) proposed that bilinguals' language and literacy skills in the L1 and L2 are interdependent. According to his *Linguistic Interdependence Hypothesis*, competencies acquired in one language can transfer to another insofar as there is adequate exposure and motivation to learn the language (Cummins 1981). Stemming from this framework, Geva and Ryan (1993) proposed the *Common Underlying Cognitive Processes* framework, arguing that common cognitive processes underlie parallel skills in the L1 and L2. In the more recent *Transfer Facilitation Model*, transfer is defined as the "automatic activation of well-established first-language competencies, triggered by second-language input" (Koda 2008, p. 78). According to Koda (2008), L1 transfer occurs continually throughout L2 development as a product of increased L2 input. The linguistic distance between L1 and L2 is believed to influence the relative ease with which transfer occurs. Languages that share many linguistic features presumably require fewer adjustments to transferred L1 competencies, and require less L2 input, before L2 metalinguistic awareness and related reading sub-skills can develop (Koda 2008). While this model describes the transfer of skills from L1 to L2 only, there is also evidence that transfer can occur in the reverse direction (Chung et al. 2019). According to the *Interactive Transfer Framework* (Chung et al. 2019), children's relative proficiency in the L1 and L2 influences the direction of transfer, with transfer occurring from the more proficient language to the less proficient one. In addition to proficiency, language complexity has been identified as a factor that influences transfer, in that bilinguals' heightened sensitivity to complex linguistic structures encountered in one language may facilitate transfer to another language (Chung et al. 2019). Although a growing body of evidence demonstrates that a range of L1 metalinguistic skills predicts L2 reading (see Chung et al. 2019 for a review), there remains a gap in our understanding of whether, how, and in what direction syntactic awareness contributes to reading skills across languages due to inconsistent findings in the literature.

While few in number, studies examining the effect of L1 syntactic awareness on L2 reading comprehension have demonstrated evidence of transfer (Siu and Ho 2015, 2020; Sohail et al. 2022; Swanson et al. 2008). Among these is Sohail et al.'s (2022) study involving Grade 1 students in French immersion. Syntactic awareness in English, the children's L1, significantly contributed to French reading comprehension after controlling for age, nonverbal reasoning, phonological awareness, vocabulary, word reading, and within-language syntactic awareness. The authors suggested that these early French L2 learners were able to tap into syntactic awareness in their stronger language, English, in order to bolster reading comprehension in their developing L2. These results are corroborated by two studies involving Chinese-English bilinguals whose two languages are typologically distant (Siu and Ho 2015, 2020). In their study of English L2 learners in Grades 1 and 3, Siu and Ho (2015) found that Chinese syntactic awareness contributed to English reading comprehension after controlling for age, nonverbal reasoning, vocabulary, word reading, and working memory (Siu and Ho 2015). This effect was also observed longitudinally in a recent investigation by the same authors: Chinese syntactic awareness, measured in the first and third grades, predicted English reading comprehension one year later (Siu and Ho 2020).

In contrast, studies examining the effect of syntactic awareness on reading comprehension among Latinx children in the USA found no evidence of transfer (Leider et al. 2013; Proctor et al. 2012; Swanson et al. 2008). In Swanson et al.'s (2008) study with third graders, L2 English syntactic awareness predicted L1 Spanish reading comprehension after controlling for phonological awareness and vocabulary in both languages, as well as Spanish syntactic awareness. Although Spanish syntactic awareness predicted English reading comprehension after controlling for Spanish phonological awareness and vocabulary only, the model was no longer significant when English controls were added (Swanson et al. 2008). The authors attributed their findings to instructional factors, arguing that because all academic instruction was in English, the L2, L1 Spanish skills had less opportunity to develop. L1 predictors, such as syntactic awareness, may therefore have been unable to contribute to L2 reading over and above the effects of L2 predictors. At the same time, two other studies involving Spanish-speaking children in Grades 2 to 5 in English-medium schools found either no significant cross-language effects (Proctor et al. 2012) or significant but negative cross-language effects (Leider et al. 2013). Taken together, these studies suggest a need for additional research examining the effect of syntactic awareness on reading comprehension among bilingual children to clarify the nature of transfer and the specific factors that facilitate it.

There are many similarities between English and French syntax that may facilitate cross-language transfer. For example, both languages follow a subject-verb-object word order (e.g., *I eat chocolate* and *Je mange du chocolat*), and in both languages, a noun is preceded by a definite or indefinite article (e.g., *a cat* or *the boy* and *un chat* or *le garçon*). However, there are also notable differences between the syntactic structures of the two languages. For example, word order in French switches to subject-object-verb when the object is a pronoun (e.g., *Je peins la maison* 'I am painting the house' versus *Je la peins* 'I [it] am painting'). While adjectives precede nouns in English (e.g., *a small bird*, *a red bird*) and may precede nouns sometimes in French (*un petit oiseau*), the reverse order also exists in French (e.g., *un oiseau rouge*). Thus, French allows for somewhat more variation in word order than English. From a theoretical perspective, the presence of structural similarities and differences between their languages may heighten English-French bilingual children's awareness of syntax. This hypothesis aligns with Kuo and Anderson's (2010, 2012) *structural sensitivity theory*, which contends that exposure to two languages increases the salience of linguistic structures among bilinguals and allows them to form more abstract representations of language. This bilingual advantage has been demonstrated in studies on children's syntactic awareness (e.g., Siu and Ho 2022).

1.3. The Present Study

In the present study, we examined the within- and cross-language relations between syntactic awareness and reading comprehension in children who were in their third year of French immersion. We examined Grade 3 students because it is around this age that oral language skills become increasingly important in facilitating children's reading comprehension (e.g., Gough and Tunmer 1986), as children are beginning to encounter complex sentences in the books they are expected to read. With almost three years of instruction in French, immersion students' experience with a variety of syntactic structures and rules is likely to have promoted robust development of syntactic awareness.

Our study addressed two research questions. Our first question examined whether syntactic awareness is related to reading comprehension *within* each of the children's two languages. Specifically, we asked if French syntactic awareness predicts French reading comprehension and if English syntactic awareness predicts English reading comprehension.

Given that our participants were in their third year of French immersion and had developed a degree of oral proficiency in French, we expected to find a link between French syntactic awareness and French reading comprehension. Although English was not the language of academic instruction, our participants were immersed in English outside of the classroom. We, therefore, hypothesized that a within-language relation between syntactic awareness and reading comprehension in English was likely.

Our second question examined cross-language relations between syntactic awareness measured in English or French and reading comprehension in the other language. Specifically, we sought to determine whether English syntactic awareness contributes to French reading comprehension and whether French syntactic awareness contributes to English reading comprehension through cross-language transfer. Findings from previous studies on the role of L1 syntactic awareness in L2 reading in bilingual populations have been mixed. However, there is a theoretical rationale for expecting the transfer of L1 syntactic awareness to L2 reading (e.g., Koda 2008). Likewise, given theoretical and empirical support for cross-language transfer from the L2 to L1 (Chung et al. 2019), we also expected to find transfer from French syntactic awareness to English reading comprehension. We based this prediction on the assumption that French immersion children in Grade 3 have developed a sufficient level of syntactic awareness in French to make a positive contribution to reading comprehension across languages. Notably, few studies on syntactic awareness have examined whether transfer occurs from L2 to L1. Clarifying the direction of transfer can advance our understanding of how emerging bilinguals leverage syntactic awareness across languages to facilitate reading comprehension, particularly in the context of instruction that is L2-based.

2. Materials and Methods

2.1. Participants

Seventy-six third-grade students (53% females, Mage = 8 years, 2 months, *SD* = 4.3 months) attending a Canadian French immersion school participated in the study in the spring of Grade 3. All children had received academic instruction solely in French since Grade 1. Most of the participants (87%) were born in Canada; among the ten participants born outside of Canada, the mean age of immigration was two years. Regarding language background, no child spoke French as the L1. English was reported as the L1 for 58% (n = 44) of the participants (i.e., they had little to no exposure to a language other than English at home). For these children, French was the L2. The remaining 32 participants were exposed to a language other than English most often in the home (i.e., >50% of the time). This group was linguistically diverse (n = 17 for Russian, n = 4 for Chinese, n = 3 for Spanish and Hebrew, n = 2 for Serbian, and n = 1 each for Azerbaijani, Hungarian, and Korean). For these children, English was the L2 and French the L3. Background information collected at an earlier time point (Grade 1) within the context of a larger study indicated that 92% of mothers in the sample were at least college- or university-level graduates, and 72% of parents read daily with their children.

2.2. Measures

Participants were assessed on word reading, receptive vocabulary, syntactic awareness, and reading comprehension skills through parallel measures in English and French. All participants also completed a nonverbal reasoning measure. Each of the measures is described in detail below.

Nonverbal reasoning. Children's nonverbal reasoning was assessed in first grade using all four subtests of the Matrix Analogies Test (MAT; Naglieri 1985): Pattern Completion, Reasoning by Analogy, Serial Reasoning, and Spatial Visualization. Each subtest consisted of 16 items involving standard progressive matrices that were to be completed using one of the six choices provided for each item. Testing for each subtest continued until four consecutive errors were committed. The number of correct responses on all subtests was summed to produce a raw score for the MAT. Cronbach's reliability alpha was 0.94 for this task.

Word reading. English word reading was assessed using the letter and word identification subtest of the Woodcock Johnson-III Test of Achievement (WJ-III; Woodcock et al. 2001). Participants were required to point to correct answers or read visually presented letters and words. There were 76 test items on this task and a discontinue rule of six consecutive errors within a single page. The total number of correct responses represented the raw score. For the English task, Cronbach's reliability alpha was 0.93.

French word reading was assessed via an experimental task designed by Au Yeung et al. (2015). The task consisted of 120 items of increasing difficulty. If a child produced fewer than five correct responses within a set of eight words, testing was discontinued. A raw score was obtained by summing the number of words that had been read correctly. Cronbach's reliability alpha for the French word reading task was 0.98.

Receptive vocabulary. Form A of the Peabody Picture Vocabulary Test-IV (PPVT-IV; Dunn and Dunn 2007) was used to assess children's receptive vocabulary in English. Participants were asked to point or verbally identify one of four pictures that best matched an aurally presented word. The PPVT contained two training items and 228 test items, divided into 19 sets of 12 items each, with a discontinuation rule of eight errors within a set. A raw score was obtained for each participant by subtracting the total number of errors from the ceiling item. Cronbach's reliability alpha was 0.95 for this measure.

French receptive vocabulary was assessed using Form A of the Échelle de Vocabulaire en Images Peabody (EVIP; Dunn et al. 1993). The test consisted of 170 items, and its task requirements paralleled the PPVT-IV. Note that, as the test is originally normed for native French speakers, we opted to test our participants from the first item rather than at the age-based starting point. The task was discontinued upon reaching six errors within eight consecutive items. The raw score for the EVIP was the number of correct responses provided by the participant. For the French vocabulary measure, Cronbach's reliability alpha was 0.97.

Syntactic awareness. Error-correction tasks were created in English and French to assess children's syntactic awareness (see Appendix A for the full tasks). Participants were presented with sentences containing syntactic errors and asked to "fix the sentence up so that it sounds right". All sentences were presented both aurally and visually. A raw score was computed for each language by summing the correct number of responses on each task. Items for the English task were adopted from an existing syntactic awareness measure designed by Deacon and Kieffer (2018). Participants received three practice items and 16 test items (e.g., "*The teacher the story read to the children*" to be corrected to "*The teacher read the story to the children*"). The French error-correction task involved a similar procedure. It included 2 practice items and 18 test items (e.g., "*On va à la maison très vite*" to be corrected to "*On va très vite à la maison*"). Cronbach's reliability alpha was 0.73 for the English measure and 0.64 for the French measure.

Reading comprehension. The comprehension subtests of Form 3 of Level C (48 items) of the Gates-MacGinitie Reading Tests Second Canadian Edition (GMRT-II; MacGinitie and MacGinitie 1992) were used to assess English reading comprehension. It was the only measure administered in a group session, albeit independently, with participants receiving 20 min to read passages and complete as many questions as possible. Each participant's raw score was the total number of correct answers. Cronbach's reliability alpha was 0.93 for this task.

French reading comprehension was tested using a parallel, experimental task. Form 4 of Level C of the Gates-MacGinitie Reading Tests Second Canadian Edition (GMRT-II; MacGinitie and MacGinitie 1992) was translated into French and administered following the same procedure as the English measure. The task consisted of 48 items. Cronbach's alpha was 0.82.

2.3. Procedure

Participants completed the nonverbal reasoning test in the fall of Grade 1. All other measures for the present study were administered in the spring of Grade 3. The order of task administration was randomized across participants. Trained graduate students administered the majority of measures individually to each participant, with the exception of the two reading comprehension tasks, which were administered in a small group setting. English-language tasks were preceded by instructions in English, while French-language tasks were preceded by instructions in both French and English to ensure children's comprehension of task requirements.

3. Results

Table 1 details the mean, range, standard deviation, and internal consistency reliability (i.e., Cronbach's alpha) for each of the measures. One univariate outlier was identified and removed from the sample. The English L1 children scored significantly above their English L2 peers on English receptive vocabulary ($p = 0.037$). However, the mean standard score for the English L2 group on this measure was 100.75, which is the expected mean score of the monolingual norm. *T*-tests also demonstrated that the two groups did not differ in performance on receptive vocabulary in French or on measures of syntactic awareness, word reading, and reading comprehension in French or English (all $p > 0.264$). Furthermore, a Box's *M* test detected no significant difference in variance-covariance patterns between the language groups on all measures (Box's $M = 47.28$, $F = 0.906$, $p = 0.652$). Given that the English L2 children were proficient in English and the two groups performed similarly on most measures, the two groups were combined in all remaining analyses.

Table 1. Means, Standard Derivations (SDs), and Cronbach's Alpha (α) of All Measures (n = 72).

Measure	Min	Max	*M*	*SD*	α
Nonverbal Reasoning	1	48	19.76	11.87	0.94
French Word Reading	37	119	83.47	21.89	0.98
French Vocabulary	0	121	80.07	26.20	0.97
French Syntactic Awareness	1	13	5.08	2.65	0.64
French Reading Comprehension	9	45	20.24	7.55	0.82
English Word Reading	35	71	53.27	8.05	0.93
English Vocabulary	92	164	141.28	14.93	0.95
English Syntactic Awareness	2	16	11.16	2.90	0.73
English Reading Comprehension	7	45	25.69	10.21	0.93

Due to observations of skew in the distributions of some variables, hierarchical regression analyses were conducted using the bootstrapping method embedded in SPSS, with a default bootstrap sample size of 1000. The bootstrapping method is a technique for deriving robust estimates for statistics such as significance tests, confidence intervals, and standard errors in the presence of any deviations from normality (Field 2009). A consistent pattern of results was obtained with and without the bootstrapping method; therefore, the results below are reported using raw data.

Bivariate correlations (i.e., Pearson correlations) between all measures are presented in Table 2. Syntactic awareness in each language was significantly correlated with all other variables. With respect to the within-language relations of interest, syntactic awareness and reading comprehension were moderately correlated with one another in French ($r = 0.54$) and in English ($r = 0.49$). Cross-linguistically, English syntactic awareness was moderately correlated with French reading comprehension ($r = 0.35$), and French syntactic awareness was moderately correlated with English reading comprehension ($r = 0.56$). English and French syntactic awareness were also moderately correlated ($r = 0.39$).

Table 2. Pearson Correlations Among All Measures (n = 72).

Measure	1	2	3	4	5	6	7	8
1. Nonverbal Reasoning	-							
2. French Word Reading	0.03	-						
3. French Vocabulary	0.34 **	0.30 **	-					
4. French Syntactic Awareness	0.27 *	0.38 ***	0.40 ***	-				
5. French Reading Comprehension	0.28 *	0.36 ***	0.41 ***	0.45 ***	-			
6. English Word Reading	0.40 ***	0.53 ***	0.28 **	0.33 **	0.41 ***	-		
7. English Vocabulary	0.47 ***	0.21	0.55 ***	0.39 ***	0.36 **	0.43 ***	-	
8. English Syntactic Awareness	0.38 ***	0.40 ***	0.50 ***	0.39 ***	0.35 **	0.48 ***	0.48 ***	-
9. English Reading Comprehension	0.40 ***	0.36 **	0.61 ***	0.56 ***	0.75 ***	0.50 ***	0.62 ***	0.49 ***

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

3.1. Within-Language Effects of Syntactic Awareness on Reading Comprehension

Hierarchical regressions were conducted to test the unique associations between syntactic awareness and reading comprehension after controlling for other known contributors to reading comprehension, including word reading and vocabulary (National Reading Panel 2000). The regression analyses with French and English reading comprehension as the dependent variables are displayed in Tables 3 and 4, respectively.

Table 3. Hierarchical Linear Regression Predicting Grade 3 French Reading Comprehension.

Steps and Predictors		French Reading Comprehension				
		***B* (*SE*)**	***β***	**R^2**	**Adjusted R^2**	**ΔR^2**
1.	Nonverbal reasoning	0.08 (0.07)	0.12	0.08 **	0.06 *	0.08 *
2.	Language status	−0.09 (1.59)	−0.01	0.05	0.06	0.00
3.	French word reading	0.07 (0.04)	0.19	0.20 ***	0.17 ***	0.12 ***
4.	French vocabulary	0.05 (0.04)	0.19	0.26 *	0.22 *	0.06 *
5.	French syntactic awareness	0.73 * (0.34)	0.26 *	0.31 **	0.26 **	0.05 *
6.	English syntactic awareness	0.10 (0.34)	0.04	0.31	0.25	0.00

Note: *B* (*SE*) = unstandardized coefficient and standard error from the final model; *β* = standardized regression coefficient from the final model. Estimates of R^2, Adjusted R^2, and ΔR^2 from each step of the model. * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

As shown in Table 3, nonverbal reasoning, language status (French L2, French L3), French word reading, and French vocabulary were entered in steps one through four of the regression model predicting French reading comprehension. Together, these variables explained 24% of the variance in French reading comprehension. In the fifth step, French syntactic awareness was entered to determine its within-language effect on reading comprehension. French syntactic awareness constituted a significant and unique contributor, explaining approximately 9% of the additional variance. We computed an interaction term as the product of language status and French syntactic awareness and entered this interaction term in the final step. The interaction term was non-significant ($p > 0.363$), indicating that the variance contributed by syntactic awareness to French reading comprehension was not influenced by participants' language status (i.e., French L2/L3). Therefore, the original regression model without the interaction terms is reported.

Table 4. Hierarchical Linear Regression Predicting Grade 3 English Reading Comprehension.

Steps and Predictors		English Reading Comprehension				
		***B* (*SE*)**	***β***	***R*²**	**Adjusted *R*²**	**Δ*R*²**
1.	Nonverbal reasoning	0.10 (0.08)	0.11	0.16 ***	0.15 ***	0.16 ***
2.	Language status	−2.09 (1.86)	−0.10	0.17	0.15	0.01
3.	English word reading	0.21 (0.12)	0.17	0.31 ***	0.28 ***	0.14 ***
4.	English vocabulary	0.18 * (0.07)	0.28 *	0.46 ***	0.42 ***	0.15 ***
5.	English syntactic awareness	0.19 (0.36)	0.06	0.47	0.43	0.02
6.	French syntactic awareness	1.55 *** (0.41)	0.37 ***	0.56 ***	0.52 ***	0.08 ***

Note: B (SE) = unstandardized coefficient and standard error from the final model; β = standardized regression coefficient from the final model. Estimates of R^2, Adjusted R^2, and ΔR^2 from each step of the model. * $p \leq 0.05$; *** $p \leq 0.001$.

The second hierarchical regression model, with English reading comprehension as the outcome variable, paralleled the first model and is presented in Table 4. Nonverbal reasoning, language status, English word reading, and English vocabulary comprised the first four steps. Together, these variables accounted for 43% of the explained variance. English syntactic awareness was entered into the model next. It explained no unique variance to the outcome variable. An interaction term between language status and English syntactic awareness was not significant ($p > 0.162$) when it was entered in the final step. As such, the model without the interaction is reported.

Commonality analyses were conducted to explore the different patterns that emerged in each language regarding the unique contributions of syntactic awareness to reading comprehension. A commonality analysis supplements regression analyses by decomposing the regression effect into its unique and common effects (Nimon and Reio 2011), making it possible to determine the unique and shared contributions of syntactic awareness to reading comprehension in each language. A commonality analysis was performed to examine the possible combinations of predictor variables—nonverbal reasoning as well as word reading, vocabulary, and syntactic awareness in French—and the percentage of the total variance in French reading comprehension they explain. In predicting French reading comprehension, 41.08% of the regression effect was uniquely explained by these four predictors. Among these, French syntactic awareness explained the largest proportion, contributing a unique 16.30% to the model. Its common effect with the other variables was smaller, sharing 11.96%, 8.26%, and 4.22% of the variance, respectively, with each of French word reading, vocabulary, and nonverbal reasoning. The common effect of the three French predictors—syntactic awareness, word reading, and vocabulary—together was 12.84%. The common effect of all four variables was 4.31%.

We also performed a commonality analysis examining English reading comprehension. The results indicated that 35.42% of the regression effect was uniquely explained by the three English predictors, with the largest proportion (24.35%) accounted for by English vocabulary. English syntactic awareness contributed a unique 3.59% to the model. Among the common effects, English syntactic awareness shared 8.75%, 4.68%, and 0.37% of the variance, respectively, with English vocabulary, English word reading, and nonverbal reasoning. Together, the three English variables—syntactic awareness, vocabulary, and word reading—contributed a common effect of 11.37%. With nonverbal reasoning added, the four variables had an additional common effect of 15.98%. The large amounts of variance shared with co-predictors might explain why, despite its strong correlation with the outcome variable, syntactic awareness did not emerge as a unique within-language contributor to reading comprehension in English.

3.2. Cross-Language Effects of Syntactic Awareness on Reading Comprehension

To determine whether syntactic awareness contributes to reading comprehension across languages, the cross-language syntactic awareness variable was entered as a sixth and final step into each of the aforementioned hierarchical regression models (see Tables 3 and 4). It therefore followed the control variables of nonverbal reasoning, language status, and within-language word reading, vocabulary, and syntactic awareness. Partialling out the effects of within-language syntactic awareness was important in determining the unique contribution, if any, of syntactic awareness measured in the other language.

Results for the model examining French reading comprehension showed that English syntactic awareness made no additional contribution to French reading comprehension, over and above the variance explained by the controls in French. On the other hand, French syntactic awareness, entered as the final variable in the regression model, contributed a significant 9% of unique variance to the overall model of English reading comprehension. With French syntactic awareness included, the regression model explained a total of 53% of the variance in English reading comprehension.

Again, we tested the interaction between language status and cross-language syntactic awareness in each regression model. The interaction term was not significant in either model (all $p > 0.254$). Since the relation between syntactic awareness and reading comprehension did not vary by participants' language status, the models without the interaction terms are reported.

4. Discussion

This study focused on the effects of syntactic awareness on reading comprehension among children in French immersion who were either majority English-speakers, for whom French is the L2, or minority language children from a range of L1 backgrounds, for whom English is the L2 and French the L3. We investigated whether the relation between syntactic awareness and reading comprehension exists within each of English and French, as well as across the two languages. Regarding within-language effects, our results indicated that French syntactic awareness made a significant contribution to children's reading comprehension in third grade after controlling for nonverbal reasoning, word reading, and vocabulary. Across languages, French syntactic awareness also emerged as a significant contributor to English reading comprehension, above and beyond the variance explained by controls and English syntactic awareness. English syntactic awareness, on the other hand, did not contribute significantly to English or French reading comprehension beyond relevant controls.

The finding that French syntactic awareness contributes to within-language reading comprehension in Grade 3 among L2 and L3 learners of French is consistent with the results of studies involving native francophone children in Grades 1 and 6 (Gaux and Gombert 1999; Plaza and Cohen 2003) and extends them to third-grade learners of French as an additional language in an immersion setting. The finding is also consistent with a study that involved Portuguese L1 fifth graders learning French as an L2 in a majority French-speaking environment (Simard et al. 2014). The present study extends the Simard et al. findings to a French immersion population for whom exposure to French was restricted to the classroom. These results suggest that, at least for third-grade French immersion students, French language instruction promotes a level of proficiency in French syntactic awareness that is sufficient to support within-language reading comprehension.

From a theoretical perspective, our French-language findings align with expectations regarding the role of syntax in reading. As part of the linguistic system, syntax is directly involved in comprehension processes (Perfetti and Stafura 2014). Awareness of word order is theorized to be critical to readers' ability to parse complex sentences into smaller, more manageable "chunks", which are then recombined to form text-level representations (Deacon and Kieffer 2018; Perfetti and Stafura 2014). The ability to parse complex sentences may also reduce the cognitive load associated with reading, increasing a text's comprehensibility and the reader's ability to monitor comprehension (Perfetti et al. 2005; Tunmer et al. 1987). At the same time, our finding that French syntactic awareness predicted within-language reading comprehension diverges from previous studies involving non-native French speakers that have found no significant relationship between syntactic awareness and reading comprehension in French. These include Lefrançois and Armand (2003) and Sohail et al. (2022), who examined 11-year-olds and 6-year-olds, respectively. In both of these previous studies, however, children were non-francophones in the early stages of French acquisition who had been enrolled in French language programs for under 10 months. In contrast, the French L2/L3 third graders in the present study and the French L2 fifth graders in Simard et al. (2014) were children who had been attending school in French over a period of years and can be expected to have achieved a greater degree of French language proficiency. We argue that this relatively advanced proficiency—both in French oral language and in French reading—enabled these readers to draw on syntactic awareness developed in French to facilitate text comprehension. Our findings suggest that it may take a certain level of proficiency and experience with a language, such as experience with different sentence structures, before syntactic awareness emerges as a significant contributor to reading comprehension. Future studies using a longitudinal design are needed to reveal developmental changes in the relation between syntactic awareness and reading comprehension at varying levels of language and reading proficiency.

Turning to English reading comprehension, the absence of within-language effects of syntactic awareness on reading comprehension was unexpected. Although significantly correlated, we found that English syntactic awareness failed to account for variance in English reading comprehension once nonverbal reasoning, word reading, and vocabulary were accounted for in the regression model. Our result differs from previous findings of significant within-language relations reported in studies involving English monolingual speakers of similar ages (e.g., Deacon and Kieffer 2018) and among L2 learners of English (e.g., Farnia and Geva 2013; Lesaux et al. 2006, 2007; Low and Siegel 2005; Tong et al. 2021). Instructional factors provide a possible explanation for the discrepant results; in the early elementary years, French language and literacy skills are a primary focus of instruction in French immersion, with instruction delivered entirely in French. English language arts (i.e., instruction in English reading, writing, oral communication, and media literacy) are not introduced until a later grade, meaning that English proficiency is developed in less formal—and less linguistically complex—interactions outside of school (Cummins 2008). One can expect, then, that the children had limited exposure to the more linguistically challenging (i.e., syntactically complex) English featured in the reading materials typically introduced in school and in the measure we used to assess reading comprehension. We argue that, because of limited experience with syntactically complex text in English, the result of French-only instruction, children may not yet have learned to rely on English syntactic cues available to them when reading in that language.

Notably, English syntactic awareness was observed to have some effect on English reading comprehension in combination with other predictors, despite the fact that it did not emerge as a unique predictor. It is plausible that, due to the lack of formal English language arts instruction, English syntactic awareness in these students was undifferentiated from other within-language skills, such as word reading and vocabulary, with which it shared large amounts of variance in the model predicting English reading comprehension. The argument that language skills are largely undifferentiated in the early stages of acquisition is supported by empirical evidence (Foorman et al. 2015). As a result, its unique contribution was very small (3%). This contrasts strikingly with the results of the commonality analysis in French, which demonstrated a much larger unique contribution of French syntactic awareness (16%) to French reading comprehension.

As for cross-language transfer, results revealed a significant crossover effect of French syntactic awareness on English reading comprehension after controlling for cognitive and within-language predictors. Our findings reveal that, in Grade 3, children in French immersion rely more heavily on French syntactic awareness when reading for meaning in English than on the parallel skill in English. This may seem surprising, since English is the stronger of the children's two languages. Indeed, we note the children's higher mean performance on our measure of English versus French syntactic awareness (i.e., 69.8% and 28.2% correct responses, respectively), suggesting stronger awareness of syntax in English than in French. At the same time, English syntactic awareness in Grade 3 did not predict French reading comprehension beyond relevant controls that included French predictors. This finding corroborates that of Swanson et al. (2008), who found no evidence of cross-language transfer of L1 (Spanish) syntactic awareness to L2 (English) reading comprehension among Grade 3 Spanish-English bilinguals, but contrasts with the prior finding of transfer of English (L1) syntactic awareness to French (L2) reading comprehension in a study involving French immersion students in Grade 1 (Sohail et al. 2022). Thus, in the present study, we report reverse transfer (i.e., transfer from the weaker to the stronger language) of French syntactic awareness to English reading comprehension in the absence of the contribution from English to French reported in Sohail et al. (2022). This shift in direction may be reflective of a gradual development of L2 syntactic awareness in French immersion students over the elementary school years. Notably, a similar shift in the direction of transfer from L1 to L2 in the earlier primary grades to L2 to L1 in later grades was observed in a longitudinal study on morphological awareness involving French immersion students in Grades 1 to 3 (Deacon et al. 2007). We hypothesize that by the time French immersion students reach Grade 3, they have acquired substantial language and literacy skills in their L2 through instruction. It is therefore plausible that syntactic awareness in French eclipsed the same skill in English in contributing to French reading comprehension.

Overall, our results may best be interpreted through the lens of the RAND Reading Study Group's "heuristic for thinking about reading comprehension" (RAND Reading Study Group 2002, p. 12). The heuristic considers the influence on text processing of the reader's purpose for reading and the context within which reading occurs. Within this framework, the externally imposed requirement to read academic texts in French in school for the purpose of demonstrating understanding is expected to require the reader to deploy higher-order, strategic processing skills, including syntactic awareness, to satisfy task demands. This contrasts sharply with the degree of processing that may be deployed when reading for an internally generated purpose, such as reading a self-selected book for pleasure. Reading for pleasure outside of school is an activity that children in French immersion would likely choose to do in English since reading in French remains an arduous task, even in Grade 3. We speculate that when the children found themselves in the English testing situation, in which their purpose for reading was externally imposed, children drew on skills and knowledge acquired in French within the familiar context of school to assist them in meeting task demands that were largely inconsistent with their experience of reading in English.

The current findings, together with the results reported by Deacon et al. (2007) and Sohail et al. (2022), lead us to further speculate that cross-language transfer of metalinguistic skills may initially be enabled by reader-level factors such as language proficiency (specifically, the discrepancy in levels of L1 and L2 proficiency), but that it gradually comes under the influence of contextual factors (purpose for reading, motivation for reading, expected consequences of reading) as L2 proficiency increases as a result of instruction. At this point, the direction of transfer may shift. Research following French immersion children into the middle elementary grades, the point at which they begin formal instruction in reading and writing in English, is needed to further explore the patterns and direction of transfer of metalinguistic skill to L1/L2 reading comprehension. We expect that when reading for the purpose of demonstrating understanding in English is required of children, and school instruction offers them the means to achieve this end, cross-language effects from French to English may diminish to the point that comprehension in both languages is largely predicted by within-language effects. That said, our current findings are noteworthy because they support the mutual facilitation of English and French in French immersion and suggest that receiving instruction in French enables children to be more strategic about English reading.

The present findings bear implications for theoretical models of cross-language transfer. They concur with theories of transfer proposed by both Cummins (1979, 1981) and Koda (2008) in that syntactic awareness is a metalinguistic skill that may transfer between languages despite surface-level differences between English and French. Importantly, our findings extend the *Transfer Facilitation Model* by demonstrating cross-language transfer from L2 to L1 and challenging its unidirectional (L1 to L2) hypothesis. In this regard, our results are consistent with the *Interactive Transfer Framework*, which contends that contextual factors (e.g., instruction) influence the direction of transfer in young bilinguals (Chung et al. 2019). It is possible that this pattern of results is specific to educational settings where L1 instruction is absent, as was the case in the French immersion context of this study. Again, future studies are recommended to examine within- and cross-language patterns in the relationship between syntactic awareness and reading comprehension once English language instruction is introduced.

Finally, although our focus was on the transfer of syntactic awareness to reading comprehension, our study provides evidence of transfer at the construct level. Syntactic awareness in English and French were moderately correlated ($r = 0.39$), lending support to the *Common Underlying Cognitive Processes* framework proposed by Geva and Ryan (1993), which argues that common cognitive proficiencies underlie performance on parallel skills across languages and account for the observed correlations among higher-order L1–L2 skills (Geva 2014). Interpreted within this framework, our results indicate that syntactic awareness may represent a language-general process that, once developed in a language, can be leveraged in an additional language to facilitate reading (Geva 2014).

At the same time, our findings provide evidence of language-specific processes. The moderately strong correlations suggest some, but not complete, overlap between the constructs in English and French. Additionally, contributions emerged from French syntactic awareness to English reading comprehension after controlling for English syntactic awareness; however, there were no contributions in the reverse direction. This pattern of transfer suggests that children are also tapping into language-specific awareness of syntactic structure to support reading comprehension. Whereas phonological awareness is a metalinguistic skill that is recognized as a language-general construct (Geva and Ryan 1993), syntactic awareness may be defined by both language-general and language-specific features that influence biliteracy.

Some methodological limitations to the present study must be considered. First, our results are correlational and restricted to a single time point, and, as such, causational relations between syntactic awareness and reading comprehension cannot be determined. Second, the Cronbach's alpha reliability of our French syntactic awareness measure was below ideal levels. Despite this, it emerged as a significant variable in both within- and cross-language analyses, suggesting the robustness of this skill. Third, the participants of this study came from relatively high socioeconomic status (SES), as is common among students in French immersion programs. This narrow SES range may preclude generalizations of results to bilingual students in lower SES contexts, and future studies involving bilingual populations should recruit participants from across a broader range of SES backgrounds.

Nonetheless, our findings offer educational implications for bilingual students, particularly in light of the reading comprehension challenges faced by additional language learners (e.g., Verhoeven 2000; Farnia and Geva 2013). The findings of our study suggest that helping bilingual students gain awareness of sentence structure in their L2 may benefit reading comprehension in each of their languages. As such, it is important to target instruction specifically for the development of syntactic awareness. Within the context of immersion education, such instruction would ideally be embedded in authentic learning tasks that support its communicative goals. Importantly, a limited number of studies have demonstrated the effectiveness of syntactic interventions (see Mackay et al. 2021 for a review). Our study is also relevant for educators, parents, and researchers who are concerned that the French immersion curriculum may negatively impact students' English development. Our findings indicate that, in Grade 3, children draw on French metalinguistic skills to support not only French reading comprehension, but English reading comprehension as well. This is consistent with the program's aim of promoting additive bilingualism (Swain and Lapkin 2005).

Author Contributions: Conceptualization, J.S., K.H.-B., H.D. and X.C.; formal analysis, J.S.; writing—original draft preparation, J.S.; writing—review and editing, K.H.-B., H.D. and X.C.; visualization, J.S.; supervision, X.C.; project administration, X.C.; funding acquisition, J.S. and X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Social Sciences and Humanities Research Council of Canada, grant number 435-2013-1745 and 752-2016-1190.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Social Sciences, Humanities and Education Research Ethics Board at the University of Toronto (protocol code 24374 in 2009).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available because public sharing of the data was not approved by the Research Ethics Board and because participants had not provided consent for their data to be publicly shared at the time of testing.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Practice and Test Items on the English Syntactic Awareness Task.

Sentence Presented	Correct Response
Practice Items	
The girl the door opened.	The girl opened the door.
The boy is playing with he.	The boy is playing with him.
Test Items	
The boy jumped over log.	The boy jumped over the log.
He cleaned him shoes.	He cleaned his shoes.
The boy found the book what you lost.	The boy found the book that you lost.
What the girls are doing?	What are the girls doing?
John gave the crayon for Mary.	John gave the crayon to Mary.
Peter goes sometimes to church.	Peter sometimes goes to church *or* Peter goes to church sometimes.
I wonder how old is he.	I wonder how old he is.
The boy forgot his uniform who plays baseball.	The boy who plays baseball forgot his uniform *or* The boy who forgot his uniform plays baseball.
The teacher the story read to the children.	The teacher read the story to the children.
She will be angry if you will break it.	She will be angry if you break it.
Herself likes to dress Celina.	Celina likes to dress herself.
Found in the ocean are whales.	Whales are found in the ocean.
Interested in music Mary wasn't.	Mary wasn't interested in music
She swims not.	She doesn't swim.

Table A1. *Cont.*

Sentence Presented	Correct Response
Were eaten by the dog the cookies.	The cookies were eaten by the dog.
With Alex the girl is going to the party.	The girl is going to the party with Alex *or* The girl is going with Alex to the party.

Table A2. Practice and Test Items on the French Syntactic Awareness Task.

Sentence Presented	Correct Response
Practice Items	
La fille ouvre le porte.	La fille ouvre la porte.
Mon maman est gentille.	Ma maman est gentille.
Test Items	
Ce crayon est mon.	Ce crayon est le mien.
Il a donné le cadeau à lui.	Il lui a donné le cadeau.
Marie a fait un gâteau puis elle a mangé le.	Marie a fait un gateau, puis elle l'a mangé.
J'ai voyagé sur un train.	J'ai voyagé en train.
En automne, j'aime regarder les rouges feuilles.	En automne, j'aime regarder les feuilles rouges.
Je dois laver mes mains.	Je dois me laver les mains.
Nous allons à le parc ce matin.	Nous allons au parc ce matin.
Elle pas fait son travail.	Elle ne fait pas son travail *or* Elle n'a pas fait son travail.
La jupe est vert.	La jupe est verte.
Le garçon a regardé à mon livre.	Le garçon a regardé mon livre.
L'enfant est triste qui a perdu son chat.	L'enfant qui a perdu son chat est triste.
L'ami de moi a un chien.	Mon ami a un chien.
Quoi avez-vous fait aujourd'hui?	Qu'avez-vous fait aujourd'hui? *or* Qu'est-ce que vous avez fait aujourd'hui?
Tout mangé as-tu?	As-tu tout mangé? *or* Est-ce que tu as tout mangé?
L'école que je vais est loin de la maison.	L'école où je vais est loin de la maison.
Elle a vu le roi et reine.	Elle a vu le roi et la reine.
Je ne sais pas qu'est-ce qu'il veut.	Je ne sais pas ce qu'il veut.
On va à la maison très vite.	On va très vite à la maison.

References

Au Yeung, Karen, Kathleen Hipfner-Boucher, Xi Chen, Adrian Pasquarella, Nadia D'Angelo, and Storm Hélène Deacon. 2015. Development of English and French language and literacy skills in EL1 and EL French immersion students in the early grades. *Reading Research Quarterly* 50: 233–54. [CrossRef]

Baker, Linda, Alisa Zeliger-Kandasamy, and Laura DeWyngaert. 2014. Neuroimaging evidence of comprehension monitoring. *Psychological Topics* 23: 167–87.

Bowey, Judith. 1986. Syntactic awareness and verbal performance from preschool to fifth grade. *Journal of Psycholinguistic Research* 15: 285–308. [CrossRef]

Caplan, David, and Gloria Waters. 1999. Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences* 22: 77–94. [CrossRef]

Chung, Sheila Cira, Xi Chen, and Esther Geva. 2019. Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics* 50: 149–61. [CrossRef]

Cummins, James. 1979. Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research* 49: 222–51. [CrossRef]

Cummins, James. 1980. The construct of language proficiency in bilingual education. In *Current Issues in Bilingual Education*. Edited by James Alatis. Washington, DC: Georgetown University Press, pp. 81–103.

Cummins, James. 1981. The role of primary language development in promoting educational success for language minority students. In *Schooling and Language Minority Students: A Theoretical Framework*. Edited by California State Department of Education. Los Angeles: National Dissemination and Assessment Center, pp. 3–49. [CrossRef]

Cummins, James. 2008. BICS and CALP: Empirical and theoretical status of the distinction. In *Encyclopedia of Language and Education*, 2nd ed. Edited by Brian Street and Nancy Hornberger. New York: Springer, vol. 2, pp. 71–83.

Deacon, Storm Hélène, and Michael Kieffer. 2018. Understanding how syntactic awareness contributes to reading comprehension: Evidence from mediation and longitudinal models. *Journal of Educational Psychology* 110: 72–86. [CrossRef]

Deacon, Storm Hélène, Lesley Wade-Woolley, and John Kirby. 2007. Crossover: The role of morphological awareness in French immersion children's reading. *Developmental Psychology* 43: 732–46. [CrossRef]

Demont, Elisabeth, and Jean Emile Gombert. 1996. Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology* 66: 315–32. [CrossRef]

Dunn, Lloyd, and Leota Dunn. 2007. *Peabody Picture Vocabulary Test*, 4th ed. Bloomington: NCS, Pearson.

Dunn, Lloyd, Claudia Theriault-Whalen, and Leota Dunn. 1993. *Echelle de Vocabulaire en Images Peabody: Adaptation Française du Peabody Picture Vocabulary Test*. Toronto: PsyCan.

Durgunoğlu, Aydin Yücesan. 2002. Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia* 52: 189–204. [CrossRef]

Farnia, Fataneh, and Esther Geva. 2013. Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading* 36: 389–421. [CrossRef]

Field, Andy. 2009. *Discovering Statistics Using SPSS: (and Sex and Drugs and Rock 'n' Roll)*, 3rd ed. London: SAGE Publications, p. 782.

Foorman, Barbara, Sarah Herrera, Yaacov Petscher, Alison Mitchell, and Adrea Truckenmiller. 2015. The structure of oral language and reading and their relation to comprehension in Kindergarten through Grade 2. *Reading and Writing* 28: 655–81. [CrossRef] [PubMed]

Gaux, Christine, and Jean Emile Gombert. 1999. Implicit and explicit syntactic knowledge and reading in pre-adolescents. *British Journal of Developmental Psychology* 17: 169–88. [CrossRef]

Geva, Esther. 2014. Introduction: The cross-language transfer journey—A guide to the perplexed. *Written Language & Literacy* 17: 1–15. [CrossRef]

Geva, Esther, and Ellen Ryan. 1993. Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning* 43: 5–42. [CrossRef]

Gough, Philip, and William Tunmer. 1986. Decoding, reading, and reading disability. *Remedial and Special Education* 7: 6–10. [CrossRef]

Graesser, Arthur, Danielle McNamara, and Jonna Kulikowich. 2011. Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher* 40: 223–34. [CrossRef]

Hipfner-Boucher, Kathleen, and Xi Chen. 2016. Cross-language transfer of metalinguistic and cognitive skills in second language learning. In *Reading in a second language*. Edited by Xi Chen, Vedran Dronjic and Rena Helms-Park. New York: Routledge/Taylor Francis, pp. 99–132.

Kintsch, Walter. 1998. *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.

Koda, Keiko. 2008. Impacts of prior literacy experience on second-language learning to read. In *Learning to Read across Languages: Cross-Linguistic Relationships in First- and Second-Language Literacy Development*. Edited by Keiko Koda and Annette M. Zehler. Mahwah: Routledge, pp. 68–96.

Kuhn, Melanie R., and Steven A. Stahl. 2003. Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology* 95: 3–21. [CrossRef]

Kuo, Li-Jen, and Richard C. Anderson. 2008. Conceptual and methodological issues in comparing metalinguistic awareness across languages. In *Learning to Read across Languages: Cross-Linguistic Relationships in First- and Second-Language Literacy Development*. Edited by Keiko Koda and Annette M. Zehler. Mahwah: Routledge, pp. 39–67.

Kuo, Li-Jen, and Richard C. Anderson. 2010. Beyond cross-language transfer: Reconceptualizing the impact of early bilingualism on phonological awareness. *Scientific Studies of Reading* 14: 365–85. [CrossRef]

Kuo, Li-Jen, and Richard C. Anderson. 2012. Effects of early bilingualism on learning phonological regularities in a new language. *Journal of Experimental Child Psychology* 111: 455–67. [CrossRef]

Lefrançois, Pascale, and Françoise Armand. 2003. The role of phonological and syntactic awareness in second-language reading: The case of Spanish-speaking learners of French. *Reading and Writing* 16: 219–46. [CrossRef]

Leider, Christine Montecillo, Patrick Proctor, Rebecca Silverman, and Jeffrey Harring. 2013. Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Reading and Writing: An Interdisciplinary Journal* 26: 1459–85. [CrossRef]

Lesaux, Nonie, Andre Rupp, and Linda Siegel. 2007. Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology* 99: 821–34. [CrossRef]

Lesaux, Nonie, Orly Lipka, and Linda Siegel. 2006. Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing* 19: 99–131. [CrossRef]

Low, Pauline, and Linda Siegel. 2005. A comparison of the cognitive processes underlying reading comprehension in native English and ESL speakers. *Written Language & Literacy* 8: 131–55. [CrossRef]

MacGinitie, Walter, and Ruth MacGinitie. 1992. *Gates-MacGinitie Reading Tests*, 2nd ed. Toronto: Nelson Canada.

Mackay, Elizabeth, Elise Lynch, Tamara Sorenson-Duncan, and Storm Hélène Deacon. 2021. Informing the science of reading: Students' awareness of sentence-level information for reading comprehension. *Reading Research Quarterly* 56: S221–S230. [CrossRef]

Mokhtari, Kouider, and Dale Neiderhauser. 2013. Vocabulary and syntactic knowledge factors in 5th grade students' reading comprehension. *International Electronic Journal of Elementary Education* 5: 157–70.

Montgomery, James, Beula Magimairaj, and Michelle O'Malley. 2008. Role of working memory in typically developing children's complex sentence comprehension. *Journal of Psycholinguistic Research* 37: 331–54. [CrossRef]

Muter, Valerie, Charles Hulme, Margaret Snowling, and Jim Stevenson. 2004. Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology* 40: 665–81. [CrossRef]

Naglieri, Jack. 1985. *Matrix Analogies Test-Expanded Form*. San Antonio: The Psychological Corporation.

Nagy, William. 2007. Metalinguistic awareness and the vocabulary comprehension connection. In *Vocabulary Acquisition: Implications for Reading Comprehension*. Edited by Richard Wagner, Andrea Muse and Kendra Tannenbaum. New York: Guilford Press, pp. 52–77.

Nation, Kate, and Margaret Snowling. 2000. Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics* 21: 229–41. [CrossRef]

National Reading Panel. 2000. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction. Available online: http://www.nichd.nih.gov/publications/pubs/nrp/documents/report.pdf (accessed on 16 February 2022).

Nimon, Kim, and Thomas Reio Jr. 2011. Regression commonality analysis: A technique for quantitative theory building. *Human Resource Development Review* 10: 329–40. [CrossRef]

Ontario Ministry of Education. 2013. The Ontario Curriculum: French as a Second Language. Available online: https://www.dcp.edu.gov.on.ca/en/curriculum/elementary-fsl (accessed on 23 July 2022).

Paradis, Johanne, Fred Genesee, and Martha Crago. 2011. *Dual Language Development and Disorders: A Handbook on Bilingualism & Second Language Learning*. Baltimore: Paul H. Brookes Publishing.

Perfetti, Charles, and Joseph Stafura. 2014. Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading* 18: 22–37. [CrossRef]

Perfetti, Charles, Nicole Landi, and Jane Oakhill. 2005. The acquisition of reading comprehension skill. In *The Science of Reading: A Handbook*. Edited by Margaret Snowling and Charles Hulme. Oxford: Blackwell Publishing, pp. 227–47. [CrossRef]

Plaza, Monique, and Henri Cohen. 2003. The interaction between phonological processing, syntactic awareness, and naming speed in the reading and spelling performance of first-grade children. *Brain and Cognition* 53: 287–92. [CrossRef]

Poulsen, Mads, Jessie Leigh Nielsen, and Rikke Vang Christensen. 2022. Remembering sentences is not all about memory: Convergent and discriminant validity of syntactic knowledge and its relationship with reading comprehension. *Journal of Child Language* 49: 349–65. [CrossRef]

Proctor, Patrick, Rebecca Silverman, Jeffrey Harring, and Christine Montecillo. 2012. The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing: An Interdisciplinary Journal* 25: 1635–64. [CrossRef]

RAND Reading Study Group. 2002. *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. Santa Monica: RAND Reading Study Group.

Scarborough, Hollis. 2002. Connecting early language and literacy to later (dis)abilities. Evidence, theory and practice. In *Handbook of Early Literacy Research*. Edited by Susan Neuman and David Dickinson. New York: Guilford Press, pp. 97–110.

Simard, Daphnée, Denis Foucambert, and Marie Labelle. 2014. Examining the contribution of metasyntactic ability to reading comprehension among native and non-native speakers of French. *International Journal of Bilingualism* 18: 586–604. [CrossRef]

Siu, Carrey Tik-Sze, and Connie Suk-Han Ho. 2015. Cross-language transfer of syntactic skills and reading comprehension among young Cantonese–English bilingual students. *Reading Research Quarterly* 50: 313–36. [CrossRef]

Siu, Carrey Tik-Sze, and Connie Suk-Han Ho. 2020. A longitudinal investigation of syntactic awareness and reading comprehension in Chinese-English bilingual children. *Learning and Instruction* 67: 101327. [CrossRef]

Siu, Carrey Tik-Sze, and Connie Suk-Han Ho. 2022. Investigating effects of bilingualism on syntactic processing: Testing structural sensitivity theory. *Language Learning* 72: 534–75. [CrossRef]

Sohail, Juwairia, Tamara Sorenson-Duncan, Poh Wee Koh, Storm Hélène Deacon, and Xi Chen. 2022. How syntactic awareness might influence reading comprehension in English-French bilingual children. *Reading and Writing* 35: 1–25. [CrossRef]

Stenner, Jackson, and C. Swartz. 2012. A Causal Rasch Model for Understanding Comprehension in the Context of Reader-Text-Task. Paper presented at Annual Meeting of the National Council on Measurement in Education, Vancouver, BC, Canada, April 12–16.

Swain, Merril, and Sharon Lapkin. 2005. The evolving sociopolitical context of immersion education in Canada: Some implications for program development. *International Journal of Applied Linguistics* 15: 169–86. [CrossRef]

Swanson, Lee, Kelly Rosston, Michael Gerber, and Emily Solari. 2008. Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology* 46: 413–29. [CrossRef] [PubMed]

Tong, Xiuhong, Joyce Lok Yin Kwan, Shelley Xiuli Tong, and Storm Hélène Deacon. 2021. How Chinese-English bilingual fourth graders draw on syntactic awareness in reading comprehension: Within- and cross-language effects. *Reading Research Quarterly* 52: 409–29. [CrossRef]

Tunmer, William, Andrew Nesdale, and A. Douglas Wright. 1987. Syntactic awareness and reading acquisition. *British Journal of Developmental Psychology* 5: 25–34. [CrossRef]

Verhoeven, Ludo. 2000. Components in early second language reading and spelling. *Scientific Studies of Reading* 4: 313–30. [CrossRef]

Wagoner, Shirley. 1983. Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly* 18: 328–46. [CrossRef]

Wassenburg, Stephanie, Lisanne Bos, Björn de Koning, and Menno van der Schoot. 2015. Effects of an inconsistency-detection training aimed at improving comprehension monitoring in primary school children. *Discourse Processes* 52: 463–88. [CrossRef]

White, Elwyn Brooks. 1952. *Charlotte's Web*. New York: Harper & Brothers.
Woodcock, Richard, Kevin McGrew, and Nancy Mather. 2001. *Woodcock-Johnson III Tests of Achievement*. Rolling Meadow: Riverside.
Yang, Yu-Fen. 2006. Reading strategies or comprehension monitoring strategies? *Reading Psychology* 27: 313–43. [CrossRef]

languages

MDPI

Article

The Effect of Code-Switching Experience on the Neural Response Elicited to a Sentential Code Switch

Angélique M. Blackburn [1,*] and Nicole Y. Y. Wicha [2]

1 Department of Psychology & Communication, Texas A&M International University, Laredo, TX 78041, USA
2 Department of Neuroscience, Developmental and Regenerative Biology, The University of Texas at San Antonio, San Antonio, TX 78249, USA; nicole.wicha@utsa.edu
* Correspondence: angelique.blackburn@tamiu.edu; Tel.: +1-(956)-326-2666

Abstract: Switching between languages, or codeswitching, is a cognitive ability that multilinguals can perform with ease. This study investigates whether codeswitching during sentence reading affects early access to meaning, as indexed by the robust brain response called the N400. We hypothesize that the brain prioritizes the meaning of the word during comprehension with codeswitching costs emerging at a different stage of processing. Event-related potentials (ERPs) were recorded while Spanish–English balanced bilinguals (n = 24) read Spanish sentences containing a target noun that could create a semantic violation, codeswitch or both. Self-reported frequency of daily codeswitching was used as a regressor to determine if the cost of reading a switch is modulated by codeswitching experience. A robust N400 to semantic violations was followed by a late positive component (LPC). Codeswitches modulated the left anterior negativity (LAN) and LPC, but not the N400, with codeswitched semantic violations resulting in a sub-additive interaction. Codeswitching experience modulated the LPC, but not the N400. The results suggest that early access to semantic memory during comprehension happens independent of the language in which the words are presented. Codeswitching affects a separate stage of comprehension with switching experience modulating the brain's response to experiencing a language switch.

Keywords: bilingualism; bilingual comprehension; code switching; code-switching positivity; event related potentials; N400; LAN; LPC; semantic processing; sentence reading

Citation: Blackburn, Angélique M., and Nicole Y. Y. Wicha. 2022. The Effect of Code-Switching Experience on the Neural Response Elicited to a Sentential Code Switch. *Languages* 7: 178. https://doi.org/10.3390/languages7030178

Academic Editor: John W. Schwieter

Received: 1 May 2022
Accepted: 27 June 2022
Published: 11 July 2022

1. Introduction

Codeswitching—the act of switching between two languages—is a pervasive phenomenon amongst bilinguals around the world. Bilinguals produce codeswitches for a variety of reasons, from emotional expression to linguistic need (for reviews, see Appel and Muysken 2005; Dewaele 2004; Gonzalez-Reigosa 1973; Gumperz and Hernandez-Chavez 1972; Harris 2004; Marian and Kaushanskaya 2008; Myers-Scotton 1993; Pavlenko 2005, 2008). Importantly, in the communicative exchange, although there are certain constraints around which words can be switched (Dussias 2001; Vaughan-Evans et al. 2020), when a switch will occur is relatively unpredictable for the interlocutor. Nevertheless, in order to comprehend the utterance, they must keep pace by quickly and efficiently integrating the codeswitched word into the ongoing sentence.

Comprehending a codeswitch can incur a cost in the form of a delay in processing. For example, bilinguals read passages more slowly when they contain codeswitches than when they do not, both when reading aloud (Kolers 1966) or reading silently (Macnamara and Kushnir 1971). However, the neurolinguistic process that contributes to this codeswitching cost is just beginning to be understood. This study specifically addresses the interaction between codeswitching and processing words for meaning within a greater sentence context. We have taken advantage of the precise temporal resolution of event related potentials (ERP) to measure the timing and type of effect that switching has on accessing word meaning in the bilingual brain.

Reading a codeswitched word in a sentence context reliably elicits a modulation of two ERP components, a left lateralized negativity (LAN) followed by a late positive component (LPC) (Kutas et al. 2009; Beatty-Martínez et al. 2018; Treffers-Daller et al. 2021; Van Hell et al. 2018; Van Hell et al. 2015). These ERP components are thought to reflect working memory (as per Kluender and Kutas 1993) or syntactic analysis (Friederici 1995; Vaughan-Evans et al. 2020) and sentence integration processing costs, respectively (Friederici 1995; Salmon and Pratt 2002; Leckey and Federmeier 2020; Kuperberg 2007; Van Petten and Luka 2012; Brouwer et al. 2017). Both the LAN and the LPC are larger in amplitude when encountering a codeswitch than when reading a word in the same language as the ongoing sentence. In contrast, processing the meaning of words modulates a different component, called the N400, which is a centro-parietal negativity that peaks around 400ms after the onset of any meaningful or potentially meaningful stimulus (Kutas and Hillyard 1980a, 1980b). It is typically larger in amplitude for words that are not supported by the context compared to more predictable words, and can be thought of as an index of the current state of semantic memory (for a review see Kutas and Federmeier 2011). It has been proposed that codeswitches influence the integration of a word into the building sentence context, as well as recognition of the word (Grainger and Beauvillain 1987, 1988; Grainger and O'regan 1992), but there is mixed evidence as to whether codeswitching can modulate the N400.

In their seminal study, Moreno et al. (2002) tested whether the cost of reading a codeswitch occurs at initial access to meaning (e.g., lexico-semantics), or at later stages in comprehension (e.g., integration of the word into the sentence context). ERPs were recorded while Spanish–English bilinguals read simple English sentences (e.g., "The driver of the speeding car was given a . . . ") that could end with an expected word (ticket), the Spanish translation of the expected word (*multa*), or a plausible English lexical switch (citation). Moreno et al. hypothesized that if codeswitches cause difficulty at the level of processing meaning, reading a codeswitched word would elicit a modulation of N400 amplitude. However, codeswitched words did not modulate the N400. Instead, translations of the expected words (i.e., codeswitches) elicited a negativity with a left anterior distribution characteristic of a LAN. The LAN was followed by a positive-going increase in amplitude between 450–850 ms after stimulus onset, a classic LPC, reflecting the processing cost of the codeswitch. In contrast, lexical switches in the same language elicited a typical N400 modulation over central-parietal electrodes, with larger amplitude for these less expected words compared to the expected continuation.

Other ERP studies of single-word (Ng et al. 2014) and multiple-word (Litcofsky and Van Hell 2017) intra-sentential codeswitching have shown a similar pattern, with a LAN, followed by a LPC (although the LAN is sometimes interpreted as a different component; (Ruigendijk et al. 2016). Notably, the LPC is only present for codeswitches in a sentence context, providing support that the LPC reflects a general processing cost for integrating a switched word into the ongoing sentence context. The LAN appears for isolated words as well as sentences (Molinaro et al. 2011) and might even appear in response to codeswitches in monolinguals (Ruigendijk et al. 2016), reflecting a process that is not specific to comprehending the word. Notably, neither the LAN nor the LPC to a codeswitch appear to be modulated by the amount of context preceding it, such as words appearing earlier or later in a story (Ng et al. 2014). This further suggests that this LAN/LPC complex in response to codeswitches is not sensitive to lexico-semantic information.

There is evidence, however, that these components, namely the LPC, can be sensitive to bilingual specific factors, such as the direction in which the switch occurs (Litcofsky and Van Hell 2017; Fernandez et al. 2019; Liao and Chan 2016), the frequency with which bilinguals experience language switches in daily life (Gosselin and Sabourin 2021; Valdés Kroff et al. 2020), and even the presence of another bilingual during comprehension (Kaan et al. 2020). Moreover, individual variability, such as differences in language proficiency, modulate the codeswitching late positivity (Moreno et al. 2002; Van Der Meij et al. 2011). For example, the more proficient bilinguals are in a language, the smaller the LPC to switches into that language (Ruigendijk et al. 2016; but see Moreno et al. 2002; Van Der Meij et al. 2011). Similarly, the LPC can be modulated asymmetrically depending on the direction of the switch, typically with an increased LPC only when switching into a weaker language (Litcofsky and Van Hell 2017; Fernandez et al. 2019), although the LPC has also been observed in both directions of switch in proficient bilinguals (Liao and Chan 2016), or only when switching into a dominant language (Van Der Meij et al. 2011). The LPC is more robust for codeswitches on nouns than verbs (Ng et al. 2014), is observed even to switched function words (Kaan et al. 2020; Gosselin and Sabourin 2021) and may depend on syntactic structure of the sentence (matrix) language (Vaughan-Evans et al. 2020). Given these findings, the LPC likely reflects processes related to sentence comprehension, more broadly, rather than codeswitching, per se (Yacovone et al. 2021). In turn, we might expect factors such as frequency of experiencing a codeswitch to affect the cost of integrating a codeswitch into the ongoing sentence, as measured by the LPC.

Based on their findings, Moreno et al. (2002) concluded that codeswitches had little to no effect on processing the meaning of an expected continuation (see also review by Moreno et al. 2008). Their N400 results provide strong evidence that sentence context facilitates access not only to related words in the same language, but to the equivalent translations in another language. This is in line with theoretical models of bilingual lexical access that allow for both within- and between-language spread of activation in memory, such as the BIA+ model (Dijkstra and Van Heuven 2002). This is also in line with research suggesting that bilinguals prioritize processing the meaning of words independent of the language in which they are presented (Ng and Wicha 2013).

However, given that Moreno et al. (2002) only used translations of the expected continuation, which were supported by the sentence context, it is less clear if processing of an unexpected word would be equally unaffected by a language switch, e.g., how far does cross-linguistic activation spread. Several studies in the two decades since this seminal study have provided mixed evidence to suggest that codeswitching might actually incur an additional cost in accessing a word from semantic memory under certain circumstances (Proverbio et al. 2004; Gosselin and Sabourin 2021; Liao and Chan 2016; Yacovone et al. 2021; Van Der Meij et al. 2011; Litcofsky and Van Hell 2017; Valdés Kroff et al. 2020; Ruigendijk et al. 2016).

Unlike Moreno et al. (2002), Proverbio et al. (2004) and Van Der Meij et al. (2011) both reported N400 modulations to a codeswitch. However, in both cases there were methodological choices that might explain the modulation. Proverbio et al. (2004) observed overall larger N400 amplitude for sentence final codeswitches (referred to as mixed sentences) in multilingual professional interpreters. This was a main effect collapsed across sensible and senseless (semantically incongruous) sentences. Interestingly, the lack of an interaction between congruity and codeswitching indicates that codeswitching did not additionally modulate the N400 to senseless words. The items were blocked, such that all sentences either contained a codeswitch or did not, and sentences (not just the final word) appeared in both English and Italian in a single block. The significant differences in the language switching experience of the population (professional interpreters being experienced switchers) and methodology make it difficult to compare their findings to other studies (Moreno et al. 2008). Nevertheless, this study introduced the possibility that codeswitching can affect semantic level processing.

Van Der Meij et al. (2011) also reported an N400 modulation in response to codeswitches in late learners of English (L2) reading English sentences with occasional Spanish (L1) switches (cf. Litcofsky and Van Hell 2017). This N400 increase was small, perhaps because the switched words were adjectives rather than the typical nouns, and was larger for learners with greater proficiency in English. They interpreted the code-switch effect on the N400 as activation cost of lexical forms in the less active language. However, a comparison of late learners to early bilinguals should be interpreted with caution, given that the N400 could hypothetically be specific to learners (e.g., McLaughlin et al. 2010). The LPC to codeswitches was also larger in amplitude for higher compared to lower proficiency individuals and was interpreted as updating the context in light of an unexpected event. High proficiency bilinguals also generated a LAN (similar to Moreno et al. 2002), which the authors argued might be related to integrating the grammatical rules of the two languages. However, the results of the LPC and LAN are inconsistent with later studies which show larger effects when switching into the less dominant language (Fernandez et al. 2019; Litcofsky and Van Hell 2017).

As for whether switches can affect semantic level processes, as indexed by the N400, a handful of recent studies has provided some clues. The presence of an N400 modulation in response to a codeswitch may depend on the direction of the switch, with reports of an N400 amplitude modulation only when switching into a less dominant language (Liao and Chan 2016). This suggests that lexical-semantic costs from a switch occur when there is less proficiency in accessing information from that language. Similarly, the frequency of switching may also modulate N400 amplitude with some evidence for the presence of an N400 only for less habitual switchers, although even habitual switchers can show an N400 amplitude modulation for codeswitches when listening to sentences (although not when reading; Fernandez et al. 2019; Litcofsky and Van Hell 2017). Switches into a language that is less frequently used for reading might also decrease the distribution of the N400 effect, with less frontal negativity than non-switches, which may reflect a change in frontal ERP effects related to integration (Ng et al. 2014).

The current study takes another look[1] at the impact of codeswitching on accessing the meaning of words in a sentence, as indexed by the N400 response, as well as how the frequency of switching in daily life affects all ERP responses to a switch. Early balanced Spanish–English bilinguals read Spanish sentences for comprehension while scalp ERPs were recorded. One noun in each sentence could appear in Spanish or English (codeswitch), and either be semantically congruent or incongruent with the preceding context (see example below). We controlled for some of the potential methodological confounds in the previous studies. Namely, in our study, the same stimuli appeared in all 4 conditions, all 4 conditions appeared mixed (rather than blocked by condition), and the position of the target in the sentence was unpredictable and never sentence final. Using a fully crossed design, which manipulates semantic congruity and the occurrence of a codeswitch on the same word, allowed us to isolate the ERP components elicited by each process and determine at what stages in comprehension, if any, these processes interact.

Two recent studies measured the effect of codeswitching on processing nouns with high or low expectancy within a sentence context (Yacovone et al. 2021; Valdés Kroff et al. 2020). Yacovone et al. (2021) showed an N400 modulation for codeswitches into the less dominant language that was equivalent to the N400 modulation for low expectancy words. However, codeswitches did not additionally modulate the N400 for unexpected words (i.e., no interaction with semantic processing). In contrast, Valdés Kroff et al. (2020) showed no effect of codeswitching on the N400 in proficient early bilinguals, only revealing the typical effect of expectancy on the N400 with smaller amplitude for nouns supported by context. Given that the current study tests proficient early bilinguals, it is possible that codeswitches may not affect lexico-semantic processing, as in Valdés Valdés Kroff et al. (2020). However, in both studies the unexpected nouns were plausible continuations. Here, we determine if codeswitching affects semantic processing of nouns that are plausible, where sentence context supports the interpretation, and implausible, where sentence context does not support an interpretation, in sentences with moderate to high sentence constraint.

A secondary goal of the current study is to determine the extent to which codeswitching in daily life affects comprehension of a switched word. In both of the studies just mentioned, the bilinguals self-reported as habitual switchers (e.g., frequently switched between languages in daily life), and frequency of switching within these frequent switchers only modulated an early positivity, possibly related to target categorization (e.g., the P300, see Valdés Kroff et al. 2020). The frequency with which a bilingual habitually codeswitches has been suggested to modulate the language-control needed to switch between languages when speaking (Blackburn 2013, 2017; Green and Abutalebi 2013; Green and Wei 2014;

Kang et al. 2015; Morales et al. 2015; Wu and Thierry 2013), and even alter brain networks involved in switching between other types of tasks (see Blackburn 2019 for a more in depth discussion of the neural impact of language environment; cf. Hernández et al. 2013; Yang et al. 2016; Yim and Bialystok 2012). A number of studies have found that daily code-switching habits modulate the switching costs incurred during language production (e.g., Kheder and Kaan 2016; Soveri et al. 2011). In particular, the cost of producing a code-switch, operationalized as the difference in response time between switched and non-switched picture naming trials, is decreased in individuals who codeswitch often in their daily lives (Prior and Gollan 2011).

Codeswitching habits appear to affect not only production, but also comprehension of a codeswitch. For example, compared to bilinguals who do not switch frequently, spoken language translators show no behavioral switch cost when reading sentences with a codeswitch into their first language (Ibáñez et al. 2010). In addition, bilinguals who frequently codeswitch become sensitive to specific code-switching patterns (Valdés Kroff et al. 2016, 2017). These studies imply that people who switch often produce and comprehend switches more easily than bilinguals who do not habitually codeswitch (see Blackburn 2018 for a review).

The effect of switching on the LPC, however, is less clear. There is some evidence that expertise, such as in playing chess, leads to larger LPC amplitude, more broadly, as a reflection of improved processing (Wright et al. 2013). Similarly, Gosselin et al. reported an LPC modulation to codeswitches only for more frequent switchers. This is opposite the finding that codeswitching elicits a N400 and LAN only in individuals who switch less frequently (Gosselin and Sabourin 2021), although the authors suggested that the LPC may have also occurred in non-habitual switchers but was obscured by the overlapping N400. In brief, the effect of switching habits has been reported rarely and inconsistently across these studies.

Here, we tested whether codeswitching frequency modulates the amplitude of the LPC, or the earlier negativities, LAN and N400, using a survey developed to measure codeswitching habits. We used the Assessment of Code Switching Experience Survey (ACSES) (Blackburn 2013) to measure how often each participant codeswitches in their daily life. Codeswitching frequency was then correlated with the amplitude of the ERP effects. In the current study, semantic violations were expected to elicit a classic N400 congruity effect and codeswitches were expected to elicit a LAN and LPC. The critical questions are (1) whether codeswitching alone elicits an N400 modulation, as in Yacovone et al. (2021), (2) whether codeswitches additionally modulate the N400 congruity effect or if processing the meaning of the word takes priority, as in Ng and Wicha (2013), and (3) whether switching habits would modulate any of the stages of processing indexed by the LAN, N400 and LPC in early balanced bilinguals.

2. Materials and Methods

2.1. Participants

Twenty-four (9 female, 15 male) healthy right-handed Spanish–English balanced bilinguals were included in the analysis. Data from one additional participant was excluded due to poor performance and excessive EEG artifacts. Participants were on average 21.5 years old, ranging from 19 to 29 years (SD = 2.3). Participants completed a series of language measures before the ERP recording, including a comprehensive Language History Questionnaire adapted from the LEAP-Q (Marian et al. 2007), the Assessment of Code Switching Experience Survey (ACSES) (Blackburn 2013), and the Woodcock Passage Comprehension test in Spanish and English (Woodcock 1991). All participants reported Spanish as their native language. Participants had an average score of 2.86 (SD = 1.22, range = 1–5.37) for 'frequency of code-switching' based on the ACSES, where 1 was "never switches" and 7 was "always switches". The age of acquisition (AOA) for English was on average 8.08 years, ranging from 0 to 19 years. Spanish and English daily use was comparable; participants reported speaking English 51.25% (SD = 20.76%) and Spanish 48.75% (SD = 20.76) on an average day [paired-samples comparing English and Spanish use $t_{23} = 0.30$, $p = 0.771$).

Balanced bilinguals in this study were defined as those with Spanish and English scores on the Woodcock Passage Comprehension test within one Cognitive-Academic Language Proficiency (CALP) level of each other (Woodcock 1991). The average Relative Mastery Index (RMI) score, on a scale of 0–100 proficiency, was 80/90 for English (SD = 11.1, range = 55/90–97/90) and 82/90 for Spanish (SD = 14.6, range = 45/90–98/90). A paired-samples t-test revealed no significant difference in the RMI scores between the two languages ($t_{23} = 2.07$, $p = 0.220$), indicating that the participants were balanced in proficiency.

2.2. Stimuli and Procedure

Materials were 128 sentence pairs in Spanish adapted from Wicha et al. (2004), each containing a target noun of medium to high cloze probability ranging between 50–100% with a mean of 81% (previously normed by a separate sample of native Spanish speakers). The targets were singular, non-cognate nouns. They were never sentence final and could appear anywhere in either the first or second sentence of the pair. The target noun in each sentence pair could be either the expected Spanish noun based on sentence context or not (semantically congruent/incongruent), or the English translation of these. This created four conditions as follows:

Control Sentence: *Caperucita Roja llevaba la comida para su abuela en una* ***canasta*** *muy bonita. Pero el lobo llegó antes que ella*. [Little Red Riding Hood carried the food to her grandmother in a pretty **basket**. But the wolf arrived before her.]

Codeswitch: *Caperucita Roja llevaba la comida para su abuela en una* ***basket*** *muy bonita. Pero el lobo llegó antes que ella*.

Semantic Violation: *Caperucita Roja llevaba la comida para su abuela en una* ***manzana [apple]*** *muy bonita. Pero el lobo llegó antes que ella*.

Code Switch–Semantic Violation: *Caperucita Roja llevaba la comida para su abuela en una* ***apple*** *muy bonita. Pero el lobo llegó antes que ella*.

Every sentence pair and target appeared in all 4 conditions across 4 lists, so that participants saw only one version of each stimulus. Each list contained 32 sentence pairs per condition. Target nouns were first grouped to match grammatical gender, animacy, word length, and frequency (as closely as possible). Word frequencies were obtained from Davis and Perea (2005) (based on the LEXESP database) for Spanish, and from Davis (2005) (based on the CELEX database) for English. The target word for each sentence pair was interchanged with another sentence pair in its group, so that the target for one sentence became the semantic violation for the next. The sentences were then normed for acceptability by 15 native Spanish speakers who did not take part in the ERP study (8 males, 7 females; ages 19 to 39, mean 25 years, SD = 5.55; paid to participate). Participants in the norming study had at least 5th grade proficiency in Spanish on the Woodcock Passage Comprehension test (62.79 to 100% correct, M = 80.90%, SD = 10.12; RMI M = 69/90, SD = 21.12, Range = 23/90–99/90) (Woodcock 1991). Participants in the norming study read all 128 experimental sentences containing a semantic violation in Spanish (no codeswitch), up to and including the target word, as well as 50 filler sentences that did not contain violations and a potential sentence to substitute for the experimental set, if needed. They were asked to rate the sentence for plausibility on a scale of 1–5, where 5 indicates that the sentence made perfect sense and 1 indicates that the target word in the sentence did not make sense. Sentences with violations were rated between 1.4 and 3.1 (M = 2.0, SD = 0.39) and filler sentences without violations were rated between 4.1 and 5.0 (M = 4.79, SD = 0.23), indicating that the sentences containing violations were perceived as implausible continuations of the sentence.

The ERP recording took place in a single session in an electrically shielded, sound-attenuating chamber. Participants were seated approximately 19 inches from a 19-inch CRT monitor. The task was to read the sentences silently for comprehension. Participants were given instructions and a short practice with a token from each stimulus condition. Twenty-four trials (randomly chosen) were followed by a probe sentence in Spanish related to the sentence pair just presented. Participants made a true-false judgment by pressing a button on a keyboard for these probe sentences only. An equal number of probes occurred in each condition, and in the first or second sentence in the pair. The probe sentences were designed to be moderately difficult and were used to ensure attentiveness to the task and sentence content.

The target noun appeared in the first sentence of the pair on 66 trials (early) and in the second sentence on 62 trials (late). To help participants plan eye blinks (to avoid eye movement artifact during the targets), the trials were blocked by whether the target appeared early or late, then randomized within a block. The order of presentation was counterbalanced so that half of the participants began with the early target sentence pairs and half with the late target sentence pairs. The sentence containing the target word was presented one word at a time (300 ms with 200 ms ISI) in the center of the screen, preceded by a 1-s fixation cross (200 ms ISI). The last word of the sentence was followed by a 700 ms ISI before the onset of the next sentence pair. Stimuli were presented as white text on a black background in the Verdana font (letter matrix = 30 by 68 pixels). Participants were asked to minimize blinking during the target sentence and allowed to blink naturally during the other sentence, which was presented as a whole and stayed on the screen until the participant pressed a button to advance.

2.3. ERP Parameters

Continuous electroencephalogram (EEG) was recorded from 32 tin electrodes: 26 embedded in a standard cap (Electro-Cap Inc., Eaton, OH, USA) in geodesic array (as per M. Kutas lab convention), and 6 free electrodes to record eye artifacts and mastoid process references. Blinks were monitored using electrodes placed under each eye, and horizontal eye movements through a bipolar recording from the outer canthi of each eye. The data was recorded using an SAI analog bio-amplifier with a band pass filter set from 0.01 to 100 Hz. Electrode impedances were maintained below 5 kΩ. Data was recorded to a left mastoid reference on-line and re-referenced to the algebraic sum of the left and right mastoids (Wicha et al. 2004). The output of the bio-amplifier was fed into a 32 channel 12-bit analogue-to-digital converter and sampled at 250 Hz. ERPSYSTEM software (developed by P. J. Holcomb, Boston, MA, USA) was used to simultaneously present visual stimuli to the participant and deliver event and timing codes to the data acquisition PC at the onset of each word.

Trials with artifacts due to eye movements, excessive muscle activity, or amplifier saturation were eliminated offline before averaging using uniformly applied algorithms. Approximately 8.71% of the total trials (range = 1.5–14.5%) were rejected, with roughly equal loss of data across conditions. Data was averaged for each experimental condition from the onset of the target word, relative to a 100 ms pre-stimulus baseline. A digital band-pass filter set from 0.1 to 30 Hz was used prior to running analyses to reduce EEG content irrelevant to the components of interest.

3. Results

One participant was eliminated due to poor performance on the probe questions (58.3% accuracy). Accuracy for the other 24 participants on judging the truth of the sentence probes was high (mean = 85.5%, range = 70.8–95.8%), indicating good comprehension of the sentences and attentiveness to the task.

Figure 1 shows grand average ERPs (n = 24) to the target noun for each of the 4 conditions: control, semantic incongruity, codeswitch, and codeswitched semantic incongruity. Visual inspection of the waveform reveals equivalent early sensory components across the conditions—P1-N1-P2—followed by slow wave components. Overall, semantically incongruous nouns elicited an enhanced negativity between 300 and 500 ms post-stimulus onset (N400) relative to congruous nouns. Peak latency analysis revealed that the N400 peaked at 399 ms (SD = 28.09 ms) post stimulus onset.

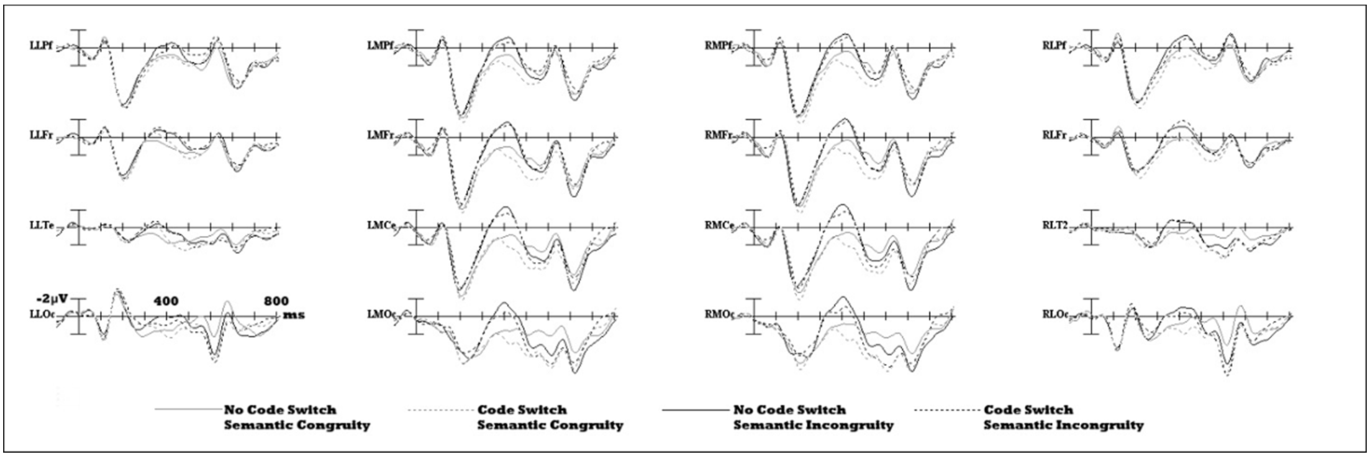

Figure 1. ERP waveforms. Raw waveforms are presented for each condition, time-locked to the onset of each stimulus.

Visual inspection indicates that there may be two distinct but overlapping positivities, a codeswitching positivity which onsets earlier than an LPC to semantic violations. Relative to non-switched Spanish nouns, codeswitched English nouns elicited an enhanced positivity between 320 to 650 ms post-stimulus onset, which overlapped with the N400. Semantic incongruities also elicited a late positive component from 500–800 ms post-stimulus onset. Therefore, an onset latency analysis was performed to obtain reliable onset latencies for the codeswitch positivity and the later semantic violation positivity.

Difference waves were obtained for the Code Switch Effect in the absence (codeswitch minus control) and presence of a semantic incongruity (codeswitched semantic incongruity minus semantic incongruity with no codeswitch) [Figure 2]. Likewise, the Semantic Congruity Effect was calculated as a difference wave in the absence (semantic incongruity minus control) and presence of a codeswitch (codeswitched semantic incongruity minus codeswitch without a semantic incongruity). The onset latency was calculated as the time-point of the first positive deflection exceeding 100 ms in duration based on the difference waves. The onset of the codeswitch effect (codeswitches minus non-switches) was similar in the presence (324 ms) and absence (319 ms) of a semantic incongruity, and overlapped in time with the N400. The LPC onset to a semantic incongruity was later than that of a codeswitch and was similar in the absence (447 ms) and presence (424 ms)[2] of a codeswitch.

Time windows for subsequent analysis were selected based on latency onsets. Due to overlap between the LPC and the codeswitching positivity, the LPC was analyzed in two separate windows, the LPCa from 500–650 ms and the LPCb from 650–800 ms (see Barber and Carreiras (2005) for a similar analysis of the LPC, i.e., P600a/P600b). Mean amplitude was therefore, measured in three non-overlapping time windows—300 to 500 ms (N400), 500 to 650 (overlap between the codeswitching positivity and the semantic positivity[3]—the LPCa), and 650 to 800 ms (LPCb), plus a non-orthogonal fourth window encompassing the entire code-switching positivity (320–650 ms) for comparison of the codeswitching effect to other studies in the literature.

Mean amplitude in each time window was subjected to omnibus repeated-measures ANOVAs with 2 levels of Semantic Congruity (congruous, incongruous), 2 levels of Code Switching (no switch, codeswitch) and 26 levels of Electrode. Interactions by Electrode were further analyzed using a 16-electrode subset with 2 levels of Semantic Congruity, 2 levels of Code Switching, 2 levels of Hemisphere (right, left), 2 levels of Laterality (medial, lateral), and 4 levels of Anteriority (prefrontal, frontal, central, and occipital) to determine the distribution of the effects (Wicha et al. 2004). Pairwise comparisons to explain interaction effects were conducted when appropriate. Regions of interest were selected based on the distributional analyses in order to measure localized effects. Effects were considered significant at $p < 0.05$, with repeated measures with greater than one degree of freedom reported after Greenhouse-Geisser correction.

3.1. N400

A main effect of Semantic Congruity, where responses to semantic incongruities were more negative than congruities, was observed [$F_{(1,23)} = 79.791$, $p < 0.001$] between 300 to 500 ms. The significant interaction between Electrode and Semantic Congruity [$F_{(2.9,66.6)} = 9.574$, $p < 0.001$] was further analyzed for distributional differences (main effect of Electrode, [$F_{(4.4,100.7)} = 4.342$, $p = 0.002$]). The distributional analysis revealed interactions of Semantics and Laterality [$F_{(1,23)} = 60.793$, $p < 0.001$], Semantics, Laterality and Hemisphere [$F_{(1,23)} = 6.339$, $p = 0.019$], and Semantics, Laterality and Anteriority [$F_{(2.3,52.1)} = 7.365$, $p = 0.001$], indicating that the effect of Semantic Congruity was widely distributed across the scalp and maximal over right medio-central sites, consistent with the N400 literature. Visual inspection hints that the semantic incongruity alone elicits a larger N400 than the semantic violation in the presence of a codeswitch, especially over medial posterior sites. However, based on the difference waves, there was no difference between the semantic congruity effect in the presence of a codeswitch (M = −1.004 μV, SE = 0.147) compared to the effect in the absence of a codeswitch (M = −0.815 μV, SE = 0.154) [$F_{(1,23)} = 0.720$, $p = 0.405$]. The apparent visual difference stems from a more positive waveform for codeswitches (M = 1.981 μV, SE = 0.314) relative to the control (M = 1.384 μV, SE = 0.366), which trends towards significance during the N400 time window ($p = 0.056$).

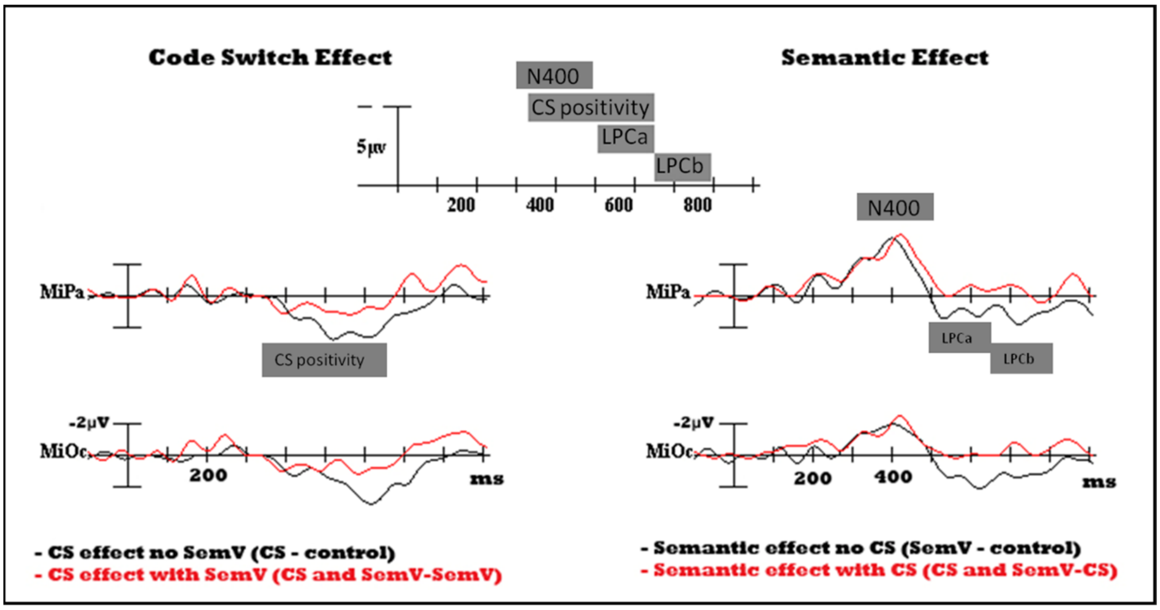

Figure 2. Difference waves of the codeswitch and semantic effects. Difference waves for the codeswitch effect without a semantic violation represent a point-by-point subtraction of the grand averaged response to a codeswitch (CS) minus the control; difference waves for the codeswitch effect with a semantic violation represent a point-by-point subtraction of the grand averaged response to a codeswitched semantic violation (CS and SemV) minus the semantic violation alone (SemV). Difference waves for the semantic effect without a codeswitch are calculated as the response to a semantic violations (SemV) minus the control; the semantic effect with a codeswitch is calculated as the response to a codeswitched semantic violation (CS and SemV) minus the response elicited to codeswitches alone (CS). These waves indicate that code-switches elicit a positivity with a latency onset around 320 ms. This onset overlaps with the first half of the LPC (LPCa) elicited to a semantic violation, but not the LPCb.

There was no significant main effect of Code Switching [$F_{(1,23)}$ = 3.820, p = 0.063], no interaction between Code Switching and Semantic Congruity [$F_{(1,23)}$ = 0.725, p = 0.403], and no interaction between Semantic Congruity, Code Switching and Electrode [$F_{(4.3,98.7)}$ = 1.334, p = 0.261], meaning that codeswitching did not affect semantic processing in this window. A significant interaction between Code Switching and Electrode [$F_{(3.9,88.9)}$ = 3.727, p = 0.008] was found, so a distributional analysis was performed. An interaction between Code Switching and Laterality [$F_{(1,23)}$ = 7.676, p = 0.011] indicated that responses to codeswitches were more positive than non-switches over medial sites [$F_{(1,23)}$ = 5.260, p = 0.031], but not lateral sites [$F_{(1,23)}$ = 1.339, p = 0.259]. An interaction between Code Switching and Anteriority indicated that codeswitches elicited a positivity compared to non-switches that was significant only over posterior sites [$F_{(1,23)}$ = 11.529, p = 0.002]. Together this indicates that the Code Switching effect is maximal over medial posterior sites. No other interactions were significant.

3.2. Left Anterior Negativity (LAN) Region of Interest Analysis

A negativity over left lateral sites was visible, similar to the one observed for codeswitches in Moreno et al. (2002) [Figure 3]. Based on Moreno's analysis, we performed a LAN region of interest analysis from 275–400 ms. A t-test was conducted to compare the mean amplitude of the left anterior electrodes (LLFr, LDCe, and LLTe) for the codeswitch and control condition. Responses to codeswitches were more negative in amplitude compared to non-switches [$t_{(23)}$ = 2.248, p = 0.034].

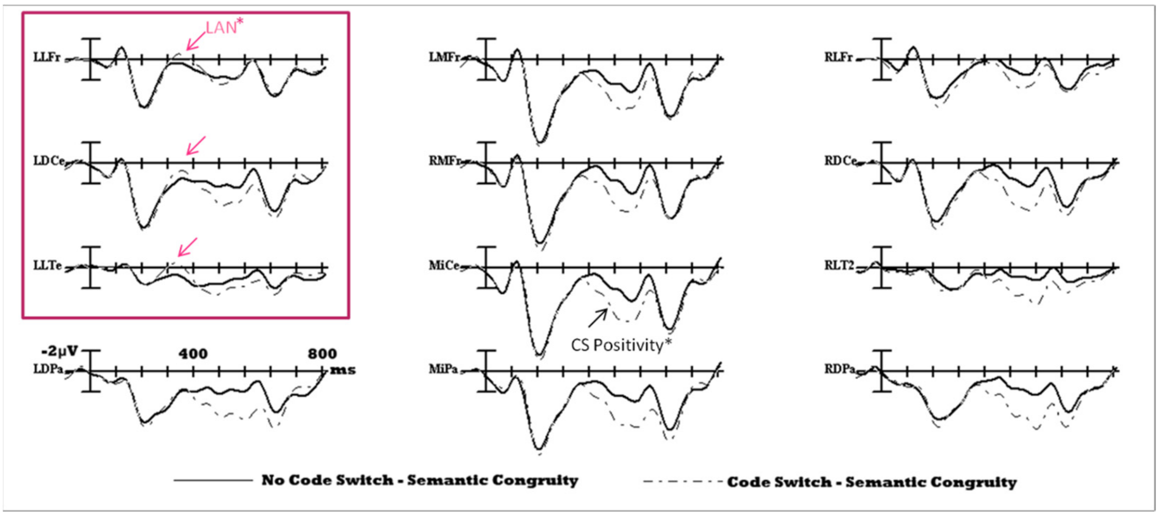

Figure 3. Electrophysiological correlates of codeswitching. A code-switching positivity can be observed from 320–650 ms post stimulus onset. A negativity is also observed over left anterior lateral sites (* $p < 0.05$).

3.3. LPCa Mean Amplitude 500–650 ms

In the LPCa window, a main effect of Code Switching was found [$F_{(1,23)}$ = 10.403, p = 0.004], with codeswitches eliciting more positive amplitude than non-switches. A main effect of Electrode [$F_{(4.0,91.1)}$ = 22.158, $p < 0.001$] and an interaction of Code Switching by Electrode [$F_{(3.6,83.6)}$ = 12.034, $p < 0.001$] led to distributional analyses. Interactions of Code Switching with Hemisphere [$F_{(1,23)}$ = 6.223, p = 0.020] and Code Switching with Laterality [$F_{(1,23)}$ = 5.372, p = 0.030], and a three way interaction of Code Switching by Hemisphere by Laterality [$F_{(1,23)}$ = 7.002, p = 0.014] indicated that the Code Switching Effect (greater positivity for switches than non-switches) was significant over both medial and lateral sites in both hemispheres, but most positive for medial sites over the right hemisphere. An interaction of Code Switching by Anteriority [$F_{(1.3,30.1)}$ = 20.754, $p < 0.001$] indicated that the codeswitching effect was only significant over central and occipital sites and largest over occipital sites. Thus, the Code Switching effect is strongest over medial posterior sites in this time window and the LPCa window (500–650 ms) overlaps in time with the Code-Switching Positivity window (320–650 ms).

No main effect of Semantics was found in this time window [$F_{(1,23)} = 0.053$, $p = 0821$]. However, an interaction of Semantics by Electrode [$F_{(3.8,86.6)} = 5.161$, $p = 0.183$] led to distributional analyses. Interactions between Semantic Congruity and Anteriority [$F_{(1.3,30.6)} = 10.559$, $p = 0.001$] and Semantic Congruity, Laterality and Anteriority [$F_{(2.4,54.6)} = 4.474$, $p = 0.012$] revealed that semantic incongruities elicit a significantly larger positivity than semantic congruities at medial and lateral occipital sites [$F_{(1,23)} = 5.081$, $p = 0.034$; $F_{(1,23)} = 13.750$, $p = 0.001$, respectively] and lateral central sites [$F_{(1,23)} = 5.203$, $p = 0.032$]. While the Code Switching effect continued to be more positive for switches than non-switches, the Semantic effect reversed from the previous N400 window, with incongruous being now more positive than congruous words.

Interactions were also found between Code Switching by Semantics [$F_{(1,23)} = 13.820$, $p = 0.001$] and Code Switching by Semantics by Electrode [$F_{(4.2,97.4)} = 2.567$, $p = 0.040$]. Pairwise comparisons with Bonferroni correction revealed that the codeswitching effect (larger positivity for codeswitched versus non-switched words) was significant only in the absence of a semantic violation ($p < 0.001$), and not in the presence of a semantic violation ($p = 0.541$) [Figure 4]. Interactions of Code Switching by Semantics by Laterality [$F_{(1,23)} = 13.158$, $p = 0.001$] and a four-way interaction of Code Switching, Semantics, Laterality, and Anteriority [$F_{(2.4,55.5)} = 3.550$, $p = 0.028$] revealed that a widespread codeswitching effect is found in the absence of a semantic violation – largest over medial occipital sites and significant at all but lateral prefrontal sites. However, in the presence of a semantic incongruity, the effect is only significant at occipital sites, both lateral ($p = 0.034$) and medial ($p = 0.028$).

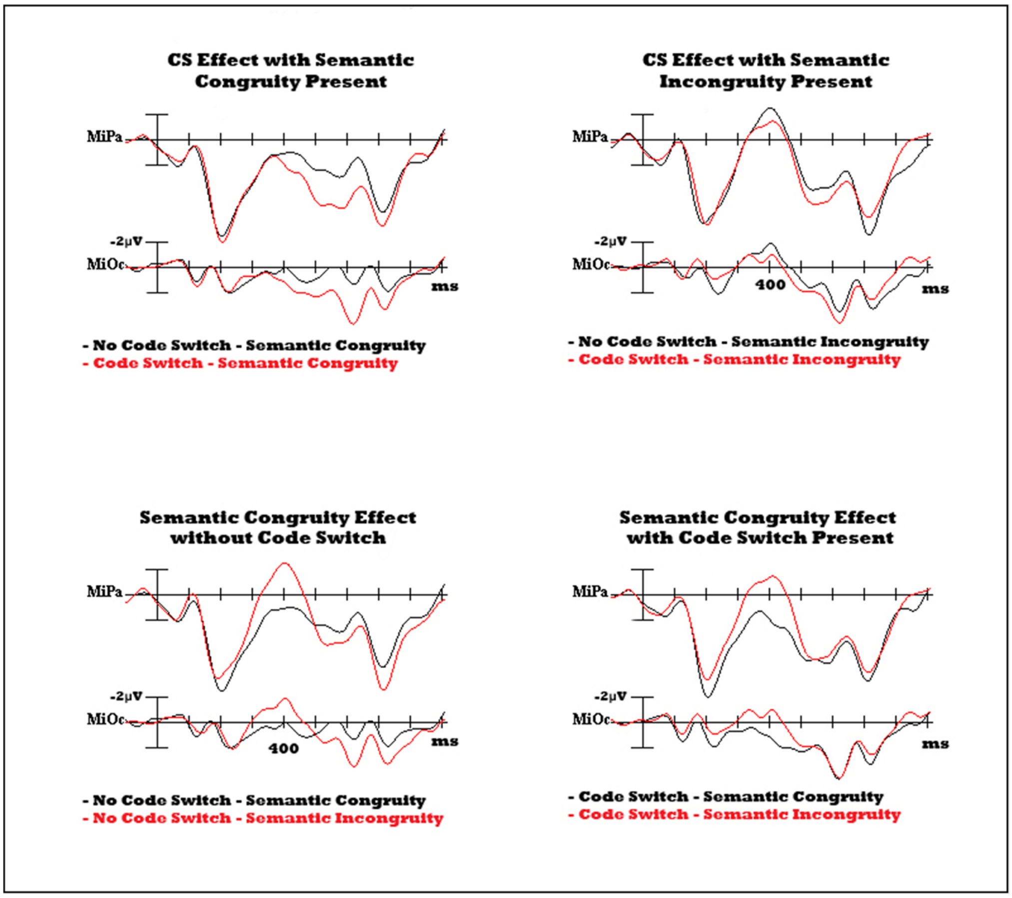

Figure 4. Interactions of code-switching and semantic congruity.

Conversely, the semantic congruity effect (larger positivity for semantic incongruities than congruities) was significant only when there was no codeswitch ($p = 0.040$). Interactions of Code Switching by Semantics by Laterality [$F_{(1,23)} = 13.158$, $p = 0.001$] and a four-way interaction of Code Switching, Semantics, Laterality, and Anteriority [$F_{(2.4,55.5)} = 3.550$, $p = 0.028$] revealed that in the absence of a codeswitch, the semantic effect is significant over central ($p_{\text{medial}} = 0.023$; $p_{\text{lateral}} = 0.002$) and occipital sites ($p_{\text{medial}} = 0.002$; $p_{\text{lateral}} < 0.001$), and largest over medial electrodes. In the presence of a codeswitch, there was a trend towards a reversal of this effect (semantic congruities are more positive than incongruities) ($p = 0.066$) [Figure 4], which became significant at frontal and prefrontal sites, and was largest over medial prefrontal sites ($p = 0.013$). No other interactions were significant.

3.4. LPCb Mean Amplitude 650–800 ms

No main effects of Semantics [$F_{(1,23)} = 1.385$, $p = 0.251$] or Code Switching [$F_{(1,23)} = 0.304$, $p = 0.587$] were found in this time window. However, an interaction of Semantics and Code Switching [$F_{(1,23)} = 6.582$, $p = 0.017$], along with subsequent pairwise comparisons, revealed that the Semantic effect (semantic incongruities are more positive than congruities) is significant in the absence of a codeswitch ($p = 0.044$), but not in the presence of one ($p = 0.574$). The Code Switch effect (switches are more positive than non-switches) was marginally significant ($p = 0.065$) in the absence of a semantic violation and not significant in the presence of one ($p = 0.275$). Main effects of Electrode [$F_{(3.5,79.5)} = 15.942$, $p < 0.001$] and interactions of Semantics with Electrode [$F_{(3.7,84.2)} = 3.047$, $p = 0.025$] and Code Switching by Semantics by Electrode [$F_{(3.5,80.5)} = 3.306$, $p = 0.019$] led to distributional analyses. Interactions of Code Switching by Semantics by Laterality [$F_{(1,23)} = 9.700$, $p = 0.005$] and Code Switching by Semantics by Anteriority [$F_{(1.3,30.8)} = 4.602$, $p = 0.030$] indicated that semantic incongruities are more positive than congruities only in the absence of a codeswitch and only over medial ($p = 0.031$) occipital sites ($p < 0.001$) and central sites ($p = 0.003$). The codeswitching effect is only significant in the absence of a semantic incongruity over lateral sites ($p = 0.046$), but not medial sites ($p = 0.097$). An interaction of Semantics by Anteriority [$F_{(1.3,29.5)} = 5.544$, $p = 0.019$] indicated that semantic incongruities elicited a positivity compared to congruities only over occipital sites ($p = 0.003$), with marginal significance over central sites ($p = 0.061$). No other theoretically relevant interactions were significant.

3.5. Code Switching Positivity 320–650 ms

In this expanded window, main effects of Code Switching [$F_{(1,23)} = 9.249$, $p = 0.006$] and Semantics [$F_{(1,23)} = 27.919$, $p < 0.001$] reached significance. Amplitudes to semantic incongruities were more negative than non-violations, and codeswitches were more positive than non-switches. A main effect of Electrode [$F_{(4.7,108.9)} = 11.549$, $p < 0.001$] and interactions of Electrode with Semantics [$F_{(3.5,80.0)} = 4.433$, $p = 0.004$] and Electrode with Code Switching [$F_{(4.5,103.9)} = 7.752$, $p < 0.001$] led to distributional analyses. The interaction of Code Switching by Semantics by Electrode was not significant.

An interaction of Code Switching by Anteriority [$F_{(1.3,29.5)} = 12.402$, $p = 0.001$] indicated that the Code Switching Effect was only significant over central and posterior sites, and maximal over posterior sites. Interactions of Code Switching and Laterality [$F_{(1,23)} = 8.498$, $p = 0.008$] and Code Switching and Hemisphere [$F_{(1,23)} = 4.752$, $p = 0.040$] indicated that the Code Switching Effect was significant over both medial and lateral sites and in both hemispheres, but larger over right medial sites.

An interaction of Semantics by Laterality [$F_{(1,23)} = 22.643$, $p < 0.001$] and Semantics by Laterality by Anteriority [$F_{(2.4,54.4)} = 6.121$, $p = 0.002$] indicated that semantic incongruities elicit a larger negativity compared to congruous words, and this effect was significant over all sites except lateral occipital sites. The negativity for semantic violations is largest over medio-central sites—consistent with the N400.

Finally, an interaction of Code Switching by Semantics [$F_{(1,23)} = 5.316$, $p = 0.030$] and Code Switching by Semantics by Laterality [$F_{(1,23)} = 8.104$, $p = 0.009$] indicated that the semantic incongruities elicit a negativity compared to congruities both in the presence

($p < 0.001$) and absence ($p = 0.031$) of a codeswitch, but the effect is largest in the presence of a codeswitch. In both cases, the semantic effect is larger over medial than lateral sites. The Code Switching effect is only significant in the absence of a semantic violation ($p = 0.001$) but not in the presence of a semantic violation ($p = 0.455$). It is largest over medial sites but significant over all sites. This interaction occurs in part due to overlap of this time window with both the N400 and the onset of the LPC to a semantic violation. No other theoretically relevant interactions were significant. The analysis in this larger window shows the effects of semantics on the N400 and the effects of codeswitching and semantics on the LPC, however, the more refined analysis of the smaller windows shows differences across the LPCa and LPCb windows.

3.6. Individual Differences

As in Moreno et al. (2002), we performed regression analyses to explore whether the main effects, N400 responses to semantic violations and LPC responses to codeswitches, were correlated with any of our language proficiency measures or frequency of daily codeswitching measured with ACSES.

3.6.1. N400 (Semantic Congruity)

Regression analyses were performed on N400 amplitude of the semantic congruity effect (i.e., the averaged amplitude of all 26 electrodes in the semantic incongruity condition minus the control condition) and N400 peak latency using the following covariates: Woodcock RMI in English, Woodcock RMI in Spanish, and ACSES scores. There were no correlations of any factors with N400 mean amplitude. Spanish proficiency measured by the Woodcock RMI correlated negatively with N400 peak latency [$r_{(22)} = -0.421$, $R^2 = 0.177$, $F_{(1,22)} = 4.741$, $p = 0.040$], indicating that the N400 effect is earlier in bilinguals with greater proficiency in the base language of the sentences, similar to Moreno et al. (2002). Semantic access may occur earlier in bilinguals with greater proficiency. There was no correlation between peak latency with English proficiency or with ACSES scores.

3.6.2. LCPa and LCPb

No significant correlations were found for the LPCa or LPCb windows (measured separately) between Spanish language proficiency and the codeswitch effects of latency or amplitude (mean of all electrodes; codeswitch minus control condition) or for the semantic congruity effect (mean of all electrodes; semantic incongruity minus control). English RMI was positivity correlated with both the Code Switch effect [$r_{(22)} = 0.478$, $R^2 = 0.229$, $F_{(1,22)} = 6.520$, $p = 0.018$] and the Semantic effect [$r_{(22)} = 0.447$, $R^2 = 0.200$, $F_{(1,22)} = 5.493$, $p = 0.029$] during the LPCb. As English proficiency increased, the LPCb to both a codeswitch and a semantic violation increased in amplitude. No other correlations with proficiency were significant in these time windows.

ACSES scores were positively correlated with the amplitude of the Code Switch Effect, when measured from 320–650 ms [$r_{(22)} = 0.411$, $R^2 = 0.169$, $F_{(1,22)} = 4.480$, $p = 0.046$]. As ACSES scores increased, i.e., as frequency of code-switching in daily life increased, mean amplitude of the codeswitching positivity also increased [Figure 5, right panel]. The correlation between ACSES and the Code Switch Effect was also measured during the LPCa, which partially overlaps with the codeswitching positivity, and LPCb windows. The correlation was significant during the LPCa from 500–650 ms [$r_{(22)} = 0.415$, $R^2 = 0.172$, $F_{(1,22)} = 4.582$, $p = 0.044$] but not significant during the LPCb from 650–800 ms [$r_{(22)} = 0.376$, $R^2 = 0.141$, $F_{(1,22)} = 3.624$, $p = 0.070$]. No correlation between ACSES and the Semantic Congruity Effect was found during these time windows [Figure 5, left panel]. ACSES scores were not correlated with the onset of the codeswitch positivity.

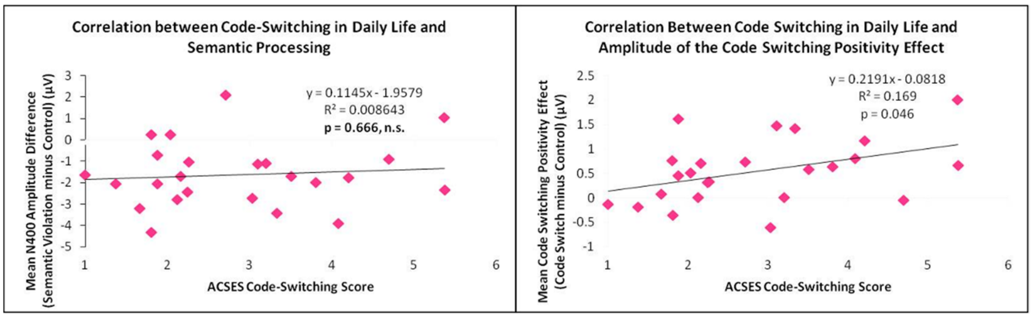

Figure 5. Correlations of frequency of code-switching with the amplitude semantic congruity effect on the N400 (**left panel**) and the amplitude of the code-switching positivity (320–650 ms; **right panel**).

4. Discussion

We used ERPs to determine the time course of processing a codeswitch and how codeswitching interacts with access to the meaning of a word. We also investigated how codeswitching experience modulates comprehension and integration of a codeswitched word. Our data suggest that semantic level processing in a sentence context is not dependent on language identification in proficient bilinguals, and that increased frequency of codeswitching mainly affects the cost of processing a switch at post lexico-semantic integration.

As expected, a widely distributed, right medio-central N400 was observed in response to semantically implausible nouns (see Kutas and Federmeier 2011; Kutas and Hillyard 1980b). This was true regardless of whether the noun was in the same or different (codeswitch) language as the sentence. That is, a codeswitch did not modulate the amplitude of the N400, either for congruent or incongruent target words. The N400 was followed by a late positivity, typical of a semantic P600 (for a review, see Bornkessel-Schlesewsky and Schlesewsky 2008). In contrast, codeswitches elicited a LAN followed by a late positivity, with both effects overlapping in time with the N400, replicating prior findings (Moreno et al. 2002; Van Der Meij et al. 2011). There was an interaction between semantics and codeswitching only at the later part of the positivity (LPCb), but not at the LAN, N400 or LPCa. This interaction was sub-additive, such that codeswitches and semantic incongruities elicited equivalent LPC amplitude individually as did words that were both codeswitched and semantically incongruous. Language proficiency in this balanced bilingual population had two effects. Proficiency in Spanish (e.g., the matrix language) modulated N400 latency, but not N400 amplitude, while proficiency in English (codeswitch) modulated the semantic P600 and the codeswitching positivity. Finally, codeswitching habits modulated the late positivity, with larger positive amplitude for individuals who are more frequent codeswitchers. Codeswitching habits did not modulate the N400. We discuss the significance of each of these findings below.

4.1. Lack of a Codeswitch Effect on the N400 Indicates That Access to Meaning Is Independent of Language Membership

Ng and Wicha (2013) found that in single word recognition, meaning is accessed prior to language membership during lexical selection. In their study, while reading sequentially presented words in English and Spanish, bilinguals were asked to press a button each time they saw a word from a specific semantic category (i.e., people words), but only if they appeared in a specific language. They found that bilinguals started to categorize the words as targets and non-targets before rejecting the words from the non-target language. This happened especially when the non-target language was the more frequently used language (English), making it harder to ignore. Their findings implied that early proficient bilinguals can prioritize the meaning of words independent of language membership (for an example of when language membership is prioritized see Hoversten et al. 2015).

Similarly, in the current study, there was no effect of language membership, namely the language in which the target word was presented, and there was no interaction of codeswitching and semantic congruence on the amplitude or latency of the N400. Codeswitched words were processed in the same time course, eliciting N400 modulations in the same latency around 300 to 500 ms, and to the same degree, eliciting indistinguishable amplitude modulations, as non-switched words. As with single word processing (Ng and Wicha 2013), this suggests that early proficient bilinguals can prioritize meaning over the language of the presented word during this early stage of comprehension indexed by the N400, at least when the sentence context supports a specific continuation. This finding is also in line with Valdés Kroff et al. (2020) where proficient bilinguals showed no modulation of the N400 to codeswitches in a similar design, and with Moreno et al. (2002) who suggested that the cost of processing a codeswitch may not occur at the semantic level and instead at a later stage reflected by a late positivity.

These findings differ, however, from studies that do show a modulation on the N400 (Table 1). Proverbio et al. (2004) reported an enhanced N400 for semantically acceptable codeswitches compared to sentences without language switches. Proverbio et al. used a blocked presentation and included multilingual professional interpreters, which could complicate generalization of their findings. Nevertheless, this was not the only study to show a codeswitching modulation on the N400. Several studies have shown an asymmetric effect of codeswitching, where switching into the non-dominant language elicits a modulation of the N400 (Liao and Chan 2016). Others have reported switch effects in both directions in habitual switchers (Yacovone et al. 2021), or even in language learners (Ruigendijk et al. 2016). Often in these studies language proficiency was measured by self-assessment, where as our population was carefully assessed for proficiency in both languages. So, perhaps it is language proficiency that affects access to the meaning of words rather than codeswitching, per se.

Some studies have suggested that switching frequency might modulate the N400, with studies showing an N400 modulation in habitual switchers (Litcofsky and Van Hell 2017) and others suggesting the opposite (Gosselin and Sabourin 2021). Using the reliable and validated ACSES survey to measure self-reported switching habits, we did not observe an effect of switching frequency on the amplitude or latency of the N400. This again suggests that differences in how proficiency is measured may contribute to when a modulation of the N400 is observed, rather than switching habits, per se.

In fact, we observed a modulation of the N400 with proficiency, partially replicating findings from Moreno et al. (2002). Increased proficiency in the language of sentence presentation (Spanish) led to earlier N400 latency. This N400 latency shift even with small proficiency differences in our more balanced bilingual population indicates that N400 *latency* is quite sensitive to language proficiency. Semantic access may occur earlier in bilinguals with greater proficiency.

In brief, codeswitching did not modulate the N400 in this study, revealing that at least for early fluent bilinguals, code-switches are not processed as semantic violations and the brain can prioritize the meaning of words regardless of which language they appear in during sentence reading.

Table 1. Summary of studies to date measuring ERPs to codeswitches during sentence comprehension. Key population and task factors are noted, including switching habits where reported, direction of switch where appropriate and format of sentences (written or auditory). Effects on N400, LAN and LPC components are specific to codeswitches; other relevant effects are noted, but additional effects may be reported in original studies.

Study	Population	Task	N400	LAN	LPC	Other
Moreno et al. 2002	English-Spanish; English dominant	Read English sentences; expected, lexical switch or Spanish (non-dominant) translation of expected word	No	Yes	Yes; modulated by Spanish proficiency earlier and smaller for more proficient	Late Frontal positivity for lexical and code switches
Proverbio et al. 2004	Multilingual interpreters	Read English and Italian sentences; blocked by mixed (codeswitch) and unmixed sentences but mixed by matrix language	Larger N400 overall for mixed (switched) than unmixed	Not reported	Not reported	Report a modulation of the N1 for codeswitches
Van Der Meij et al. 2011	Late learners of English	Read English sentences (L2) switch into L1 Spanish (dominant); targets were adjectives	Small but sig; larger for high English proficiency	Yes	Yes	Frontal positivity to CS starts earlier than posterior LPC; report ortho N250 that might be N400
Ng et al. 2014	Balanced early Spanish–English	Read English stories with embedded Spanish nouns and verbs	Reduced for switches over frontal sites but not medial central (more focal)	Yes	Yes; earlier for nouns than verbs (LPCa)	No effect of story position on codeswitch effects
Liao and Chan 2016	Mandarin-Taiwanese early bilinguals; dominant Mandarin	Spoken sentences in both languages	Only when switching into non-dominant language	Only when switching into non-dominant language	Yes in both directions	
Ruigendijk et al. 2016	Russian learners of German; intermediate and high proficiency & native German	Spoken German sentences with Russian codeswitch or German semantic violation	Yes, even for monolinguals	Not reported	Yes, larger in less proficient learners	Reported N400 possibly N2 modulations
Litcofsky and Van Hell 2017	L1 Spanish, early L2 English; Dominant L2; habitual codeswitchers	Written sentences English and Spanish translations; multi-word switch after target noun	No	No	Only when switching into non-dominant language	
Fernandez et al. 2019	L1 Spanish, early L2 English; $\frac{1}{2}$ Dominant L2; equal daily use of both; habitual codeswitchers	Auditory sentences English and Spanish with multiword codeswitch after target noun	Yes; in both switch directions	No	Only when switching into non-dominant language	
Vaughan-Evans et al. 2020	Welsh-English; proficient in both	Read English and Welsh matrix language; manipulated adjective-noun word order as per English or Welsh; codeswitch on every trial; semantic acceptability judgment	No	Yes, for both switch types	LPC only for matrix language framework switches	

Table 1. *Cont.*

Study	Population	Task	N400	LAN	LPC	Other
Kaan et al. 2020	Spanish–English, early, English dominant	Read English sentences with multiword Spanish switch (non-dominant) starting with function word; presence of bilingual or monolingual confederate	No; possibly because switch measured at function word	No	Yes; reduced with presence of bilingual confederate	Early positivity
Valdés Kroff et al. 2020	Spanish–English early proficient; frequent switchers	Reading Spanish with English switched nouns with high/low expectancy at noun 2 × 2 w/switches	No	No	Yes	Early posterior positivity (P300?) modulated by switch frequency
Gosselin and Sabourin 2021	French-English, early and proficient; habitual and non-habitual switchers	Determiner phrase switches	Only for non-habitual switchers	Only for non-habitual switchers	Only for habitual switchers	LPC might be obscured by N400 in non-habitual group
Yacovone et al. 2021	Spanish–English; English dominant; frequent codeswitchers	Spoken English stories, high/low expectancy at target noun w/Spanish (ND) switches	Yes; non-dominant switch	No	Yes, regardless of expectancy	Same modulation for codeswitch, low expectancy non-switch and low expectancy switch
This Study	Spanish–English early and balanced proficiency; vary in switching habits	Spanish sentences with Single English word switch 2 × 2 w/semantic violations	No; only main effect of congruency	Yes	Yes; reduced with more switch frequency	

4.2. Codeswitching LAN—A Different Kind of Negativity

Encountering a codeswitched word elicited a LAN, a negativity that overlaps in time with the N400, but has a different distribution and is commonly observed to codeswitched words both in a sentence context and isolated word pairs (Barber and Carreiras 2005). A LAN has also been observed in response to syntactic agreement violations and has been attributed to early detection of a mismatch between morphosyntactic features (and difficulty integrating these features into the syntactic structure; Friederici 2002; Münte et al. 1997a, 1997b), or to the taxing of working memory (Kluender and Kutas 1993). The sentences in the current study were presented in Spanish, a language with rich morphological marking for gender and number agreement. Therefore, it is possible that the mismatch between the Spanish article and the English word contributed to the LAN.

Similarly, Litcofsky and Van Hell (2017) observed a small LAN on the second codeswitched word in a sentence that began in Spanish and switched the remainder of the sentence into English. In line with this morphological mismatch hypothesis, they did not observe this LAN for switches out of English into Spanish. However, a LAN has also been observed for codeswitches embedded in English sentences, a language with poor morphological markings (Moreno et al. 2002). Because the LAN is also linked to working memory load (Kluender and Kutas 1993), Moreno et al. (2002) suggested that codeswitches incur a working memory cost, perhaps in order to integrate morphological cues (such as agreement) across languages. Thus, the LAN in the current study likely reflects this cross-language syntactic integration, while the N400 reflects initial access to semantic memory.

4.3. Overlap of the Codeswitching Positivity with the N400 Suggests Parallel Processing of Meaning and Language Membership

The temporal overlap between the N400 and codeswitching positivity suggests parallel processes. These findings, along with Ng and Wicha (2013) in single words, provide evidence that language membership does not need to be identified prior to semantic retrieval, but rather these are separable processes that can occur in parallel (although for examples where it does not occur in this order, see Hoversten and Traxler 2020; Hoversten et al. 2015, 2017).

Early models of word recognition proposed a serial word recognition system in which language membership was identified prior to lexico-semantic retrieval (Kirsner et al. 1984). Our data challenges a serial model in favor of one which allows for parallel processing of semantics and language membership. Our results are more consistent with the BIA+ model of word recognition, in which access to semantics occurs independently of identification of the language (Dijkstra and Van Heuven 2002). In BIA+, identification of language membership occurs through language nodes, or labels, which provide information about properties of the word, but do not substantially affect word activation and semantic retrieval (Van Heuven and Dijkstra 2010). According to the model, word recognition is language non-selective, meaning that potential word choices in both languages are activated (Dijkstra and Van Heuven 2002).

Note that the need for an explicit language node may not be obligatory and may be instead an emergent property of the lexical spread within the lexicon (e.g., Elman 1993). Nevertheless, our results suggest that knowledge of the language that the word belongs to is not necessary for initial access of semantic level information, as indexed by the N400. Gullifer et al. (2013) showed that being in a codeswitching context versus a single language context does not change the non-selective nature of language comprehension during sentence reading. We have furthered this finding by showing that semantic access during sentence reading is unaffected by encountering a codeswitch. These data indicate that access to semantics occurs independently of and sometimes even prior to accessing language membership, at least for early balanced bilingual adults.

4.4. P600, LCP and Other Positivities

There has been a confluence of late positive modulations reported in the sentence comprehension literature. These positivities have been observed in response to a wide variety of stimuli, including semantic violations (Chaire 2013; Wicha et al. 2004), syntactic or morphosyntactic-agreement violations (Barber and Carreiras 2005; Osterhout and Holcomb 1992; Osterhout and Mobley 1995), codeswitches (Moreno et al. 2002), and changes in form (Kutas and Hillyard 1980a). Importantly, these late positivities are also observed in the absence of an overt violation of meaning or grammar and may instead reflect continued processing of unexpected stimuli (Leckey and Federmeier 2020; Van Petten and Luka 2012). Because of the significant overlap in time, occurring around 500–900 ms after stimulus onset, and distribution over posterior scalp electrodes, it has been difficult to determine if these positivities are neurologically and functionally distinct or similar events. In turn, these late positivities are often not treated separately in the literature. However, there have been recent attempts to distinguish them (Leckey and Federmeier 2020; Van Petten and Luka 2012).

The P600 was first described as a brain response specific to syntactic violations (Osterhout and Holcomb 1992; Hagoort et al. 1993) but has since been observed in response to non-syntactic, even non-linguistic, stimuli and does not even require an overt violation (Van Petten and Luka 2012). Nevertheless, although the P600 may not be a response specific to grammar violations, encountering syntactic difficulties in a sentence context (but not in isolated word pairs; Barber and Carreiras 2005) reliably elicits this P600 response. One interpretation of this syntactic P600 is that it reflects revisiting, reanalysis or reprocessing after an anomaly has been detected (Vaughan-Evans et al. 2020; Leckey and Federmeier 2020; Kuperberg 2007; Van Petten and Luka 2012; Brouwer et al. 2017), for integration of the word within the broader sentence context (Salmon and Pratt 2002). Importantly, given that the syntactic P600 is sensitive to stimulus features, such as task relevance (e.g., making grammaticality judgments) and probability (Coulson et al. 1998b, cf Osterhout and Mobley 1995), it has been argued that this ERP component may be related to the well-studied domain-general P300 (specifically the P3b), which is associated with context or memory updating (Coulson et al. 1998a, 1998b; Donchin 1981; Verleger 1988; Gunter et al. 1997; Osterhout and Mobley 1995; Van Petten and Luka 2012).

Semantic anomalies can also elicit a late positivity, which often but not always follows the N400 (Kuperberg 2007; Van Petten and Luka 2012). This positivity, referred to as the semantic P600, post-N400 positivity or late positive component (LPC) (Leckey and Federmeier 2020; Van Petten and Luka 2012), appears in response to incongruous, overt violations of meaning, like the ones used in this study, and also to less overt manipulations, such as thematic role violations (e.g., "For breakfast, the eggs would only eat toast and jam"; Kuperberg et al. 2003). The semantic P600 does not tend to appear for sentence continuations that are unexpected but semantically plausible[4] and has been explained by similar processes as the syntactic P600, reflecting reprocessing or reanalysis after encountering a violation of meaning (Kuperberg 2007). Recent evidence in older adults, however, suggests that the semantic P600 may be separable from the P600 driven by syntactic errors (Leckey and Federmeier 2020). For example, the P600 can begin earlier for syntactic errors than to semantic incongruities and has even been described as having two separate phases, with an earlier one sensitive to syntactic information and the later one sensitive to both syntax and semantics (Wicha et al. 2004; Guajardo and Wicha 2014; Barber and Carreiras 2005).

The codeswitching positivity overlaps in time and distribution with the semantic and syntactic P600s (see Table 1) (Litcofsky and Van Hell 2017; Moreno et al. 2002; Van Der Meij et al. 2011). This codeswitching positivity has also been associated with sentence reanalysis, similar to the P600 (Litcofsky and Van Hell 2017), or as indexing the unexpected nature of the switched word (Moreno et al. 2002). However, it is unclear if and how the codeswitching positivity commonly elicited to language switches in a sentence context relates to either of these P600 responses. We explore below how our data informs what kind of effect this codeswitching positivity might be.

4.5. The Codeswitching Positivity vs. the Semantic P600

Semantic incongruities in the current study elicited a typical semantic P600 from 500–800 ms post-stimulus onset, consistent with previous findings where overt semantic incongruities are presented (Chaire 2013; Wicha et al. 2004). Codeswitched nouns also elicited a late positivity that started earlier than the semantic P600, but overlapped with the semantic P600 in distribution, maximal over medial posterior electrodes. Critically, semantic violations that were also codeswitched did not elicit any additional modulation of the LPC/semantic P600 (Figure 4), indicating that these positivities are not additive. The larger amplitude for the semantic P600 than the codeswitch positivity may indicate that reprocessing is more difficult for meaning violations than for language switches. The ease of integrating a codeswitch is in line with studies showing that the processing cost for a codeswitch can be eliminated under natural circumstances and meaning is obtained without delay for codeswitches (Chan et al. 1983; Gullifer et al. 2013). In theory, this would be facilitated if words from both languages were activated during word recognition (Caramazza and Brones 1979; Dijkstra and Van Heuven 2002), which has previously been demonstrated even in single language sentence contexts (Duyck et al. 2007; Schwartz and Kroll 2006; Titone et al. 2011; Van Assche et al. 2011; Van Assche et al. 2012). Words from both languages may be activated and considered within the sentence context – resulting in faster integration of codeswitches than semantic violations.

Does this mean that the codeswitching LPC and the semantic P600 are the same positivity? This is not clear. We found that proficiency in the language of the code-switched word (English) modulated the amplitude of the semantic P600 as well as the codeswitching positivity, specifically the later LPCb but not the earlier LPCa. As proficiency in English (codeswitch) increased, the amplitude of the positivity to both a codeswitch or a semantic incongruity increased. This might suggest that this later positivity relates to similar sentence reprocessing for both semantic anomalies and unexpected codeswitches. Yet, the direction of this modulation is curious. One might expect that better proficiency in a language would elicit easier reprocessing and therefore reduced positivity, opposite of what we observed. In fact, both Moreno et al. (2002) and Litcofsky and Van Hell (2017) observed exactly this. As proficiency in the switched language increased the codeswitching positivity to the target word decreased.[5] Similarly, Ruigendijk et al. also observed larger codeswitching positivity in less proficient learners of a language (Ruigendijk et al. 2016), although Liao and Chan (2016) observed a codeswitching positivity in both directions for bilingual switching into their more proficiency and into their less proficient languages.

Litcofsky and Van Hell (2017) measured the effect of proficiency in a highly fluent, but less balanced Spanish–English population, most of whom codeswitched habitually. They observed an asymmetrical effect, with a codeswitching positivity only when switching into the weaker language. They inferred that the weaker language may entail more effortful processing and sentence-level restructuring after encountering a switch. This interpretation was corroborated by EEG analysis in the time-frequency domain, which showed a decrease in lower beta band oscillations (15–18 Hz) only when switching into the weaker language. Decreases in the lower beta band have been observed to sentence-level processing, such as in response to syntactic violations. In contrast, they observed power increases in theta band oscillations (4–7 Hz), related to lexico-semantic processing and lexical inhibition, when switching into the dominant language. They suggested that switching into a lower proficiency language might require effortful sentence-level restructuring and shifting between languages, while switching into the dominant language results in a release of L1 inhibition at the lexical-semantic level that was previously required to process sentences in the weaker language (Litcofsky and Van Hell 2017; Van Hell et al. 2018).

These inconsistencies in the effect of language proficiency across studies is puzzling. Given that our sample was balanced in proficiency across their languages, it is possible that the effects of proficiency in our data are tenuously linked to small fluctuations in proficiency. However, a different proficiency effect, the reduction in N400 latency with increased proficiency, is consistent with previous findings. Perhaps instead our proficiency effect reflects suppression of a stronger language; greater suppression would be needed as proficiency increases, eliciting larger LPC amplitude as an index of access to this inhibited language.

4.6. What Is the Codeswitching Positivity?

The behavioral literature shows a processing cost for reading a codeswitch (Macnamara and Kushnir 1971). This delay may be reflected in the codeswitching positivity. What is the codeswitching positivity?

Some clues can be gleaned from outside the language literature. It has been suggested that the codeswitching positivity may reflect task-set reconfiguration (Moreno et al. 2008). During non-language task switching, a sustained differential positivity (D-Pos) peaks approximately 300 to 400 ms following a cue to switch between two tasks relative to a cue not to switch (Karayanidis et al. 2003; Nicholson et al. 2005). This differential positivity is maximal over central-parietal sites, has been localized to the superior parietal cortex, and is associated with activating task rules during reconfiguration to a new task (Jamadar et al. 2009). D-Pos is followed by activation in the dorsolateral prefrontal cortex, a region postulated to initiate a cortical-subcortical loop involved in switching between both languages and non-language tasks (Abutalebi and Green 2008). A similar component to D-Pos for codeswitches compared to non-switches has been identified during language production (Blackburn 2013). However, activity in the dorsolateral prefrontal cortex following D-Pos might play a larger role during production, as it is not observed during comprehension of a codeswitch, at least not in the absence of a processing cost when reading short two-word sentences with switches (Phillips and Pylkkanen 2019).

The code-switching positivity observed here is similar to D-Pos in time-course and distribution and may reflect a language reconfiguration process in order to switch between languages. We suggest based on evidence that motor representations are modeled during auditory comprehension (Bangert et al. 2006), that the code-switching positivity may reflect activation of the motor representation for switching between languages. Two distinct processing streams have been implicated in auditory processing: a ventral pathway involved in mapping input to meaning and a dorsal sensorimotor pathway that enables mapping of the input to articulatory motor representations (see also Friederici and Gierhan 2013; Hickok and Poeppel 2007; Poeppel et al. 2012). It is possible that semantic processing occurs via the ventral stream and is reflected by the N400 and subsequent semantic P600. In contrast, codeswitch processing may occur along the dorsal pathway, resulting in a codeswitching positivity that reflects mapping of the input to the motor output necessary to articulate a switch into another language.

To test this prediction, factors known to increase D-Pos and/or articulation difficulty, such as stimuli that evoke both tasks (e.g., false cognates or words with similar cross-language orthography but different pronunciation) can be manipulated in a bilingual sentence reading paradigm. Ongoing research is investigating how code-switching experience modulates both D-Pos and the codeswitching positivity, to determine if these components are functionally similar.

Another possibility is that the codeswitching positivity is similar to a P600 or LPCa elicited in response to a syntactic violation (Hagoort et al. 1993; Osterhout and Holcomb 1992; cf. Van Der Meij et al. 2011). Codeswitches elicit a pattern similar to the LAN-P600 pattern commonly obtained for syntactic violations (Barber and Carreiras 2005; Osterhout and Mobley 1995). Similarly, processing sentential codeswitches has elicited neural oscillations at a frequency comparable to those elicited to syntactic violations (Litcofsky and Van Hell 2017). It has been suggested that the LPC may in fact represent two processes with two separate generators: a syntactic early phase and a later integration phase (Barber and Carreiras 2005). The differential positivity for codeswitches versus non-switches was only significant during the first half of the LPC. Although codeswitches do not necessarily create syntactic anomalies, perhaps the codeswitches are processed initially as syntactic anomalies. This explanation may be in line with the three processing steps proposed by

Friederici and Kotz (2003): an initial structure building phase reflected by the LAN elicited to a codeswitch, semantic integration reflected by the N400, and a late syntactic integration phase reflected by the P600/LPC.

Finally, it has been suggested that both the semantic P600 and the codeswitching positivity are variants of a late P3 component (Van Der Meij et al. 2011; but see Leckey and Federmeier 2020 who argue that the semantic and syntactic effects are not the same). The LPC is often interpreted as a P3 which increases in amplitude for anomalous stimuli (Coulson et al. 1998b). Nieuwenhuis et al. (2005) proposed that late positivities such as the P3 may reflect phasic activity of the locus coeruleus-noradrenergic system, which works to increase sensitivity of neurons to incoming stimuli and to focus attention (Nieuwenhuis et al. 2005). Locus coeruleus activity is driven by task-relevant decision-making processes, i.e., identification of task-relevant stimuli and mapping these to appropriate responses, and is typically amplified in response to stimuli that are relevant to the current task.

According to this hypothesis, individuals have an internal model of the external world or current task structure, e.g., reading sentences in Spanish. When the current expectations are violated by a stimulus that does not match the task, e.g., an English word or a word that violates semantic expectations built by the sentence context, the locus coeruleus responds with phasic firing and a large P3 results. Increased locus coeruleus phasic firing facilitates responding to the outcome of the decision making process and results in less distraction by irrelevant stimuli and better performance on the task (Nieuwenhuis et al. 2005).

The positivities observed to codeswitches and semantic anomalies may reflect the brain activity evoked by the need to update the context in light of an unexpected word and may reflect both attention and decision making processes associated with this context updating (Polich 2007; Van Der Meij et al. 2011). If this is the case, both the codeswitching positivity and the semantic P600 may involve activation of a general anomaly detector and all or part of these positivities may involve the same generator (Barber and Carreiras 2005; Coulson et al. 1998b; Van Der Meij et al. 2011). This possibility is not entirely different from the task reconfiguration account described above, as the P3 is sensitive to task relevance (Coulson et al. 1998a) and D-Pos may also be a variant of the P3.

4.7. Code Switching Experience Modulates the Code Switching Positivity

Previous studies have found that increased code-switching experience in daily life reduces the response time cost of producing a code-switch (Prior and Gollan 2011). Of equal importance is the fact that codeswitches may be unexpected by the interlocutor and are known to incur a cost in comprehension time. A major question is whether this processing cost is attenuated with code-switching experience. We tested whether individual differences in codeswitching behavior tune neural sensitivity to processing a sentential codeswitch.

A correlation between how frequently an individual codeswitches and the amplitude of the codeswitching positivity was observed. Specifically, increased switching also increased the amplitude of the codeswitch positivity, suggesting that bilinguals who codeswitch more often in daily life may dedicate more resources to processing codeswitches. In contrast, there was no correlation between code-switching experience and semantic retrieval during the N400. This effect of codeswitching experience was specific to processing the language switch itself.

A larger codeswitching positivity for individuals who are more experienced with codeswitching may appear counterintuitive given findings that expertise often results in more efficient processing (e.g., during motor learning, see Landau and D'Esposito 2006). However, both the interpretation that the codeswitching positivity may be a form of a P3 or P600 and that it may reflect a task-decision stage of processing, leads to a prediction that the positivity would be positively correlated with experience. First, the P600 amplitude is positively correlated with proficiency and behavioral sensitivity to syntactic anomalies (McLaughlin et al. 2010). Thus, code-switching experience may increase sensitivity to a codeswitch. Second, the P3 amplitude is linked to attention control and task performance and is susceptible to both developmental change and individual differences (Johnstone et al. 1996; see Kok 2001 for a review; Ladish and Polich 1989). Larger P3 amplitudes are linked to more accurate and faster responses (Nieuwenhuis et al. 2011), and the P3 amplitude is larger in children with better capacity allocation—the ability to invest resources in a task (Jonkman et al. 2000). Wright et al. (2013) found that the P3 amplitude is larger during task performance for individuals who are experts at performing that task. They concluded that experts may engage selective attention more effectively during the task. Thus, if the codeswitching positivity is related to the P3 family of positivities, then this increase in the amplitude with more experience/expertise with codeswitching would be consistent with what is observed for P3 amplitude.

Similarly, individuals with experience codeswitching may be more inclined to allocate mental capacity (resources) into processing a codeswitch and may be more attentive to codeswitches than those less likely to encounter codeswitching in their daily lives. As discussed above, the enhanced P3 in daily codeswitchers may reflect an increased phasic locus coeruleus response which serves to focus attention and facilitate decision making processes (e.g., integration of the word into the sentence). The codeswitching positivity has also been likened to the positivity elicited during task-set reconfiguration, D-Pos (Moreno et al. 2008). We are currently testing the prediction that as code-switching experience increases, D-Pos elicited during both language and non-language task switching will also increase. In addition, these findings open the door to further studies linking this neurocognitive research paradigm with sociolinguistic research tradition. For instance, we are interested in testing differences in the code-switching positivity for habitual code-switchers when they encounter a contextually predictable codeswitch.

It should be noted that we operationalized codeswitching frequency using the ACSES survey, which measures switching within conversations (i.e., being in a dense codeswitching context). Bilinguals were not categorized according to their language environment, as has been the custom in other studies (for a review, see Blackburn 2018). Thus, we were not able to distinguish between bilinguals who refrain from switching because they spend time in environments where only one language is appropriate (single-language context) and those who switch frequently across conversations but not within conversations (dual-language switching) (Green and Abutalebi 2013). Hofweber et al. (2016) suggested that bilinguals who spend time in a dense codeswitching context should exhibit increased monitoring skills and control when processing language switches, as a result of often having to be ready to anticipate a language switch. Further research is necessary to determine if the codeswitching positivity differs for bilinguals who spend time in a dual-language context when both languages are appropriate vs. those who spend time in a single-language context, or according to the specific type of codeswitching bilinguals encounter. Nonetheless, the findings herein validate ACSES and show that this bilingual phenomenon is independent of other measures, such as language proficiency.

5. Conclusions

In conclusion, processing of a codeswitch does not appear to interfere with semantic retrieval. Rather, in support of models such as BIA+, our findings indicate that identification of language membership and semantic retrieval occur in parallel and likely involve separate generators. A codeswitching positivity, which overlaps in time with the N400 to a semantic violation, was observed following a LAN to switches compared to non-switches. A late positivity with similar distribution and overlapping time course to the codeswitching positivity was observed for semantic incongruities relative to congruities. The co-occurrence of a codeswitch and semantic violation resulted in a positivity that did not differ substantially from the positivity elicited to each individually. It is possible that these positivities reflect a general P3-like anomaly detector and reprocessing due to the unexpected event. Alternatively, the codeswitch positivity may reflect reconfiguration of the language at this later reprocessing stage, or less likely, the treatment of a code-switch as a syntactic violation in line with a syntactic P600. Codeswitching experience was found to modulate the codeswitching positivity, much like expertise and increased attention have been found to increase the amplitude of the P3 component. While further investigation is needed to elucidate the generators of the codeswitching positivity and the semantic P600, it is clear that the codeswitching positivity reflects a process distinct from and parallel to semantic retrieval, and that this process is affected by experience with codeswitching in daily life.

Author Contributions: Conceptualization, N.Y.Y.W. and A.M.B.; methodology, N.Y.Y.W. and A Blackburn; formal analysis, A.M.B. and N.Y.Y.W.; investigation, A.M.B.; resources, N.Y.Y.W.; data curation, A.M.B. and N.Y.Y.W.; writing—original draft preparation, A.M.B.; writing—review and editing, N.Y.Y.W.; supervision, N.Y.Y.W.; funding acquisition, N.Y.Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Eunice Kennedy Shriver National Institute of Child Health & Human Development and National Institute of General Medical Sciences through award SC1HD060435 and by the UTSA College of Science through funds awarded to N.Y.Y.W. Computational support was provided by the UTSA Computational System Biology Core, funded by the National Institute on Minority Health and Health Disparities through award G12MD007591. A.M.B. was supported in part by the Texas A&M International University 2020–2021 University Research Grant.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of University of Texas at San Antonio (protocol #06-201; 2006).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1 We present our research as it fits in the current body of literature. However, it should be noted that our original hypotheses were proposed and the research was conducted prior to the literature presented in Table 1 (Blackburn 2013).

2 Semantic violations with a codeswitch did not elicit an LPC in every electrode. Only electrodes in which a positivity was elicited were included in the onset calculation. N.B. Due to local maxima/minima in individual data, onset latencies were calculated on the grand averaged difference waves. Therefore, these values do not contain individual subject variance and do not allow for statistical comparisons.

3 We refer to the LPCa and LPCb elicited in response to a semantic violation here as the semantic positivity.

4 Failed semantic predictions have been reported to elicit a more frontally distributed positivity (Federmeier 2007).

5 Van Der Meij et al. (2011) found that as proficiency in the language of the sentence context increased the positivity in response to a codeswitched word also increased.

References

Abutalebi, Jubin, and David W. Green. 2008. Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes* 23: 557–82. [CrossRef]

Appel, René, and Pieter Muysken. 2005. Language use in the bilingual community. In *Language Contact and Bilingualism*. Amsterdam: Amsterdam Academic Archive.

Bangert, Marc, Thomas Peschel, Gottfried Schlaug, Michael Rotte, Dieter Drescher, Herman Hinrichs, Hans-Jochen Heinze, and Eckart Altenmüller. 2006. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage* 30: 917–26. [CrossRef] [PubMed]

Barber, Horacio, and Manuel Carreiras. 2005. Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience* 17: 137–53. [CrossRef] [PubMed]

Beatty-Martínez, Anne L., Jorge R. Valdés, and Paola E. Dussias. 2018. From the field to the lab: A converging methods approach to the study of codeswitching. *Language* 3: 19. [CrossRef]

Blackburn, Angelique 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. Ph.D. Thesis, University of Texas at San Antonio, San Antonio, TX, USA.

Blackburn, Angelique M. 2017. The Bilingual Brain. In *An Introduction to Bilingualism: Principles and Processes*, 2nd ed. Edited by Jeanette Altarriba and Roberto R. Heredia. New York: Psychology Press, pp. 107–38.
Blackburn, Angelique M. 2018. Cognitive impact of bilingualism and language habits at the borders of cultures and nations. In *Inquiries into Literacy Learning and Intercultural Competency in a World of Border Tensions*. Edited by Tonya Huber and Philip S. Roberson. Charlotte: Information Age Publishing, pp. 141–66.
Blackburn, Angelique M. 2019. Factors Affecting Cortical Representation. In *The Handbook of the Neuroscience of Multilingualism*. Edited by John W. Schwieter. Hoboken: Wiley, pp. 280–306.
Bornkessel-Schlesewsky, Ina, and Matthias Schlesewsky. 2008. An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews* 59: 55–73. [CrossRef] [PubMed]
Brouwer, Harm, Matthew W. Crocker, Noortje J. Venhuizen, and John C. J. Hoeks. 2017. A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science* 41: 1318–52. [CrossRef] [PubMed]
Caramazza, Alfonso, and Isabel Brones. 1979. Lexical access in bilinguals. *Bulletin of the Psychonomics Society* 13: 212–14. [CrossRef]
Chaire, Alondra. 2013. Morpho-syntax and the aging brain: An ERP Study of Sentence Comprehension in Older Adult Spanish Speakers. Master's Thesis, University of Texas at San Antonio, San Antonio, TX, USA.
Chan, Mun-Chee, Helen L. H. Chau, and Rumjahn Hoosain. 1983. Input/output switch in bilingual codeswitching. *Journal of Psycholinguistic Research* 12: 407–16. [CrossRef]
Coulson, Seana, Jonathan W. King, and Marta Kutas. 1998a. ERPs and Domain Specificity: Beating a Straw Horse. *Language and Cognitive Processes* 13: 653–72. [CrossRef]
Coulson, Seana, Jonathan W. King, and Marta Kutas. 1998b. Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13: 21–58. [CrossRef]
Davis, Colin J. 2005. N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods* 37: 65–70. [CrossRef] [PubMed]
Davis, Colin J., and Manuel Perea. 2005. A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods* 37: 665–71. [CrossRef]
Dewaele, Jean-Marc. 2004. The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development* 25: 204–22. [CrossRef]
Dijkstra, Ton, and Walter J. B. Van Heuven. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5: 175–97. [CrossRef]
Donchin, Emanuel. 1981. Surprise! . . . Surprise? *Psychophysiology* 18: 493–513. [CrossRef]
Dussias, Paola E. 2001. Psycholinguistic complexity in codeswitching. *International Journal of Bilingualism* 5: 87–100. [CrossRef]
Duyck, Wouter, Eva Van Assche, Denis Drieghe, and Robert J. Hartsuiker. 2007. Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33: 663–79. [CrossRef] [PubMed]
Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48: 71–99. [CrossRef]
Federmeier, Kara D. 2007. Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* 44: 491–505. [CrossRef]
Fernandez, Carla B., Kaitlyn A. Litcofsky, and Janet G. Van Hell. 2019. Neural correlates of intra-sentential code-switching in the auditory modality. *Journal of Neurolinguistics* 51: 17–41. [CrossRef]
Friederici, Angela D. 1995. The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain Lang* 50: 259–81. [CrossRef]
Friederici, Angela D. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6: 78–84. [CrossRef]
Friederici, Angela D., and Sarah M. E. Gierhan. 2013. The language network. *Current Opinion in Neurobiology* 23: 250–54. [CrossRef]
Friederici, Angela D., and Sonja A. Kotz. 2003. The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage* 20: S8–S17. [CrossRef]
Gonzalez-Reigosa, Fernando. 1973. The Anxiety-Arousing Effect of Taboo Words in Bilinguals. Ph.D. Thesis, The Florida State University, Tallahassee, FL, USA.
Gosselin, Leah, and Laura Sabourin. 2021. Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia* 159: 107922. [CrossRef]
Grainger, Jonathan, and J. Kevin O'regan. 1992. A psychophysical investigation of language priming effects in two english-french bwguals. *European Journal of Cognitive Psychology* 4: 323–39. [CrossRef]
Grainger, Jonathan, and Cecile Beauvillain. 1987. Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology* 39: 295–319. [CrossRef]
Grainger, Jonathan, and C. Beauvillain. 1988. Associative priming in bilinguals: Some limits of interlingual facilitation effects. *Canadian Journal of Psychology/Revue Canadienne de Psychologie* 42: 261. [CrossRef]
Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef]
Green, David W., and Li Wei. 2014. A control process model of code-switching. *Language, Cognition and Neuroscience* 29: 499–511. [CrossRef]

Guajardo, Lourdes, and Nicole Y. Wicha. 2014. Morphosyntax can modulate the N400 component: Event related potentials to gender-marked post-nominal adjectives. *NeuroImage* 91: 262–72. [CrossRef]

Gullifer, Jason W., Judith F. Kroll, and Paola E. Dussias. 2013. When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology* 4: 278. [CrossRef]

Gumperz, John J., and Eduardo Hernandez-Chavez. 1972. Bilingualism, Bidialectalism, and Classroom Interaction. In *Functions of Language in the Classroom*. Edited by Courtney B. Cazden, Vera P. John and Dell Hymes. New York: Teachers College Press.

Gunter, Thomas C., Laurie A. Stowe, and Gusbertus Mulder. 1997. When syntax meets semantics. *Psychophysiology* 34: 660–76. [CrossRef]

Hagoort, Peter, Colin Brown, and Jolanda Groothusen. 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes* 8: 439–83. [CrossRef]

Harris, Catherine L. 2004. Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development* 25: 223–47. [CrossRef]

Hernández, Mireia, Clara D. Martin, Francisco Barcelo, and Albert Costa. 2013. Where is the bilingual advantage in task-switching? *Journal of Memory and Language* 69: 257–76. [CrossRef]

Hickok, Gregory, and David Poeppel. 2007. The cortical organization of speech processing. *Nature Reviews/Neuroscience* 8: 393–402. [CrossRef]

Hofweber, Julia, Theodoros Marinis, and Jeanine Treffers-Daller. 2016. Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism* 6: 648–68. [CrossRef]

Hoversten, Liv J., and Matthew J. Traxler. 2020. Zooming in on zooming out: Partial selectivity and dynamic tuning of bilingual language control during reading. *Cognition* 195: 104118. [CrossRef]

Hoversten, Liv J., Trevor Brothers, Tamara Y. Swaab, and Matthew J. Traxler. 2015. Language membership identification precedes semantic access: Suppression during bilingual word recognition. *Journal of Cognitive Neuroscience* 27: 2108–16. [CrossRef]

Hoversten, Liv J., Trevor Brothers, Tamara Y. Swaab, and Matthew J. Traxler. 2017. Early processing of orthographic language membership information in bilingual visual word recognition: Evidence from ERPs. *Neuropsychologia* 103: 183–90. [CrossRef] [PubMed]

Ibáñez, Antonio J., Pedro Macizo, and M. Teresa Bajo. 2010. Language access and language selection in professional translators. *Acta Psychologica* 135: 257–66. [CrossRef] [PubMed]

Jamadar, Sharna, Alexander Provost, W. Ross Fulham, Patricia T. Michie, and Frini Karayanidis. 2009. Multiple sources underlie ERP indices of task-switching. In *ASCS09: Proceedings of the 9th Conference of the Australasiam Society for Cognitive Science*. Edited by Wayne Christensen, Elizabeth Schier and John Sutton. Sydney: Macquarie Centre for Cognitive Science, pp. 154–61.

Johnstone, Stuart J., Robert J. Barry, John W. Anderson, and Stephen F. Coyle. 1996. Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *International Journal of Psychophysiology* 24: 223–38. [CrossRef]

Jonkman, Lisa, C. Kemner, M. N. Verbaten, H. Van Engeland, G. Cafferman, Jan Buitelaar, and H. S. Koelega. 2000. Attentional capacity, a probe ERP study: Differences between children with attention-deficit hyperactivity disorder and normal control children and effects of methylphenidate. *Psychophysiology* 37: 334–46. [CrossRef] [PubMed]

Kaan, Edith, Souad Kheder, Ann Kreidler, Aleksandra Tomić, and Jorge R. Valdés Kroff. 2020. Processing code-switches in the presence of others: An ERP study. *Frontiers in Psychology* 11: 1288. [CrossRef]

Kang, Carissa, Gita Martohardjono, and Barbara Lust. 2015. Linguistic code-switching affects executive function in some bilinguals: First, First, results from a new methodology. Paper presented at the Bilingualism and Executive Function: An Interdisciplinary Approach, New York, NY, USA, April 30.

Karayanidis, Frini, Max Coltheart, Patricia T. Michie, and Karen Murphy. 2003. Electrophysiological correlates of anticipatory and poststimulus components of task switching. *Psychophysiology* 40: 329–48. [CrossRef]

Kheder, Souad, and Edith Kaan. 2016. Processing Code-Switching in Algerian Bilinguals: Effects of Language Use and Semantic Expectancy. *Frontiers in Psychology* 7: 248. [CrossRef]

Kirsner, Kim, Marilyn C. Smith, R. S. Lockhart, M. L. King, and M. Jain. 1984. The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior* 23: 519–39. [CrossRef]

Kluender, Robert, and Marta Kutas. 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5: 196–214. [CrossRef]

Kok, Albert. 2001. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38: 557–77. [CrossRef] [PubMed]

Kolers, Paul A. 1966. Reading and Talking Bilingually. *The American Journal of Psychology* 79: 357–76. [CrossRef] [PubMed]

Kuperberg, Gina R. 2007. Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research* 1146: 23–49. [CrossRef] [PubMed]

Kuperberg, Gina R. Tatiana Sitnikova, David Caplan, and Phillip J. Holcomb. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17: 117–29. [CrossRef]

Kutas, Marta, Eva M. Moreno, and Nicole Y. Y. Wicha. 2009. *The Cambridge Handbook of Linguistic Code-Switching Cambridge*. Edited by Barbara E. Bullock and Almeida Jacqueline Ed Toribio. Cambridge: Cambridge University Press.

Kutas, Marta, and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62: 621–47. [CrossRef]

Kutas, Marta, and Steven A. Hillyard. 1980a. Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology* 11: 99–116. [CrossRef]

Kutas, Marta, and Steven A. Hillyard. 1980b. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207: 203–5. [CrossRef]

Ladish, Christine, and John Polich. 1989. P300 and probability in children. *Journal of Experimental Child Psychology* 48: 212–23. [CrossRef]

Landau, Susan M., and Mark D'Esposito. 2006. Sequence learning in pianists and nonpianists: An fMRI study of motor expertise. *Cognitive, Affective, & Behavioral Neuroscience* 6: 246–59.

Leckey, Michelle, and Kara D. Federmeier. 2020. The P3b and P600(s): Positive contributions to language comprehension. *Psychophysiology* 57: e13351. [CrossRef]

Liao, Chia-Hsuan, and Shiao-Hui Chan. 2016. Direction matters: Event-related brain potentials reflect extra processing costs in switching from the dominant to the less dominant language. *Journal of Neurolinguistics* 40: 79–97. [CrossRef]

Litcofsky, Kaitlyn A., and Janet G. Van Hell. 2017. Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97: 112–39. [CrossRef] [PubMed]

Macnamara, John, and Seymour L. Kushnir. 1971. Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior* 10: 480–87. [CrossRef]

Marian, Viorica, and Margarita Kaushanskaya. 2008. Words, feelings, and bilingualism. *The Mental Lexicon* 3: 72–90. [CrossRef]

Marian, Viorica, Henrike K. Blumenfeld, and Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language Profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50: 940–67. [CrossRef]

McLaughlin, Judith, Darren Tanner, Ilona Pitkänen, Cheryl Frenck-Mestre, Kayo Inoue, Geoffrey Valentine, and Lee Osterhout. 2010. Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning* 60: 123–50. [CrossRef]

Molinaro, Nicola, Horacio A. Barber, and Manuel Carreiras. 2011. Grammatical agreement processing in reading: ERP findings and future directions. *Cortex* 47: 908–30. [CrossRef]

Morales, Julia, Carolina Yudes, Carlos J. Gómez-Ariza, and M. Teresa Bajo. 2015. Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia* 66: 157–69. [CrossRef]

Moreno, Eva M., Antoni Rodríguez-Fornells, and Matti Laine. 2008. Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics* 21: 477–508. [CrossRef]

Moreno, Eva M., Kara D. Federmeier, and Marta Kutas. 2002. Switching Languages, Switching Palabras (Words): An electrophysiological study of codeswitching. *Brain Lang* 80: 188–207. [CrossRef]

Münte, Thomas F., Andras Szentkuti, Bernardina M. Wieringa, Mike Matzke, and Sönke Johannes. 1997a. Human brain potentials to reading syntactic errors in sentences of different complexity. *Neuroscience Letters* 235: 105–8. [CrossRef]

Münte, Thomas F., Mike Matzke, and Sönke Johannes. 1997b. Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience* 9: 318–29. [CrossRef] [PubMed]

Myers-Scotton, Carol. 1993. *Social Motivations for Codeswitching*. Oxford: Clarendon Press.

Ng, Shukhan, and Nicole Y.Y. Wicha. 2013. Meaning first: A case for language-independent access to word meaning in the bilingual brain. *Neuropsychologia* 51: 850–63. [CrossRef] [PubMed]

Ng, Shukhan, Christian Gonzalez, and Nicole Y. Y. Wicha. 2014. The fox and the cabra: An ERP analysis of reading codeswitched nouns and verbs in bilingual short stories. *Brain Research* 1557: 127–40. [CrossRef] [PubMed]

Nicholson, Rebecca, Frini Karayanidis, Dane Poboka, Andrew Heathcote, and Patricia T. Michie. 2005. Electrophysiological correlates of anticipatory task-switching processes. *Psychophysiology* 42: 540–54. [CrossRef]

Nieuwenhuis, Sander, Gary Aston-Jones, and Jonathan D. Cohen. 2005. Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin* 131: 510–32. [CrossRef]

Nieuwenhuis, Sander, Eco J. De Geus, and Gary Aston-Jones. 2011. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* 48: 162–75. [CrossRef]

Osterhout, Lee, and Linda A. Mobley. 1995. Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34: 739–73. [CrossRef]

Osterhout, Lee, and Philip Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31: 785–806. [CrossRef]

Pavlenko, Aneta. 2005. *Emotions and Multilingualism*. Cambridge: Cambridge University Press.

Pavlenko, Aneta. 2008. Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism Language and Cognition* 11: 147. [CrossRef]

Phillips, Sarah F., and Liina Pylkkanen. 2019. Are Neural Effects of Composition Sensitive to Code-Switching? Paper presented at the Society for the Neurobiology of Language Conference, Helsinki, Finland, August 23.

Poeppel, David, Karen Emmorey, Gregory Hickok, and Liina Pylkkänen. 2012. Towards a new neurobiology of language. *Journal of Neuroscience* 32: 14125–31. [CrossRef]

Polich, John. 2007. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118: 2128–48. [CrossRef] [PubMed]

Prior, Anat, and Tamar H. Gollan. 2011. Good language-switchers are good task switchers: Evidence from Spanish–English and Mandarin-English bilinguals. *Journal of the International Neuropsychological Society* 17: 1–10. [CrossRef] [PubMed]

Proverbio, Alice Mado, Giuliana Leon, and Alberto Zani. 2004. Language switching mechanisms in simultaneous interpreters: An ERP study. *Neuropsychologia* 42: 1636–56. [CrossRef]

Ruigendijk, Esther, Gerd Hentschel, and Jan Patrick Zeller. 2016. How L2-learners' brains react to code-switches: An ERP study with Russian learners of German. *Second, Language Research* 32: 197–223. [CrossRef]

Salmon, Nirit, and Hillel Pratt. 2002. A comparison of sentence- and discourse-level semantic processing: An ERP study. *Brain Lang* 83: 367–83. [CrossRef]

Schwartz, Ana I., and Judith F. Kroll. 2006. Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55: 197–212. [CrossRef]

Soveri, Anna, Antoni Rodriguez-Fornells, and Matti Laine. 2011. Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology* 2: 183. [CrossRef]

Titone, Debra, Maya Libben, Julie Mercier, Veronica Whitford, and Irina Pivneva. 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: 1412–31. [CrossRef]

Treffers-Daller, Jeanine, Esther Ruigendijk, and Julia Hofweber. 2021. Editorial: Behavioral and Neurophysiological Approaches to Code-Switching and Language Switching. *Frontiers in Psychology* 12: 660695. [CrossRef]

Valdés Kroff, Jorge R., Paola E. Dussias, Chip Gerfen, Lauren Perroti, and M. Teresa Bajo. 2016. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism* 7: 163–198. [CrossRef]

Valdés Kroff, Jorge R., Patricia Román, and Paola E. Dussias. 2020. Are all code-switches processed alike? Examining semantic v. language unexpectancy. *Frontiers in Psychology* 11: 2138. [CrossRef] [PubMed]

Valdés Kroff, Jorge R., Rosa E. Guzzardo Tamargo, and Paola E. Dussias. 2017. Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches. *Linguistic Approaches to Bilingualism* 8: 98–133. [CrossRef]

Van Assche, Eva, Denis Drieghe, Wouter Duyck, Marijke Welvaert, and Robert J. Hartsuiker. 2011. The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language* 64: 88–107. [CrossRef]

Van Assche, Eva, Wouter Duyck, and Robert J. Hartsuiker. 2012. Bilingual word recognition in a sentence context. *Frontiers in Psychology* 3: 174. [CrossRef]

Van Der Meij, Maartje, Fernando Cuetos, Manuel Carreiras, and Horacio A. Barber. 2011. Electrophysiological correlates of language switching in second language learners. *Psychophysiology* 48: 44–54. [CrossRef]

Van Hell, Janet G., Carla B. Fernandez, Gerrit Jan Kootstra, Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2018. Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism* 8: 134–61. [CrossRef]

Van Hell, Janet G., Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2015. Sentential code-switching: Cognitive and neural approaches. In *The Cambridge Handbook of Bilingual Processing*. Edited by John W. Schwieter. Cambridge: Cambridge University Press, pp. 459–82.

Van Heuven, Walter J. B., and Ton Dijkstra. 2010. Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews* 64: 104–22. [CrossRef]

Van Petten, Cyma, and Barbara J. Luka. 2012. Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology* 83: 176–90. [CrossRef] [PubMed]

Vaughan-Evans, Awel, Maria Carmen Parafita Couto, Bastien Boutonnet, Noriko Hoshino, Peredur Webb-Davies, Margaret Deuchar, and Guillaume Thierry. 2020. Switchmate! An electrophysiological attempt to adjudicate between competing accounts of adjective-noun code-switching. *Frontiers in Psychology* 11: 549762. [CrossRef]

Verleger, Rolf. 1988. Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. *Behavioral and Brain Sciences* 11: 343–427. [CrossRef]

Wicha, Nicole Y. Y., Eva M. Moreno, and Marta Kutas. 2004. Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience* 16: 1272–88. [CrossRef]

Woodcock, Richard W. 1991. *Woodcock Language Proficiency Battery-Revised. English and Spanish Forms. Examiner's Manual*. Chicago: The Riverside Publishing Company.

Wright, Michael J., Fernand Gobet, Philippe Chassy, and Payal Nanik Ramchandani. 2013. ERP to chess stimuli reveal expert-novice differences in the amplitudes of N2 and P3 components. *Psychophysiology* 50: 1023–33. [CrossRef] [PubMed]

Wu, Yan Jing, and Guillaume Thierry. 2013. Fast Modulation of Executive Function by Language Context in Bilinguals. *The Journal of Neuroscience* 33: 13533–37. [CrossRef] [PubMed]

Yacovone, Anthony, Emily Moya, and Jesse Snedeker. 2021. Unexpected words or unexpected languages? Two ERP effects of code-switching in naturalistic discourse. *Cognition* 215: 104814. [CrossRef]

Yang, Hwajin, Andree Hartanto, and Sujin Yang. 2016. The Complex Nature of Bilinguals' Language Usage Modulates Task-Switching Outcomes. *Frontiers in Psychology* 7: 560. [CrossRef] [PubMed]

Yim, Odilia, and Ellen Bialystok. 2012. Degree of conversational code-switching enhances verbal task switching in Cantonese–English bilinguals. *Bilingualism: Language and Cognition* 15: 873–83. [CrossRef]

languages

MDPI

Article

Language Control and Intra-Sentential Codeswitching among Bilingual Children with and without Developmental Language Disorder

Aviva Soesman [1,2,*], Joel Walters [2] and Sveta Fichman [2]

[1] Division of Languages, Tel Aviv University, Tel Aviv-Yafo 6997801, Israel
[2] English Literature and Linguistics, Bar-Ilan University, Ramat Gan 5290002, Israel
* Correspondence: soesman@tauex.tau.ac.il

Abstract: The present study investigated bilingual language control among preschool children in a sentence repetition task containing unilingual stimuli and codeswitched stimuli within prepositional phrases (PPs). Cross-language errors, that is, codeswitches that were not part of the stimulus sentences, were taken as evidence of difficulties in language control. Specifically, we investigated cross-language errors as a function of stimulus sentence type (codeswitched or unilingual), CS site within the PP, directionality (English or Hebrew stimulus sentences), and group status (children with typical language development (TLD), and children with Developmental Language Disorder (DLD)). We also examined cross-language errors in terms of word class and locus in the sentence. The participants were 65 English (home language)–Hebrew (societal language) bilinguals with TLD and 13 with DLD, ages 5;5–6;10 (M = 5;11). Stimulus sentences contained five codeswitch conditions within prepositional phrases, for example, a codeswitched preposition (P) or a codeswitched preposition, determiner and noun (P+DET+N), and a 'no switch' condition. The stimuli were 36 English and 36 Hebrew sentences (+24 fillers) matched for semantic content and syntax. English sentences contained switches to Hebrew, and Hebrew sentences contained switches to English. The results showed more cross-language errors for codeswitched than unilingual sentence stimuli. The children with TLD showed a directionality effect, producing more cross-language errors in Hebrew sentence stimuli than in English, but the children with DLD did not. The children with DLD had more cross-language errors than their peers with TLD for English stimuli. Most cross-language errors appeared in the sentence-final, adverbial temporal phrase. Findings are discussed in terms of language co-activation and competition in order to account for the difference in performance on unilingual versus codeswitched stimuli and in light of sociopragmatic and psycholinguistic factors to account for the directionality effect among children with TLD and the lack thereof among children with DLD.

Keywords: bilingual children; intra-sentential codeswitching; language control; cross-language errors; Developmental Language Disorder (DLD)

Citation: Soesman, Aviva, Joel Walters, and Sveta Fichman. 2022. Language Control and Intra-Sentential Codeswitching among Bilingual Children with and without Developmental Language Disorder. *Languages* 7: 249. https://doi.org/10.3390/languages7040249

Academic Editor: John W. Schwieter

Received: 18 March 2022
Accepted: 14 September 2022
Published: 26 September 2022

1. Introduction

1.1. Language Control and Cross-Language Intrusions

This study examines bilingual language control among bilingual children as reflected in cross-language errors produced in a sentence repetition task. Language control enables the bilingual to monitor speech production, in particular, to use the intended language and avoid interference from the non-target language (Branzi et al. 2016). Within the field of bilingual language control, psycholinguistic studies on language switching feature prominently. These studies typically examine the effects and relative cost of processing mixed and unilingual language in experiments on single words or sentences (e.g., Costa and Santesteban 2004; Ruigendijk et al. 2016; Thomas and Allport 2000). Relatively few studies have specifically focused on the production of what has been called variously "unintended" language switches, "cross-language errors", "intrusions", or "intrusion errors."[1] The phenomenon can be described as a failure to control inappropriate language responses, for instance, a Dutch–English bilingual saying "en" instead of "and." The lack of studies on language control is especially evident in the field of child bilingualism.

Cross-language errors should not come as a complete surprise as it is now commonly accepted that even in monolingual contexts, a second language is never completely deactivated (e.g., Bobb and Wodniecka 2013; Costa 2005; Grosjean 2001; Kroll et al. 2014). There is, however, no agreement regarding the origin of these errors (Zheng et al. 2018). For example, Poulisse and Bongaerts (1994) view an unintended language switch as an instance where language users erroneously activate the L2 lemma in L1 speech based on a model where language selection takes place via the spread of activation to an L1 or L2 lexical item. Lipski (2016), in contrast, argues that unintended switching cannot be explained only by assigning the wrong language to individual lemmas but that longer stretches of morpho-syntactic units may also be activated in the non-target language due to lack of lexical access in a speaker's weaker language. Accordingly, less fluent speakers plan their speech in their stronger language, which "takes over" (p. 141) when a translation equivalent is not available. Alternatively, it has been suggested that intrusion errors may originate in erroneous language selection (at the conceptual level) in addition to erroneous word selection (at the lexical level) (Zheng et al. 2018; for an extensive discussion, see Declerck and Philipp 2015).

Language control abilities have been associated with domain-general, cognitive control, or executive functioning skills (e.g., Festman et al. 2010; Prior and Gollan 2013). Festman et al. (2010, p. 1) define cognitive control as "the ability to flexibly adapt behavior to current demands by focusing on task-relevant information and behaviors over a period of time while dealing with interference and competition." Executive functions include inhibition, shifting, and working memory, assessed by a variety of non-verbal tasks such as the Flanker task, the Dimensional Card Sort task, and the backward digit recall task.

Several studies have examined the production of cross-language errors in adult bilinguals. These studies focused on the role of language proficiency and the grammatical properties of the cross-language errors, as well as their appearance in unilingual and mixed language settings. Poulisse (2000) and Poulisse and Bongaerts (1994) reported more unintended L1 utterances in L2 speech than vice versa in story retelling tasks and spontaneous speech samples. Furthermore, lower L2 proficiency led to the use of more unintended L1 in L2, particularly function words. This finding was explained by lower activation and higher resting levels of the non-dominant L2 among beginning L2 learners as compared to more proficient bilinguals. Poulisse (2000) also suggested that less fluent speakers have less automaticity in speech production and more difficulty suppressing automatized responses. Gollan and colleagues examined intrusion errors in eye tracking experiments where bilinguals read monolingual or mixed language paragraphs aloud. Here, too, intrusion errors involved function words more than content words. These studies further reported that more errors resulted for dominant language target words than for non-dominant target words (Gollan et al. 2014; Schotter et al. 2019). Another study involving reading grammatical and ungrammatical codeswitching (CS) found that more intrusion errors occurred in passages with ungrammatical CS and in passages with a high number of switches (Gollan and Goldrick 2016). Gollan et al. (2014) reported more cross-language errors in mixed than unilingual passages. Likewise, Declerck et al. (2017) noted more intrusion errors when bilinguals were asked to produce a sentence during a language switch trial than during trials that took place in the same language as the previous trial. The authors interpreted this finding in terms of increased between-language interference in bilingual as compared to unilingual language settings.

Bilingual language activation and control have been assessed in natural speech via elicited narratives and, as noted above, with experimental paradigms involving CS (e.g., Kootstra et al. 2010; Hofweber et al. 2016). Kootstra et al. (2010), for instance, employed a picture description task with two dialogue partners, in which the confederate's CS patterns were manipulated. The results showed that in terms of syntax and tendency to codeswitch, the participants tended to adapt their CS patterns to those of their interlocutor. Green and Wei (2014) proposed the Control Process Model of Code-switching to account for different kinds of language control depending on the type of CS. For instance, dense CS, where there is frequent back-and-forth language switching within an utterance, entails a high level of language co-activation. This is described as an 'open control' mode with relatively little need for language inhibition of the non-target language (Green and Wei 2014). In contrast, insertional CS (switching of individual lexical items of one language into a grammatical frame of another) involves 'coupled control,' in which language control is ceded from one language to the other with the dominant language 'gate' open and the other 'on the latch' (Green and Wei 2014, p. 502). Hofweber et al. (2016) introduce the notion of monitoring, "a cognitive control mechanism involving the management of co-activated conflicting task-schemata allowing for flexible and rapid adaptation to changes in behavioural goals or task requirements" (p. 649). They argue that to prevent cross-language interference, monolingual contexts involve the least amount of language monitoring, while insertional and dense CS involve the most. The present paper is focused on language control during insertional CS.

There are various functions and motives for CS. Psycholinguistic motivations include a lack of word knowledge, difficulty retrieving a word, or the sense that meaning may be better conveyed in the alternate language (Genesee and Nicoladis 2009; Greene et al. 2012; Heredia and Altarriba 2001). CS has also been linked to sociolinguistic factors, including changes in social setting or topic of conversation, the role or identity of the interlocutors, and the different contexts for using the first language (L1) versus that of the L2 (Gumperz 1982; McClure 1981; Myers-Scotton 1993b; Raichlin et al. 2019). Pragmatic functions of CS include attention-getting, emphasis or focus, clarification, or disagreement (Grosjean 1982; Wei 1998; Zentella 1997).

Some of these factors might account for cross-language intrusions. For instance, a child may not be aware of the conversational partner's identity and sociolinguistic norms. Furthermore, she may have trouble with grammatical or lexical features in one of her languages due to imbalanced language skills. Sequential bilingual children are expected to codeswitch more from L1 (the minority/home language) to L2 (the majority/societal language) than vice versa since they often experience language shift where the L2 becomes the stronger language over time (Basnight-Brown and Altarriba 2007; Ebert and Kohnert 2016). This CS directionality effect has been corroborated in several studies reporting CS from a weaker into a stronger language (Greene et al. 2012; Halmari 2005; Kuzyk et al. 2020; Peynircioğlu and Durgunoğlu 2002; Ribot and Hoff 2014). Similarly, cross-language intrusions may be more common from L1 to L2.

CS directionality can also be a function of sociolinguistic factors, specifically switching from the home to the societal language. Gutiérrez-Clellen et al. (2009) found that English-dominant 5 to 6-year-olds tended to switch from Spanish (home language) to English (societal language), whereas their Spanish-dominant peers did not switch from English to Spanish. The authors attributed the directionality patterns to the children's awareness of the societal language. Kapantzoglou et al. (2021) reported that, regardless of proficiency, a majority of the 7-year-old Spanish-English children in their study codeswitched from Spanish (home language) to English (societal language), but only a very small percentage (10%) did so in the opposite direction. The authors concluded that "sociocultural" factors might outweigh "psycholinguistic" factors. Montanari et al. (2019) reported no directionality effects for children entering preschool but a clear preference for switching from the home language to the societal language a year later. These authors also ascribed this shift to the children's understanding of sociolinguistic factors such as language status and use in the children's community. Smolak et al. (2020), in a study of bilinguals aged 31–39 months, found that exposure was a predictor of CS directionality for Spanish–English in the southwest US, whereas proficiency predicted CS directionality for French–English in Montreal. Similar patterns have been reported for other language pairs. Iluz-Cohen and Walters

(2012) found more CS from English (home language) into Hebrew (societal language), which the authors attributed to the children's "sociolinguistic sensitivity" (p. 71). Likewise, Raichlin et al. (2019) reported more CS from Russian (home language) into Hebrew (societal language), which was attributed not just to the children's language proficiency but also to the children's greater exposure to the societal language, the language with the greatest use in a wide range of social settings.

1.2. Language Control in Bilingual Children

Ineffective language control in bilingual children has been examined through picture-naming/description tasks in single- and mixed-language trials (Gross and Kaushanskaya 2018, 2020, 2022; Jia et al. 2006; Kohnert 2002; Kohnert et al. 1999). A breakdown in language control in these studies was reported when children labeled or described the picture stimuli in the non-target language. These studies showed more cross-language naming errors when children were required to name pictures under more demanding, bilingual conditions (e.g., a picture naming task in which the language cue alternated between two languages) than under less demanding monolingual conditions (Gross and Kaushanskaya 2018; Jia et al. 2006; Kohnert 2002; Kohnert et al. 1999).

Gross and Kaushanskaya (2018) examined the relationship between domain-general cognitive control and domain-specific language control in children. In their 2018 study, poorer performance on a shifting task was associated with a greater number of cross-language errors. Two additional studies worthy of note for their methodological innovation further examined language switching and executive functions in bilingual children (Gross and Kaushanskaya 2020, 2022). In these studies, Spanish–English bilingual children and confederates took place in a dialogue describing pictures in turns. The confederates were introduced as either English or Spanish-speaking monolinguals. Three different contexts were assessed for cross-speaker errors (involving the use of non-target language English with a Spanish-speaking confederate, and vice versa): a unilingual Spanish session, a unilingual English session, and a dual language, Spanish–English session. The 2020 study reported no overall relationship between cognitive control and language control but a greater role for cognitive control in the dual language context than in the single language sessions. Based on Green and Abutalebi (2013), the authors attributed this greater role to higher control demands (such as inhibition and task disengagement) for dual-language settings. The 2022 study compared cross-language, inter-sentential CS (using English with a Spanish-speaking confederate and vice versa) with intra-sentential CS (inserting English words in the target language Spanish, and vice versa). The authors reported an association between executive skills and language control skills for inter-sentential CS but not for intra-sentential CS. This finding is in line with a study on French–English bilingual children's CS, which reported no association between executive skills and intra-sentential CS (Kuzyk et al. 2020).

Gross and Kaushanskaya (2020) also addressed proficiency and reported more cross-language intrusions for children with weaker overall language skills in both single and dual language sessions. More intrusions occurred when the confederate spoke the child's weaker language. The authors attributed the relationship between intrusions and language skills to less developed communicative and sociolinguistic competence and more linguistic gaps among children with lower language abilities.

Given the findings for the relative independence of cognitive and language control (particularly in intra-sentential CS), the present study focused on the unique features of language control in bilingual children by asking preschool children to repeat sentences involving intra-sentential CS.

1.3. Bilingual Children with Developmental Language Disorder (DLD)

Developmental language disorder (DLD) refers to a language disorder that is characterized by comprehension and production impairment in connection with grammar, vocabulary, and discourse. The language abilities of individuals with DLD are defined as below those of their peers, resulting in communicative, social, and academic repercussions. DLD is also defined by exclusion of motor or auditory impairments, emotional or neurological problems, and below-average non-verbal intelligence (American Psychiatric Association 2013). DLD has been estimated to occur in 7.4% of preschool children in the US (Tomblin et al. 1997).

Children with DLD are a heterogeneous group whose language profile may show difficulties in morphology, syntax, phonology, and pragmatics (Leonard 2014), with some children affected in just one language module, others in several (Friedmann and Novogrodsky 2008). Children with DLD may also have lexical–semantic problems, manifested in a delay in vocabulary acquisition and difficulty with lexical naming and novel word learning (Lahey and Edwards 1999; Rice et al. 1994).

Bilingual children with DLD present with language difficulties in both of their languages, albeit not necessarily the same difficulties (e.g., Armon-Lotem and Walters 2010). Bilingual DLD has been defined as below-norm performance in both languages on standardized or locally-normed assessments (Armon-Lotem and Meir 2016; Ebert and Kohnert 2016; Håkansson et al. 2003; Meir et al. 2015). Some of the language phenomena reported for bilingual children with DLD are omission/substitution of prepositions (Armon-Lotem et al. 2008), errors in complex structures involving non-linear relations (Meir et al. 2015), and gaps in lexical–semantic skills (McMillen et al. 2020). See Novogrodsky and Meir (2020) for a comprehensive review of bilingual DLD covering lexical, phonological, morphosyntactic, syntactic, and narrative abilities.

Some of the language difficulties apparent in bilingual children with a language disorder overlap with those observed in bilingual children with typical language development (TLD) (Degani et al. 2019; Meir and Armon-Lotem 2015; Paradis 2010). Such similarities could result in misdiagnosis of both populations: Children with DLD may be underdiagnosed, while those with typical development may be incorrectly labeled as language impaired (Bedore and Peña 2008; Rothweiler 2007). Identification of DLD among sequential bilingual children remains a challenge (Ebert and Kohnert 2016). CS abilities are under-investigated as a potential marker of bilingual DLD, and this avenue is clearly worthy of pursuit. The study we describe below on cross-linguistic intrusions is an attempt to move in this direction.

Until recently, studies on CS in children with DLD have been scarce and inconclusive. By far, the most innovative and sophisticated research on CS in bilingual children with DLD is Gross and Kaushanskaya's 2022 study described above. This study found that the children with DLD had more cross-language errors than the children with TLD when speaking in Spanish (home language) to a Spanish speaking confederate. No differences between children with TLD and DLD were reported for intra-sentential CS, with both groups inserting more English (societal language) words into Spanish than vice versa. Performance on executive skills, too, was found to be similar in the two groups. On this basis, the authors suggest that the higher rate of cross-language errors in this population may be attributed to lower pragmatic and metalinguistic skills and a greater likelihood of home language attrition.

Several other studies have examined the impact of language impairment on CS frequency and directionality. Iluz-Cohen and Walters (2012), using a narrative elicitation and story retelling task involving interlocutors introduced as unilingual or bilingual, examined English–Hebrew bilingual 5 to 7-year-olds and reported a higher rate of CS for children with DLD than children with TLD. Furthermore, while TLD children switched more from their L1 (weaker language) to L2 (dominant), children with DLD tended to switch in both directions. Similarly, O'Toole and Hickey (2012), based on interviews with therapists, noted that language-impaired children codeswitched more frequently and did so from the majority language (English) to the minority language (Irish). Greene et al. (2012), using semantic-expressive tasks with 5-year-old Spanish–English bilinguals, found that language-impaired children mixed more in English (into Spanish) than their TLD peers, while the TLD children mixed more in Spanish. Similarly, Sheng et al. (2012), using a word association task with 7 to 9-year-olds, noted that the children with TLD switched more from Spanish to English, while the language-impaired children switched more from English to Spanish. Mammolito (2015) analyzed narratives from 5 to11-year-old Spanish–English bilinguals and noted more CS for children with more severe language impairment. In contrast to these studies, Gutiérrez-Clellen et al. (2009), studying 6-year-old Spanish–English bilingual children's narratives and conversations, reported that the language-impaired children did not differ from their TLD peers in terms of frequency of CS. Finally, in a story retelling task with 5 to 7-year-old Spanish-English bilinguals, Kapantzoglou et al. (2021) found significantly more CS from English to Spanish than vice versa for both TLD and DLD children and no difference in CS frequency between the two groups. In sum, the data regarding CS frequency and directionality among children with DLD are inconsistent.

1.4. Sentence Repetition

In order to examine the production of non-targeted CS (i.e., cross-language intrusions), the present study employed a sentence repetition (SRep) task. In this task, the participants are instructed to listen carefully to sentence stimuli and to repeat them exactly as heard. SRep involves listening, phonological short-term memory, and production skills. When the sentences are of sufficient length, the task goes beyond simple parroting, requiring the decoding and encoding of underlying syntactic, semantic, and phonological representations from long-term memory (Marinis and Armon-Lotem 2015; Riches 2012). SRep has been used widely to examine language abilities in a variety of populations (Vinther 2002), including monolinguals (Komeili and Marshall 2013), bilinguals (Meir et al. 2015), and foreign language learners (Rast and Dommergues 2003). SRep has also been used to assess the grammaticality of CS via intra-sententially codeswitched sentences (Azuma and Meier 1997; Lipski 2017; Soesman and Walters 2021). Finally, SRep has been shown to be a discriminator of children with DLD from those with TLD (Armon-Lotem and Meir 2016; Ebert 2014; Pratt et al. 2021). The lower performance of children with DLD on SRep tasks has been attributed to impairments in language representation and/or processing (Marinis and Armon-Lotem 2015).

1.5. The Present Study

The current study investigates cross-language errors in an SRep task containing unilingual and intra-sententially codeswitched sentence stimuli. It operationalizes cross-language errors as those instances where a non-codeswitched item from the stimulus sentence is produced as a codeswitch. We refer to such CS as "non-targeted CS".

The present study differs from previous work on cross-language errors in several ways. First, previous research on bilingual children and language control has used picture stimuli in either single language or mixed-language trial blocks, exposing the participants to lesser or greater language co-activation and competition (e.g., Gross and Kaushanskaya 2020; Jia et al. 2006). In contrast, the present study uses a task in which there are varying degrees of language activation within a single language block with participants exposed to alternating unilingual and codeswitched stimuli. In processing and producing a codeswitched sentence, the child needs to continuously adjust the activation levels of the two languages, necessitating a relatively high activation level of both languages (Green and Abutalebi 2013; Kroll et al. 2014; Litcofsky and Van Hell 2017; Van Hell et al. 2015). Although we argue that both languages are activated throughout the task, more intense bilingual activation results when hearing and repeating a codeswitched sentence, and less intense bilingual activation when hearing and repeating a unilingual stimulus. In light of this, we hypothesize more non-targeted CS for stimulus sentences with CS, which is expected to elicit greater bilingual language competition and interference (e.g., Declerck et al. 2017; Hofweber et al. 2016).

Some studies on intrusion errors have focused on intra-sentential CS via read-aloud tasks (e.g., Gollan and Goldrick 2016) or corpus studies (Lipski 2016) with adults. Other studies have analyzed errors based on picture naming (e.g., Zheng et al. 2018) or language samples from story retelling and spontaneous speech (Poulisse 2000; Poulisse and Bongaerts

1994). However, no studies of bilingual children we are aware of have experimentally manipulated stimuli via SRep with CS at different points in a sentence.

Several features of the present study make it distinctive. First, it examines which type of CS in the stimulus sentence triggers cross-language switching. The participants hear intra-sentential codeswitches, which were incorporated at different loci within a prepositional phrase (PP). These included single noun or preposition switches or longer word stretches, such as full PPs. We hypothesize that the larger the switched constituent in the stimulus sentence, the more the language of this constituent will be activated and subsequently produced. Moreover, while previous research has looked at the linguistic status of cross-language errors in terms of function versus content words (e.g., Schotter et al. 2019), the present study also examines for which sentence constituent cross-language errors are most likely to occur.

Furthermore, the present study analyzes directionality effects by comparing non-targeted CS in English stimuli containing CS to Hebrew (the societal language) with Hebrew stimuli containing CS to English (the home language). We hypothesize that the participants will respond differently to a switch from the home language to the societal language than vice versa (Gutiérrez-Clellen et al. 2009; Iluz-Cohen and Walters 2012).

Finally, this study compares cross-linguistic errors in children with TLD and DLD. We expect more difficulty with language control among children with DLD (e.g., Gross and Kaushanskaya 2022) and hence more non-targeted CS for this group than for children with TLD.

Specifically, the study addressed the following questions:

1. **Stimulus CS condition, Language and Group**: To what extent will codeswitched sentence stimuli in five different sites elicit more non-targeted CS than non-codeswitched/unilingual stimuli? To what extent will there be a frequency difference in non-targeted CS between English and Hebrew sentence stimuli? To what extent will there be a difference in non-targeted CS between children with TLD and DLD?
2. **Sentence constituent**: In which sentence constituent will most non-targeted CS occur? To what extent will there be a frequency difference in non-targeted CS between English and Hebrew sentence stimuli? To what extent will there be a difference in non-targeted CS between children with TLD and DLD?

2. Methods

The present study is part of a larger project investigating linguistic markers of bilingual DLD based on the first author's PhD dissertation (Soesman 2018). Portions of the participants, materials, and procedure sections overlap with those reported in a study on grammatical aspects of CS in bilingual preschool children (Soesman and Walters 2021). The questions addressed here, the data analyses, and the findings on cross-linguistic errors are unique to the present study.

2.1. Participants

A total of 78 English–Hebrew bilingual children (41 girls), ages 5;5–6;10 (M = 5;11), participated in the study. Background information was obtained from a parent questionnaire. All of the children were raised in English-speaking homes and were exposed to Hebrew, the societal language, for at least 17 months. Fifty-six children were sequential bilinguals who acquired Hebrew at least three months after birth; 22 were exposed to Hebrew from birth. The participants were recruited from public, Hebrew-speaking regular, and "language" preschools in Israel. All of the children were administered Raven's Colored Progressive Matrices non-verbal IQ test (Raven 1998). Only children with standard scores above 90 were included in the study.

To evaluate language proficiency, the children were administered standardized tests in both languages, the CELF-2 Preschool for English (Wiig et al. 2004) for English and the Goralnik Screening Test for Hebrew (Goralnik 1995). The CELF includes six sub-tests: sentence structure, word structure, word classes, expressive vocabulary, concepts and following directions, and sentence repetition. The Goralnik test assesses vocabulary, sentence repetition, listening comprehension, expression, pronunciation, and storytelling. The participants who scored more than 1SD below the monolingual mean in L1 English were considered to be below the normal range for English, whereas those who scored more than 1.5 SD below the monolingual mean in L2 Hebrew were considered to be below the normal range for Hebrew. These cut-off points align with previous studies with the same population in Israel (Iluz-Cohen and Armon-Lotem 2013; Iluz-Cohen and Walters 2012; Meir et al. 2015). Since sequential bilingual children are exposed to L1 from birth, and they are exposed to L2 around the age of 24 months (see Table 1), the cut-off point is not identical for the two languages. Their Hebrew skills are expected to be lower compared to monolinguals due to lower exposure. In the literature on diagnosing bilingual children with DLD, a range between −1 SD and −2 SD is used depending on the tests, the amount of exposure, and the specific population (Altman et al. 2021; Peña et al. 2020; Thordardottir 2015.)

Children were considered typically developing (TD) if they scored within the monolingual norms in at least one of their languages. The raw scores on both tests were converted to z-scores based on the standardized criteria.

Children were classified as bilingual DLD if they met all of the following criteria:

1. They had scores below the norm on both English and Hebrew assessments (Armon-Lotem and Meir 2016; Ebert and Kohnert 2016; Håkansson et al. 2003; Kohnert 2010).
2. To ensure that low scores on the Hebrew screening test were not a result of lack of exposure to L2/Hebrew but rather due to language impairment, Length of Exposure (LoE) was taken into account for classifying a child with DLD. Only children in one of the following categories were included in the DLD group:
    a. LoE from birth AND z-scores from −1.50 to −1.74;
    b. LoE greater than 36 months AND z-scores from −1.75 and −1.99;
    c. LoE from 24–36 months AND z-scores from −2.00 to −2.99;
    d. LoE from 16–23 months AND z-scores lower than −2.99.
3. Exclusionary criteria for DLD, i.e., no auditory, visual, neurological, or emotional impairments, as well as no history of otitis media (Leonard 2014; Tallal and Stark 1981) based on information provided by parents.

Table 1 displays demographic, exposure, and proficiency background data for the two groups. The table shows that in terms of demographic data (age, gender, socioeconomic status (mother's education)), there were no significant differences between the groups. The data further indicate that there were no significant differences in length of exposure to Hebrew and age of onset (AoO) of Hebrew.

Table 1. Demographic, exposure (LoE, AoO), proficiency and other background data from children with TLD and DLD.

	TLD (*N* = 65)	**DLD (*N* = 13)**	***t***	***p*-Value**	***t*-Tests**
Age in months (SD)	70.06 (3.48)	73.77 (5.79)	2.230	0.043	DLD > TLD
Mother's education in years (SD)	16.98 (2.81)	17.45 (1.86)	0.533	0.596	-
LoE in months (SD)	51.11 (16.38)	50.00 (20.29)	0.214	0.831	-
AoO in months (SD)	18.95 (16.96)	23.77 (20.20)	0.905	0.368	-
Eng. prof. (*CELF*) z-score (SD)	−0.50 (1.00)	−1.93 (0.73)	4.896	< 0.001	TLD > DLD
Heb. prof. (*Goralnik*) z-score (SD)	−0.88 (1.40)	−2.73 (1.42)	4.346	< 0.001	TLD > DLD
Non-verbal IQ	121.42 (14.06)	108.54 (8.88)	4.27	< 0.001	TLD > DLD

LoE: Length of exposure to Hebrew; AoO: Age of Hebrew onset.

The children in the TLD group had significantly higher scores than the children in the DLD group for both English (CELF) and Hebrew proficiency (Goralnik). The children in the TLD group scored significantly higher on non-verbal IQ and were younger than the children in the DLD group. We, therefore, added the IQ standard score and Age as fixed factors to the GLMM analysis of non-targeted CS.

2.2. Materials and Design

The stimuli consisted of 36 sentences in English and 36 in Hebrew. Each sentence included an NP subject, an intransitive V, a PP consisting of P+DET+N, and a temporal phrase. The following prepositions were included in the stimulus sentences: *around* (saviv), *inside* (betox), *over* (me'al), *next to* (leyad), *before* (lifney), and *after* (axrey). The temporal phrase did not contain a preposition. Each of the 36 sentences was rearranged into a set of six conditions where the following element(s) were codeswitched: a single N, a DET+N, a single P, a P+DET, a P+DET+N. Each set also included a unilingual stimulus sentence (no-switch condition). An example of a set with the six conditions appears in (1) for English and (2) for Hebrew.

(1) An example set of English stimulus sentences with Hebrew switches

	Subject	**Verb**	**Prepositional Phrase**	**Temporal Phrase**
Non-CS	The girl	played	inside the bedroom	all day
CS-N	The girl	played	inside the XEDER	all day
CS-DET+N	The girl	played	inside HA- XEDER	all day
CS-P	The girl	played	BETOX the bedroom	all day
CS-P+DET	The girl	played	BETOX HA- bedroom	all day
CS-P+DET+N	The girl	played	BETOX HA- XEDER	all day

(2) An example set of Hebrew stimulus sentences with English switches

	Subject	**Verb**	**Prepositional Phrase**	**Temporal Phrase**
Non-CS	Ha- yalda	sixqa	betox ha- xeder	kol ha- yom
CS-N	Ha- yalda	sixqa	betox ha- BEDROOM	kol ha- yom
CS-DET+N	Ha- yalda	sixqa	betox THE BEDROOM	kol ha- yom
CS-P	Ha- yalda	sixqa	INSIDE ha- xeder	kol ha- yom
CS-P+DET	Ha- yalda	sixqa	INSIDE THE xeder	kol ha- yom
CS-P+DET+N	Ha- yalda	sixqa	INSIDE THE BEDROOM	kol ha- yom

English and Hebrew sentences were constructed to be identical in terms of syntax and semantics. The lexical items in the stimulus sentences were early-acquired vocabulary taken primarily from Hart and Risley (1995). Lexical items were adapted to the Israeli context following feedback from children, parents, and teachers and based on pilot work. The stimuli did not include English-Hebrew cognates in order to preclude cross-linguistic activation effects due to similar phonology and meaning (Van Hell et al. 2015).

For each language, six experimental lists were created with six items for each CS condition in English and six parallel items for each CS condition in Hebrew. Each list included a total of 36 test items, with one sentence from each of the 36 sets. Thus, no participant was presented with the same sentence in more than one condition. Twenty-four unilingual filler sentences were interleaved among the codeswitched sentences, yielding a total of 60 stimulus sentences. The filler sentences differed syntactically from the test sentences, and they were not coded or scored in the present study. The six conditions and distractor sentences appeared in the same, pseudorandom order in each list. Sentences from the same CS condition as well as identical lexical items did not appear consecutively. Each experimental list started with a filler sentence followed by a unilingual test sentence. The participants were assigned randomly to lists in both English and Hebrew.

2.3. Procedure

The study was approved by the university Institutional Review Board and the Israel Ministry of Education Chief Scientist's office. The participants' parents signed consent forms after they were informed about the aims of the study. Each participant completed the Raven's non-verbal IQ test, the CELF English proficiency, and the Goralnik Hebrew screening test. The sentence stimuli were recorded with a fluent bilingual female speaker of English and Hebrew reading the sentences. The recording was condcuted with an AKG C1000 condenser microphone in a sound-proof studio and digitized by means of an M Audio Delta interface. The recordings were then incorporated into a PowerPoint© presentation.

The participants were tested individually in a separate room in their kindergarten where only the tester and child were present. The stimulus sentences were presented via PowerPoint© via a laptop computer with audio headphones. The children were asked to repeat the sentences verbatim. To ensure that the participants understood the task, three practice sentences were presented at the beginning of each session. Two of these sentences were unilingual, while the third contained a codeswitch. After the child repeated the third sentence, the experimenter verified that the child had noticed that the last sentence was special since it contained words from both languages. After repeating 10 sentences, the child received a sticker as a means of encouragement.

Half of the participants were tested on the English sentences in the first session and on the Hebrew stimuli a week later. The other half were tested in Hebrew first and English a week later. All the interactions between the tester and the child were in English during the English session and in Hebrew during the Hebrew session. Each session was audio-recorded on a Toshiba laptop via Audacity© software. In addition, the children's responses were recorded manually.

2.4. Coding and Analysis

The recordings were transcribed and coded off-line following each session. Cross-language errors were analyzed for all instances of non-targeted CS. Non-targeted CS was defined as any language switching by the participants that (1) was not present in the original stimulus sentence, (2) occurred in a sentence constituent outside of the PP for codeswitched sentence stimuli (i.e., in the NP subject phrase, in the V phrase or in the temporal phrase in sentence-final position), and (3) occurred in any sentence constituent within the unilingual sentence stimuli. Codeswitches in the PP were excluded from the analysis since the PP was the very site of the **targeted** CS, and the focus of the current study was on **non-targeted** CS. (For a detailed discussion on children's modifications of switches within the PP, see Soesman and Walters (2021)).

For the English stimulus sentences, a non-targeted switch meant a switch into Hebrew; for the Hebrew stimulus sentences, it meant a switch into English. Examples of non-targeted codeswitches are underlined and are given in *UPPERCASE* in (3) and (4). Codeswitches in the stimulus sentences appear in **bold**.

(3). Stimulus sentence: The builder works inside the **bayit** every morning
'The builder works inside the house every morning.'
Non-targeted CS: The builder works inside the bayit <u>*KOL*</u> morning
'The builder works inside the house every morning.'

(4). Stimulus sentence: ha- yalda sixqa **inside the bedroom** kol ha- yom
the girl play-PAST inside the bedroom whole the day
'The girl played inside the bedroom the whole day.'
Non-targeted CS: ha- yalda sixqa inside the bedroom <u>*ALL THE TIME*</u>

The frequency of non-targeted CS was computed for each of the six CS conditions (no switch, CS-N, CS-DET+N, CS-P, CS-P+DET, and CS-P+DET+N). A non-targeted codeswitch consisting of more than one word within a phrase (e.g., "all the time") was counted as a single switch. Non-targeted elements that were self-corrected were not included in the analysis.

All of the analyses were performed in R (R Core Team 2012). The main measure in all analyses was the frequency of non-targeted codeswitches. Generalized Linear Mixed Models (GLMM) analyses with Poisson distribution were performed, examining the effect of fixed factors and random intercepts. Poisson distribution allows analyzing frequencies (counts), and the mixed models approach allows accounting for individual performance. The following fixed factors were tested in all analyses: Language (English/Hebrew), Group (TLD/DLD), and their interaction. For RQ1, we also tested the effect of CS condition (non-CS, CS-N, CS-DET+N, CS-P, CS-P+DET, CS-P+DET+N). For RQ2, we tested for Sentence constituent (Noun, Verb, Temporal phrase). The following covariates were also tested: Age, AoO, and non-verbal IQ. The participants were included as random intercepts. In all analyses, the Laplace approximation was set to 0 (nAGQ = 0), since with this option on (nAGQ = 1), the model failed to converge. In order to determine the contribution of each fixed factor, likelihood ratio tests were performed, and based on the results, optimal models were created. The estimates of each fixed factor are reported in Appendix A. Significant interactions were explored using the *emmeans* package (Lenth et al. 2022) by means of pairwise contrasts. Predicted frequencies of non-targeted CS based on significant interactions were plotted using the *plot model* function of the *sjPlot* library (Lüdecke et al. 2022) for ease of interpretation.

3. Results

By way of background, a large majority of children (73%) produced at least one non-targeted switch in either English, Hebrew, or both languages. Approximately one-third of all children produced at least one non-targeted switch in English sentences, whereas a little more than half did so in Hebrew. Moreover, while 71% of the TLD children had at least one non-targeted switch, in the DLD group, this number reached 85%. Further, while the majority in the DLD group (7 out of 11) produced non-targeted switches in both English and Hebrew sentences, in the TLD group, only a small minority (approximately 10%) did.

3.1. Frequency of Non-Targeted CS as a Function of CS Condition, Language and Group

The frequencies of non-targeted CS for unilingual and codeswitched sentences are presented in Table 2.

Table 2. Mean frequency, range, and sum of non-targeted CS for six CS conditions in English and Hebrew stimuli for children with TLD and DLD.

	TLD (N = 65)			**DLD (N = 13)**		
	Mean Frequency	**Range**	**Sum (No. of Children)**	**Mean Frequency**	**Range**	**Sum (No. of Children)**
Condition						
English						
Non-CS	0.08	0–2	5 (N = 4)	0.15	0–2	2 (N = 1)
N	0.09	0–2	6 (N = 4)	0.38	0–2	5 (N = 4)
DET+N	0.11	0–1	7 (N = 7)	0.54	0–4	7 (N = 4)
P	0.15	0–2	10 (N = 8)	0.38	0–2	5 (N = 3)
P+DET	0.06	0–2	4 (N = 3)	0.15	0–1	2 (N = 2)
P+DET+N	0.14	0–2	9 (N = 7)	0.46	0–3	6 (N = 4)
Total Eng.	0.11		41	0.34		27

Table 2. *Cont.*

	TLD (*N* = 65)			DLD (*N* = 13)		
	Mean Frequency	**Range**	**Sum (No. of Children)**	**Mean Frequency**	**Range**	**Sum (No. of Children)**
Condition						
Hebrew						
Non-CS	0.03	0–1	2 (*N* = 2)	0	0–0	0
N	0.25	0–3	16 (*N* = 10)	0.54	0–3	7 (*N* = 4)
DET+N	0.35	0–6	23 (*N* = 11)	0.38	0–3	5 (*N* = 3)
P	0.29	0–3	19 (*N* = 11)	0.46	0–2	6 (*N* = 5)
P+DET	0.23	0–3	15 (*N* = 12)	0.08	0–1	1 (*N* = 1)
P+DET+N	0.58	0–5	38 (*N* = 19)	0.77	0–3	10 (*N* = 9)
Total Heb.	0.29		113	0.37		29

Table 2 suggests that more non-targeted CS resulted for the CS conditions than the unilingual condition; more non-targeted CS resulted for Hebrew stimulus sentences than for English (for the TLD children), and more non-targeted CS resulted for children with DLD than children with TLD.

In order to test the effects of CS Condition, Language, Group, and their interactions, a GLMM analysis was conducted. Age, AoA, and non-verbal IQ were also tested as potential covariates. The full list of the fixed factors and their order, as well as the outcome of the likelihood ratio tests, are presented in Table A1 in Appendix A. Based on the results of the likelihood ratio tests, the optimal model was created which included the following fixed factors that came out significant: Condition, $\chi 2 = 53.21$, $p < 0.001$, Language, $\chi 2 = 26.65$, $p < 0.001$, and Group*Language, $\chi 2 = 8.44$, $p = 0.004$. The Condition*Language interaction was not significant, $\chi 2 = 10.80$, $p = 0.06$, nor were Group*Condition*Language, $\chi 2 = 17.34$, $p = 0.30$, Age, $\chi 2 = 0.81$, $p = 0.37$, AoO, $\chi 2 = 0.37$, $p = 0.55$ and non-verbal IQ, $\chi 2 = 0.61$, $p = 0.43$. These covariates were thus excluded from further analyses. Table 3 presents the model's summary. The fixed factors explained 19% of the variance, and the random factors explained an additional 27%.

Table 3. Estimates of the model predicting the frequency of non-targeted CS.

Predictors	***Incidence Rate Ratios***	**CI**	***p***
(Intercept)	0.08	0.03–0.19	**<0.001**
Condition [CS-N]	3.78	1.81–7.88	**<0.001**
Condition [CS-P]	4.44	2.16–9.16	**<0.001**
Condition [CS-P+DET]	2.44	1.13–5.31	**0.024**
Condition [CS-DET+N]	4.67	2.27–9.59	**<0.001**
Condition [CS-P+DET+N]	7.00	3.48–14.07	**<0.001**
Language	1.07	0.64–1.81	0.789
Group	0.24	0.11–0.53	**<0.001**
Language * Group [TLD]	2.57	1.36–4.84	**0.004**
	Random Effects		
σ^2		2.02	
$\tau_{00\ ID}$		1.01	
ICC		0.33	
N_{ID}		78	
Observations		936	
Marginal R^2/Conditional R^2		0.193/0.462	

Note: For condition, non-targeted CS is the reference group; * denotes interaction.

The analysis indicates that a significant difference emerged among the conditions. A post-hoc analysis using the *emmeans* function from the *emmeans* library with Tukey corrections showed that less non-targeted CS occurred for the Non-CS condition than for CS-N ($p = 0.005$), CS-P ($p < 0.001$), CS-DET+N ($p < 0.001$), and CS-P+DET+N ($p < 0.001$). CS-N resulted in less non-targeted CS than CS-P+DET+N ($p = 0.04$), and CS-P+DET, too, resulted in less non-targeted CS than CS-P+DET+N ($p < 0.001$). The rest of the differences were not significant.

The analysis also yielded a significant Group effect and a significant Group*Language interaction. The interaction was explored using the *emmeans* function testing for contrasts and plotting the predicted values (see Figure 1).

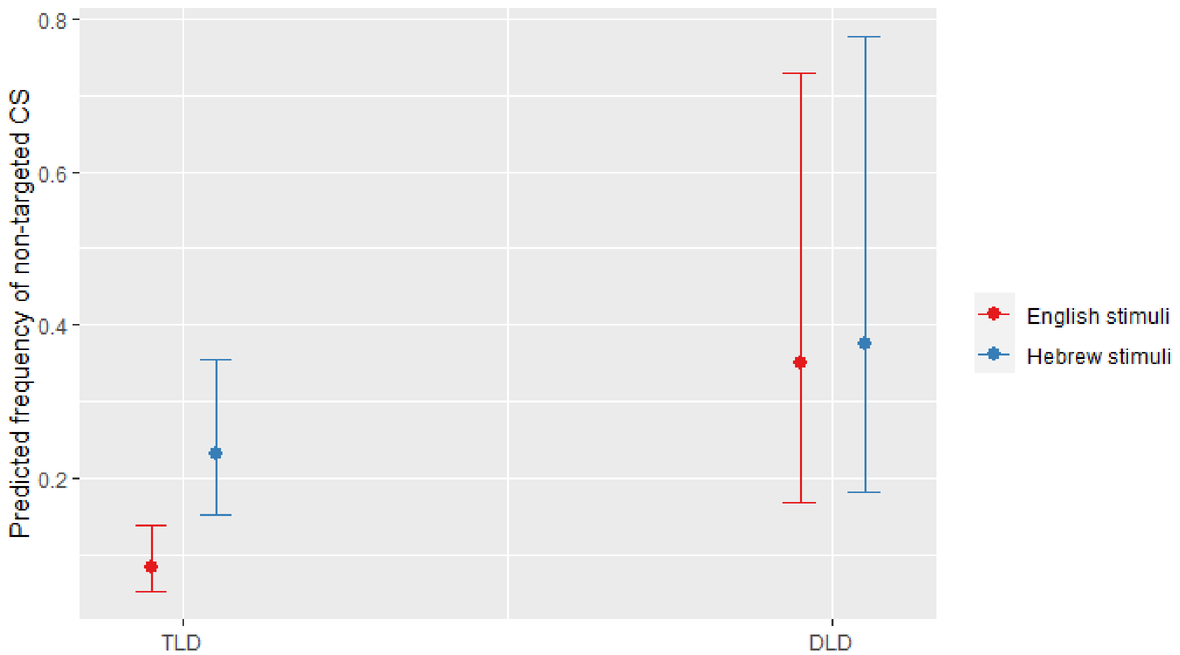

Figure 1. Predicted frequencies of non-targeted CS for the interaction of Group and Language.

Visual inspection of Figure 1 shows higher predicted frequencies of non-targeted CS for children with DLD and greater variability than for children with TLD. In addition, the graph shows no differences between English and Hebrew stimulus sentences for children with DLD but higher predicted frequencies for Hebrew stimuli than for English stimuli for children with TLD.

Post-hoc analysis with Turkey corrections confirmed that children with DLD are predicted to use a similar frequency of CS in English and Hebrew ($p = 0.79$), while children with TLD are predicted to use significantly more in Hebrew than in English ($p < 0.001$). This interaction can also be interpreted as follows: In English, children with DLD are predicted to use more non-targeted CS than children with TLD ($p < 0.001$), while in Hebrew, the two groups did not differ ($p = 0.21$).

3.2. Sentence Position of Non-Targeted CS

Table 4 demonstrates the means, ranges, and frequencies of non-targeted CS for each of the three sentence constituents: the NP Subject phrase, the Verb phrase, and the Temporal phrase. As can be seen from the table, the sentence-final temporal phrase resulted in the largest amount of non-targeted CS in both English and Hebrew sentences.

Table 4. Means, ranges, and frequencies of non-targeted CS for three grammatical constituents in English and Hebrew, for children with TLD and DLD.

	TLD (N = 65)			DLD (N = 13)		
	Mean Frequency	**Range**	**Sum (No. of Children)**	**Mean Frequency**	**Range**	**Sum (No. of Children)**
Sentence Constituent						
English						
NP Subject	0.09	0–1	6 (N = 6)	0.77	0–2	10 (N = 6)
V Phrase	0.17	0–4	11 (N = 6)	0.31	0–4	4 (N = 1)
Temporal phrase	0.32	0–5	21 (N = 11)	0.92	0–4	12 (N = 4)
Totals English			38			26
Hebrew						
NP subject	0.18	0–4	12 (N = 7)	0.54	0–4	8 (N = 5)
V phrase	0.18	0–3	12 (N = 9)	0.46	0–2	6 (N = 5)
Temporal phrase	1.35	0–13	88 (N = 25)	1.15	0–4	15 (N = 7)
Totals Hebrew			112			29

Note: The non-CS condition was excluded from this analysis since it was not relevant to the locus of non-targeted CS and there were very few instances in this condition.

A GLMM analysis was performed on the frequency of non-targeted CS, testing the effects of Sentence constituent, Language, Group, and their interactions. The full list of the fixed factors and their order, as well as the outcome of the likelihood ratio tests, are presented in Table A2 in Appendix A. The optimal model included the following fixed factors that came out significant based on the likelihood ratio tests: Sentence constituent, $\chi2 = 94.33$, $p < 0.001$, Language, $\chi2 = 29.01$, $p < 0.001$, Sentence constituent*Language, $\chi2 = 9.34$, $p = 0.009$, Group, $\chi2 = 5.86$, $p = 0.02$, Group*Sentence constituent, $\chi2 = 11.26$, $p = 0.004$, and Group*Language, $\chi2 = 6.23$, $p = 0.01$. A Group* Sentence constituent*Language interaction was not significant, $\chi2 = 3.04$, $p = 0.22$. The model's estimates are presented in Table 5. The fixed factors explained 24% of the variance, and the fixed and random factors together explained 56%.

Table 5. Estimates of the model predicting the frequency of non-targeted CS.

	Frequency		
Predictors	***Incidence Rate Ratios***	**CI**	***p***
(Intercept)	0.61	0.27–1.41	0.250
Sentence constituent [Temporal]	1.08	0.52–2.25	0.844
Sentence constituent [Verb]	0.63	0.26–1.51	0.298
Language	0.77	0.36–1.64	0.494
Group	0.10	0.04–0.28	**<0.001**
Sentence constituent [Temporal] * Language	2.09	0.94–4.68	0.072
Sentence constituent [Verb] * Language	0.86	0.32–2.30	0.762
Sentence constituent [Temporal] * Group	3.30	1.47–7.38	**0.004**
Sentence constituent [Verb] * Group	2.24	0.81–6.21	0.120
Language [Hebrew] * Group	2.37	1.21–4.66	**0.012**

Table 5. *Cont.*

	Frequency		
Predictors	***Incidence Rate Ratios***	**CI**	***p***
	Random Effects		
σ^2	1.50		
$\tau_{00\ ID}$	1.13		
ICC	0.43		
N_{ID}	78		
Observations	468		
Marginal R^2/Conditional R^2	0.238/0.564		

Note: * signals interaction.

For Sentence constituent, a post-hoc analysis using the *emmeans* function with Tukey corrections showed that non-targeted CS is predicted to be more frequent in the Temporal phrase than in the Subject position ($p < 0.001$). Non-targeted CS is also predicted to be more frequent in the Temporal phrase than in the Verb position ($p < 0.001$). The difference between the Subject and the Verb position was not significant ($p = 0.85$) for both groups and in both languages. The main effects of Language and Group were significant, as well as the interactions of Sentence constituent and Group and Language and Group. Post-hoc analyses with Tukey corrections were conducted, and the predicted frequencies for the interaction of Sentence constituent and Group are plotted in Figure 2. For the interaction of Language and Group, the results replicated those reported above for overall non-targeted CS (see Figure 1).

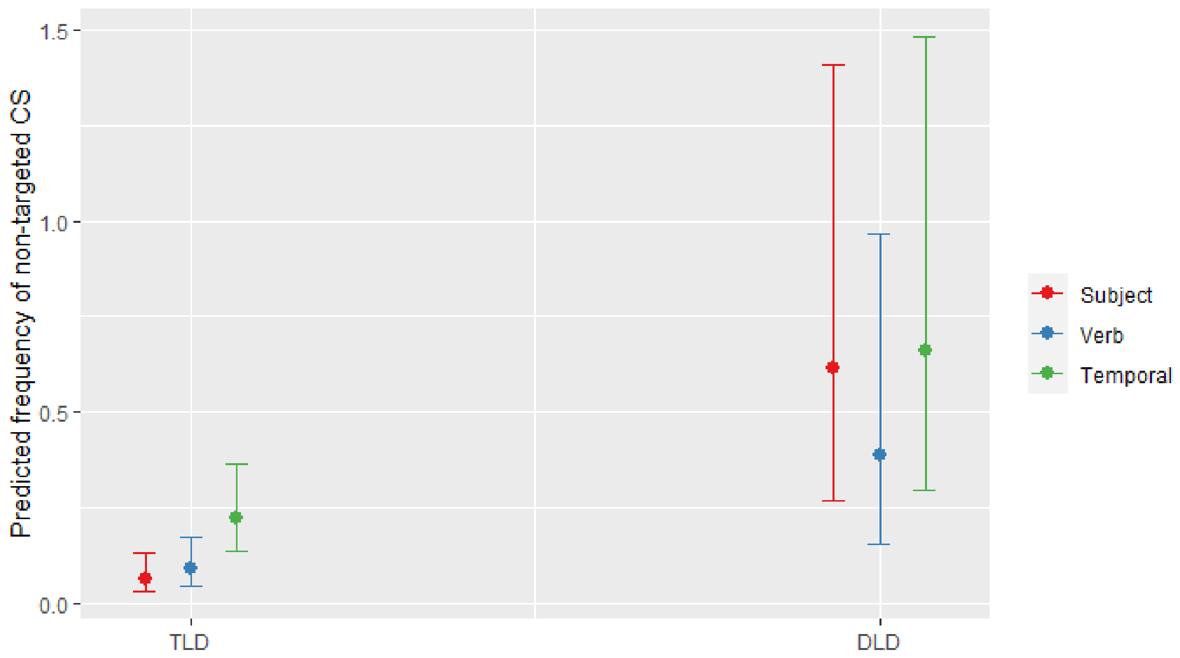

Figure 2. Predicted frequencies of non-targeted CS for the interaction of Sentence constituent and Group.

The post-hoc analysis indicated that children with DLD are predicted to use more non-targeted CS in the Temporal phrase than in the Verb ($p = 0.02$), whereas the differences between Temporal and Subject ($p = 0.33$) and between Verb and Subject ($p = 0.36$) were not significant. For TLD, the predicted values of non-targeted CS showed a significant difference between the Temporal phrase and Subject ($p < 0.001$) as well as between the Temporal phrase and Verb ($p < 0.001$), but not between the Subject and Verb ($p = 0.69$).

3.3. Non-Targeted CS within the Temporal Phrase

The category of Temporal phrase was subcategorized into grammatical constituents to assess which grammatical constituent in the temporal phrase resulted in most non-targeted CS: determiner (DET), determiner and noun (DET+N), noun (N), or other switches (for example, *tomorrow* for *etmol* 'yesterday'; or *shuv ve shuv* 'again and again' for *again and again*). These subcategories are displayed in Table 6.

Table 6. Means, ranges, and frequencies of non-targeted CS for constituents within the temporal phrase in English and Hebrew, for children with TLD and DLD.

	TLD (*N* = 65)			DLD (*N* = 13)		
	Mean	**Range**	**Sum (No. of Children)**	**Mean**	**Range**	**Sum (No. of Children)**
Temporal Phrase Constituent						
English						
DET	0.18	0–5	12 (*N* = 7)	0.38	0–4	5 (*N* = 2)
DET+N	0.09	0–1	6 (*N* = 6)	0.31	0–3	4 (*N* = 2)
N	0.03	0–1	2 (*N* = 2)	0.08	0–1	1 (*N* = 1)
Other	0.02	0–1	1 (*N* = 1)	0.15	0–1	2 (*N* = 2)
Totals English			21			12
Hebrew						
DET	0.72	0–10	47 (*N* = 18)	0.31	0–2	4 (*N* = 3)
DET+N	0.48	0–10	31 (*N* = 11)	0.09	0–3	7 (*N* = 4)
N	0.02	0–1	1 (*N* = 1)	0.03	0–1	1 (*N* = 1)
Other	0.14	0–3	9 (*N* = 6)	0.02	0–2	3 (*N* = 2)
Totals Hebrew			88			15

To further explore the non-targeted CS within the Temporal phrase, a GLMM analysis was performed with the following fixed factors: Constituent, Language, Constituent*Language, Group, Group*Constituent, Group*Language, and Group* Language*Constituent. The following factors were significant: Constituent, Language, and Group*Language. The results of the likelihood ratio tests are included in Table A3 in Appendix A. The summary of the optimal model is presented in Table 7. The fixed effects explained 30% of the variance, and the random effects explained another 29%.

For the constituents within the temporal phrase, a post-hoc analysis indicated that for both groups, predicted frequencies of non-targeted CS on DET are more frequent than on N ($p < 0.001$) and on other constituents in the Temporal phrase ($p < 0.001$). Similarly, the analysis of non-targeted CS showed that predicted frequencies of non-targeted CS on Det+N were more frequent than on N ($p < 0.001$) and on other constituents in the Temporal phrase ($p < 0.001$).

For the interaction of Group and Language, based on post-hoc analysis with Tukey corrections, children with DLD are predicted to produce a similar amount of non-targeted CS in both languages ($p = 0.56$), whereas children with TLD are predicted to produce more non-targeted CS in Hebrew than in English ($p = < 0.001$). This Group by Language interaction has the same structure as the interaction displayed above in Figure 1.

Table 7. Estimates of the model predicting the frequency of Non-targeted CS withing the Temporal phrase.

Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	0.36	0.15–0.91	**0.031**
Constituent [Det+N]	0.71	0.49–1.02	0.065
Constituent [N]	0.07	0.03–0.18	**<0.001**
Constituent [Other]	0.22	0.13–0.39	**<0.001**
Language	1.25	0.59–2.67	0.565
Group	0.26	0.09–0.74	**0.011**
Language * Group	3.35	1.37–8.21	**0.008**
	Random Effects		
σ^2		2.14	
$\tau_{00\ ID}$		1.48	
ICC		0.41%	
N $_{ID}$		78	
Observations		624	
Marginal R^2/Conditional R^2		0.302/0.587	

Note: * signals interaction.

4. Discussion

The present study investigated non-targeted CS (cross-language errors) in preschool children with and without DLD by means of an SRep task. The main findings were as follows:

1. Stimulus condition: More non-targeted CS emerged for sentence stimuli containing a codeswitch than for unilingual stimuli.
2. Directionality: More non-targeted CS occurred in Hebrew than in English stimuli for children with TLD, but not for the children with DLD, who showed no directionality difference.
3. Group Status: Children with DLD produced more non-targeted CS than children with TLD for English stimuli.
4. Sentence constituent: Most non-targeted codeswitches occurred sentence-finally, in the temporal phrase, and most involved determiners.

We discuss these four main results in terms of (1) language co-activation and competition in order to account for the difference in performance on unilingual versus codeswitched stimuli; (2) psycholinguistic and sociopragmatic factors related to home and societal language to account for the directionality effect among children with TLD and the lack thereof among children with DLD; (3) language competence and processing in children with DLD; and (4) the locus and grammatical properties of the non-targeted switches.

4.1. Codeswitched vs. Unilingual Stimuli: Frequency of Non-Targeted CS as a Function of the Type of Stimulus Sentence

Intra-sententially codeswitched stimuli resulted in more cross-language intrusions than unilingual stimuli. This result is in line with research on bilingual language co-activation. Kootstra and colleagues propose the notion of primed CS and demonstrate that a speaker's inclination to codeswitch is primed by the frequency of CS used by the conversation partner (Kootstra et al. 2010; Fricke and Kootstra 2016). Other research on language intrusions has shown that intrusions are more common when bilinguals are functioning in a bilingual setting in which they codeswitch frequently (Gollan and Goldrick 2016; Gollan et al. 2014, 2017; Grosjean 2001) or when performing language tasks in bilingual versus monolingual language trials (Gross and Kaushanskaya 2018, 2020; Jia et al. 2006; Kohnert 2002; Kohnert et al. 1999; Prior and Gollan 2013).

Processing intra-sententially codeswitched sentences (as compared to unilingual sentences) challenges language control processes due to enhanced language co-activation and competition (Bobb et al. 2016; Declerck and Philipp 2015; Green and Abutalebi 2013; Green and Wei 2014). The lower number of cross-language intrusions in unilingual sentences aligns with Green and Wei's (2014) proposal that in the monolingual language mode, the non-target language is 'locked,' whereas during intra-sentential CS, a 'coupled' control mode allows more opportunity for CS and thus for cross-language intrusions since the gate to the non-target language is not 'locked.' In line with Gollan et al.'s (2017) notion that continuous inhibition is cognitively less taxing than repetitively regulating inhibition for successive lexical items, the participants in the present study were required to exercise more control when repeating codeswitched stimuli than unilingual stimuli. Similarly, Hofweber et al. (2016) assume more language monitoring and/or competition during intra-sentential CS, that is, under conditions where the two languages are highly activated, as compared to monolingual modes. In contrast, in order to avoid unintended interference (viz., cross-language intrusions) from the second language, a monolingual mode requires "prolonged global inhibition" of the second language and minimal language monitoring.

The setting of the present study provided a high level of language co-activation. The participants were exposed to both languages throughout the experiment since the codeswitched stimuli were interleaved with the unilingual stimuli in the same session. Thus, the co-activation of two languages was probable even when processing the unilingual stimulus sentences. Indeed, the fact that there were some cases of non-targeted CS when repeating unilingual stimuli attests to this. Nonetheless, the higher incidence of cross-language intrusions for codeswitched stimuli argues for more language co-activation, more language interference, and more difficulty controlling the two languages when exposed to intra-sentential codeswitched stimuli as compared to unilingual stimuli.

More intrusions resulted when the PP in the stimulus sentence was codeswitched in its entirety than when only part of the PP contained a codeswitch, particularly for Hebrew sentence stimuli. This suggests that when participants process a comparatively longer codeswitch (P+DET+N vs. N, DET+N, or P), the language of that switch is more strongly primed and activated. This additional activation results in a greater challenge to language control and more non-targeted CS. Beyond the length of the switch, it is also possible that activation of the codeswitched language is stronger when the participants encountered a switch in a full syntactic constituent rather than part of one, where some elements remain in the non-switched language. Thus, a full constituent may lead to a boost in activation. This idea aligns with Fricke and Kootstra's (2016) suggestion that the more unambiguously words belong to a particular language, the greater the likelihood that that language will be activated.

Recent proposals maintain that distinct CS typologies affect language control processes differently (Green and Abutalebi 2013; Green and Wei 2014; Hofweber et al. 2016). For instance, according to the Control Process Model of Code-switching (Green and Wei 2014), dense CS, which is characterized by multiple switches of both grammar and lexicon within an utterance, is predicted to involve minimal language control, while insertion of single lexical items or phrases engages language control more extensively. The present study also raises the possibility that even for insertional CS, the extent of language control may fluctuate as a function of which constituent (N, DET+N, P, P+DET, P+DET+N) is codeswitched.

4.2. Directionality: Frequency of Non-Targeted CS as a Function of the Language of the Stimulus Sentence

Language control was also affected by the directionality of the codeswitch. Among children with TLD, the findings showed more non-targeted CS for Hebrew than for English stimulus sentences. More specifically, this asymmetry resulted from more frequent switching into English for Hebrew stimulus sentences than switching into Hebrew for English sentences. Unilingual stimulus sentences showed no directionality effect.

Directionality effects result from both psycholinguistic factors (lexical access difficulties in the home language) and sociopragmatic factors (extensive language exposure to the societal language and awareness of the setting and the prominence of that language).

Both sets of these factors should favor switching to the societal language. Thus, for the sequential bilingual children in the present study, similar to those in previous studies with English–Hebrew and Russian–Hebrew bilingual preschool children in Israel (Iluz-Cohen and Walters 2012; Raichlin et al. 2019), the preferred direction of switching was expected to be into Hebrew. These children grow up in a context where Hebrew is the language of daily communication in the preschool, neighborhood, and society at large, media included. Hebrew is also the language brought into the home by older siblings. English, on the other hand, is found primarily in the home. The directionality effect discussed here is in line with data reported in other language contexts, which noted more CS into the societal language than into the home language (Gutiérrez-Clellen et al. 2009; Kapantzoglou et al. 2021; Kuzyk et al. 2020; Montanari et al. 2019; Vu et al. 2010).

In contrast, the present finding of more non-targeted switching from Hebrew (the societal language) to English (the home language) for children with TLD runs counter to the preferred direction. We propose that children with TLD are not able to apply typical CS directionality patterns due to the experimental demands of sentence repetition. Typical CS directionality (into the societal language) probably involves a frequency metric in natural conditions (outside the laboratory), and it may not be accessible or relevant when the child is asked to repeat sentences with CS verbatim and in different locations. Another reason for this anomalous finding may be due to relatively stronger activation of English when processing an English codeswitch within a Hebrew stimulus sentence, i.e., when processing a codeswitch in an "atypical" direction (from the societal to the home language). It is this heightened activation of English, in competition with the psycholinguistically and sociopragmatically preferred language, Hebrew, which may have led to more cross-language intrusions into English. Thus, children with TLD may have attempted to perform the task strictly on the basis of the heightened activation of English.

The children in the DLD group did not show this directionality effect: they produced similar frequencies of non-targeted CS for English and Hebrew stimulus sentences. This is discussed in detail in the following section.

4.3. Non-Targeted CS as a Function of Group Status (TLD-DLD)

Children with DLD differed from those with TLD in two principal ways: higher rates of non-targeted CS for English stimuli and the lack of a directionality effect. They also showed a distinct pattern in terms of the sentence position of their non-targeted CS (discussed in the following section).

CS frequency. It should be noted at the outset that while the difference in predicted frequency of non-targeted CS between TLD and DLD was significant only for English stimuli, examination of individual data showed that a larger percentage of children with DLD produced non-targeted CS for both English and Hebrew stimuli.

Non-targeted CS may be a consequence of impaired linguistic competence for children with DLD. Language control is challenged due to limitations in overall language skills, which may impact sensitivity to using the required language in a particular context (Gross and Kaushanskaya 2020, 2022). Furthermore, as noted by Poulisse (2000), more cross-language errors emerge when language use is less automatized and requires more attention. These factors might be relevant for children with DLD, too. The language control difficulties of children with DLD may also stem from impaired lexical-semantic knowledge in either or both languages, as suggested by Gross and Kaushanskaya (2020, 2022). Thus, it is possible that the children with DLD in the present study activated and selected lexical items in the "wrong" language due to lexical gaps, difficulty in lexical access, and other underlying lexical-semantic difficulties.

In addition, the higher frequency of non-targeted CS among children with DLD may be a reflection of the greater prevalence of CS in their day-to-day communication. Children with DLD may be more used to codeswitching than children with TLD for reasons of lexical access difficulty and/or socio-pragmatics (Greene et al. 2012; Heredia and Altarriba 2001; Iluz-Cohen and Walters 2012; Walters 2005). Therefore, activating a lexical item in the non-target language may be an option that is more natural and acceptable to children with DLD as compared to children with TLD.

The finding that children with DLD had higher rates of cross-language intrusions is consistent with two recent studies examining language control in bilingual children and language control (Gross and Kaushanskaya 2020, 2022). In these studies, Spanish-English preschool bilinguals were asked to describe picture scenes to either a monolingual speaker of Spanish or a monolingual speaker of English. In the 2020 study, cross-language intrusions were defined as utterances in the language not understood by these speakers. The study reported an inverse correlation between language proficiency and language intrusions, implying that stronger language abilities were related to fewer intrusions. Likewise, the 2022 study found that children with DLD had more "cross-speaker switches" (cross-language intrusions in the 2020 study) when they interacted with a Spanish speaker.

Notwithstanding the differences in data collection procedures, findings regarding intra-sentential CS are very relevant here. In the 2022 study, Gross and Kaushanskaya reported no significant differences between the typically developing and language-impaired children for either frequency or directionality of intra-sentential CS. These results run counter to those of the study reported here. Gross and Kaushanskaya's (2022) results for intra-sentential CS are consistent, however, with both Gutiérrez-Clellen et al. (2009) and Kapantzoglou et al. (2021), who also found no differences between the quantity of CS for children with TLD and DLD. The inconsistency between these studies and our own probably should be attributed to the task used in the present study, where participants operated in a bilingual language mode in order to produce intra-sententially codeswitched sentences in particular locations. In the other studies, in contrast, children were engaged in more flexible, retelling, and conversational tasks either in monolingual settings (Gutiérrez-Clellen et al. 2009; Kapantzoglou et al. 2021) or in a setting in which participants were explicitly told that the interlocutors knew only one language (Gross and Kaushanskaya 2022).

It could also be argued that these latter tasks were less demanding. The SRep task, unlike conversation in monolingual settings, is an attempt to push the child's language control processes to their limits in a laboratory setting. It challenges the child's language control processes in a number of ways. The child is asked to reproduce the stimulus sentence exactly as s/he heard it, but the CS varies in where it appears in the PP in the stimulus sentence. The amount of CS in the task is far greater than the child would encounter in a natural setting or even in the experimental settings of these other studies. This has implications for language control. First and foremost, there is a high level of activation of both languages (in order to be ready to repeat the content and codeswitch in the appropriate places). Requirements to codeswitch in particular locations—along with working memory and processing limitations known to exist among children with DLD (Boerma and Blom 2020; Henry et al. 2012; Im-Bolter et al. 2006; Marton et al. 2007; Montgomery 2002) are taxing for language and cognitive control processes (cf. Thomas and Allport 2000; Van Hell et al. 2015; Von Studnitz and Green 2002, among others, for discussions on the processing costs of CS). The finding of more non-targeted CS for children with DLD shows they were less able to cope successfully with recalling the sentence content, formulating a verbatim response, and producing the codeswitched elements in their appropriate location. This multi-level semantic, morphosyntactic, and bilingual juggling act led to more cross-language intrusions in children with DLD.

The challenges to language control that CS via SRep poses for children with DLD have also been attested to in a previous study (Soesman 2018). That study found more non-targeted CS among this population for an SRep task of codeswitched single nouns. In addition, that study showed that children with DLD had a lower tendency to codeswitch elements within a PP and greater failure to codeswitch single nouns and single verbs when prompted to do so via the stimulus sentences. Like Festman et al. (2010), we consider both cross-language errors and failure to switch when prompted to do so as instances of language interference errors, indicating difficulty with language control.

Directionality effect. The other finding for children with DLD is that, in contrast to children with TLD, children with DLD did not show the directionality effect of more non-targeted CS in Hebrew stimuli as compared to English. The lack of a directionality effect may point to a different kind of difficulty in language control, one related to sociopragmatic sensitivity. For this population, the English switches in Hebrew sentences may have appeared just as commonplace as the Hebrew ones in English sentences. Therefore, there was no "red light" that turned on for the DLD children when processing a Hebrew to English CS, that is, a codeswitch in an "atypical" direction, from the societal to the home language. Children with DLD have been reported to have less developed pragmatic and social skills (Hage et al. 2021; Marton et al. 2005; Norbury et al. 2014) and to be less sensitive to language use in different social settings (Kapantzoglou et al. 2021). However, since we do not have data about the sociopragmatic skills of the children in the present study, this explanation is only a conjecture based on the population from which they were recruited.

The lack of a directionality effect is in line with other research that has shown that bilingual children with DLD do not show the pattern of bilinguals with TLD, who tend to codeswitch more from the home to the majority language than vice versa (Iluz-Cohen and Walters 2012; Sheng et al. 2012). In addition to socio-pragmatic sensitivity, this may be a function of the nature of bilingual DLD, viz., language impairment in both languages (Armon-Lotem and Meir 2016; Ebert and Kohnert 2016; Håkansson et al. 2003). Thus, since bilingual children with DLD have two weak languages, they may not have a clear preference for switching directions. Another, more speculative interpretation for the lack of an effect for directionality among children with DLD comes from the field of statistical learning. In this framework, children are said to monitor input for statistical probabilities, which helps them determine what language constructions are appropriate (Yim and Rudoy 2013). Children with bilingual DLD may have successfully monitored for the amount of CS in the experiment but may not have been statistically sensitive to the direction of the switching.

The lack of directionality differences reported here for children with DLD stands in contrast to several other studies involving this population, where more CS in the societal language than in the home language was found (Gross and Kaushanskaya 2022; Gutiérrez-Clellen et al. 2009; Kapantzoglou et al. 2021). This lack of consensus, too, may be a function of task differences. These latter studies used storytelling, conversation, and picture description tasks, which involved more interactive language use. The SRep task in the present study, along with codeswitched stimuli at different sites within a PP, did not invoke the same sociopragmatic schemata that a more interactive task would have. Thus, typical schemata for directionality may not have played a role for the language-impaired children, who were coping with the language control demands of the task.

4.4. Sentence Position of Non-Targeted CS

Another objective of the present study was to determine where in the sentence the children produced non-targeted codeswitches. The results showed that most cross-language intrusions were produced in sentence-final position, in the temporal phrase, particularly in Hebrew. One explanation for CS in this constituent is that the temporal phrase came directly after the PP, the element in the sentence stimuli that contained the target codeswitch. After processing and producing the codeswitch, the constituent immediately following this codeswitch may have been 'triggered' due to cross-linguistic priming (cf. Kootstra Gerrit Jan and van Hell 2020). Thus, in these instances, children continued using the (codeswitched) language to which they were already exposed in the PP. This finding is similar to intrusions reported in a reading-aloud task, which occurred immediately after their codeswitched antecedent (Kolers 1966). Moreover, the temporal phrase was the last element in the sentence, ruling out other candidates for intrusion.

The constituents preceding the PP–the subject and the verb–resulted in little non-targeted CS. Regarding the subject, the well-known "serial position effect" (Ebbinghaus 1913) shows better recall for items at the beginning of a list than those in the middle (e.g., Modigliani and Hedges 1987; Tan and Ward 2000). Primacy effects have been observed in children with TLD and DLD both in and outside of sentence contexts (Coady et al. 2010). For SRep tasks, fewer omissions and substitutions and better performance have been reported for sentence-initial elements than those occurring later in the sentence (e.g., Alloway and Gathercole 2005; Butterworth et al. 1986; Isham and Lane 1993). Thus, fewer cross-language intrusions of sentence-initial elements in the present study may be due to a primacy effect where sentence subjects were more accessible in the language in which they were presented.

Relatively few intrusions also occurred for the verb in the stimulus sentence. One reason may be the difficulty of codeswitching finite verbs. This difficulty has been attributed to the assumption that verbs are pivotal sentence constituents in terms of their role in case assignment and establishing the type and number of arguments (Klavans 1985; Myers-Scotton 1993a; Treffers-Daller 1993). Matras (2009) has suggested that because of this central role, codeswitching an inflected verb would lead to confusion in establishing the base language and is therefore avoided.

Within the temporal phrase, most cross-language errors involved the determiner. Other studies have reported that function words are more prone to language intrusions than content words such as nouns and verbs (Poulisse and Bongaerts 1994; Gollan and Goldrick 2016; Gollan et al. 2014; Schotter et al. 2019). Poulisse and Bongaerts (1994) propose several explanations for this content-function difference. First, since function words convey less semantic information than content words, they are less salient and may draw less attention. Non-fluent language learners, in particular, have fewer resources to attend to less salient elements and monitor their formulation. Less monitoring of function words may also be related to their shorter length, making them less prone to checking mechanisms and hence more prone to production in the 'wrong' language (cf. Gollan et al. 2014). Finally, Poulisse and Bongaerts (1994) suggest that since function words are more frequent, less activation is needed to select and produce them. This may result in more frequent cross-language errors. Schotter et al. (2019) proposed a complementary explanation, suggesting that function words are more predictable and, consequently, easier to retrieve.

Both groups had more intrusions in the temporal phrase than in the other constituents. However, children with DLD showed a pattern which differed from children with TLD regarding the sentence position of cross-language intrusions. The children with TLD made distinctions between the temporal vis-à-vis both subject and verb whereas the children with DLD only distinguished between temporal and verb in their non-targeted CS. These group differences may be due to less developed linguistic skills, specifically less sensitivity to grammatical constituents. Moreover, their efforts to cope with the demands of a codeswitched SRep task may have led them to be less attentive to grammatical properties.

4.5. Limitations and Suggestions for Future Research

The present study examined language control processes in intra-sentential CS as reflected in cross-language intrusions by means of an SRep task where only the PP was switched in the stimulus sentences. The study of language control in CS would benefit from a comparison of intra-sentential, inter-sentential, and even cross-speaker CS. Building on Green and Wei's (2014) grounding of dense CS, insertion, and alternation in different sociolinguistic contexts, a study of CS in bilingual children would help clarify the discrepancy in findings reported here for directionality in switching. This kind of study is also likely to contribute to the assessment and diagnosis of bilingual children with DLD.

Beyond the limitations in statistical power resulting from the small number of children with DLD (a problem endemic to most studies of bilingual DLD), larger numbers of children would have allowed us to look more closely at differences in language proficiency and language dominance and their role in language control. Moreover, it would allow further exploration of directionality effects in language control as a function of language competence. For instance, balanced bilinguals and Hebrew-dominant bilinguals might show different patterns in cross-language errors. Furthermore, we know from other work (Novogrodsky and Meir 2020) that L1 dominant children do not have the same morphosyntactic and narrative abilities as L2 dominant and balanced bilinguals. These differential language competencies may, in turn, affect language control abilities among children with DLD.

Another promising research avenue would be to compare the same children's CS on an SRep task with more interactive tasks such as conversation, narrative, and role-playing. Such comparisons should show a range of CS frequency and intrusion errors, providing valuable evidence about the nature of intra-sentential CS for different tasks.

In addition, it would be edifying to compare language pairs that are typologically more similar than English and Hebrew and CS for grammatical structures other than PPs. Finally, as noted at the outset, the current study set out to examine language control, not domain-general cognitive control. A comparison of the abilities for shifting and inhibition tasks could shed light on the link between impaired executive functioning and non-targeted CS in children with TLD and DLD.

5. Conclusions

Codeswitching is the signature feature of bilingualism. Since it involves the integration of multiple levels of a bilingual speaker's two languages, it is perhaps the phenomenon that can contribute most to our understanding of language control in bilingualism. In addition to structural integration of phonetic, morphosyntactic, syntactic, semantic, lexical, and pragmatic information, CS involves cross-language activation and competition as well as priming and triggering (cousins of activation) and interference and inhibition. The present paper experimentally manipulated CS in an SRep task, investigating language control as reflected in cross-language intrusions. More errors occurred for codeswitched than for unilingual sentence stimuli, highlighting increased language competition when processing intra-sentential CS. We also found that different control processes may be at work during intra-sentential CS, with greater challenges for language control when processing diverse codeswitched constituents. For children with TLD, most errors occurred in Hebrew sentences containing switches to English. For children with DLD, in contrast, relatively equal numbers appeared in both languages. This result underscores the importance of considering directionality effects and testing bilingual children in both of their languages. Children with DLD produced more cross-language errors in English stimuli than children with TLD. Occasionally, errors produced by children with DLD qualitatively differed from those made by children with TLD. These dissimilarities, along with directionality differences, point to non-targeted CS as a potential diagnostic marker for bilingual DLD.

Author Contributions: Conceptualization, A.S. and J.W.; Methodology, A.S., J.W. and S.F. Data Collection, A.S.; Formal Analysis, A.S.; Writing—Original Draft Preparation, A.S. and J.W.; Writing—Review & Editing, A.S., J.W. and S.F. Supervision, J.W.; Project Administration, J.W.; Funding Acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: A portion of this research was funded by the Israel Science Foundation (grant number 779/10).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Bar-Ilan University and the Israel Ministry of Education Chief Scientist's office.

Informed Consent Statement: Informed consent was obtained from parents of all subjects involved in the study.

Data Availability Statement: The data reported in this study are available upon request from the corresponding author. The data are not publicly available as yet.

Acknowledgments: We are thankful to the preschool children for their participation and their parents for giving us the possibility to conduct research on their children. We greatly appreciate the kindness and help of the kindergarten teachers and assistants.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Estimates of the GLMM model predicting the frequency of non-targeted CS by CS condition.

Fixed Factor	*AIC*	*BIC*	*logLik*	*Deviance*	*Chisq*	*Df*	*p*
Condition	1000.9	1034.8	−493.45	986.9	53.213	5	**<0.001**
Language	976.25	1015.0	−480.13	960.25	26.645	1	**<0.001**
Condition * Language	975.45	1038.4	−474.72	949.45	10.804	5	0.055
Group	972.1	1015.7	−477.05	954.1	6.1525	1	**0.013**
Group * Condition	978.45	1046.2	−475.23	950.45	3.6488	5	0.601
Group * Language	965.66	1014.1	−472.83	945.66	8.4373	1	**0.004**
Group * Condition * Language	978.31	1099.3	−464.15	928.31	17.359	15	0.298
Age	966.85	1020.1	−472.42	944.85	0.8144	1	0.367
AoO	967.3	1020.5	−472.65	945.3	0.3672	1	0.545
Non-verbal IQ (Raven)	967.05	1020.3	−472.53	945.05	0.6128	1	0.434

Table A2. Estimates of the GLMM model predicting the frequency of non-targeted CS by Sentence constituent.

Fixed Factor	*AIC*	*BIC*	*logLik*	*Deviance*	*Chisq*	*Df*	*p*
Sentence constituent	810.08	826.67	−401.04	802.08	94.327	2	**<0.001**
Language	783.07	803.8	−386.54	773.07	29.008	1	**<0.001**
Sentence constituent * Language	777.73	806.77	−381.87	763.73	9.3401	2	**0.009**
Group	773.87	807.06	−378.93	757.87	5.8628	1	**0.015**
Group * Sentence constituent	766.61	808.09	−373.3	746.61	11.262	2	**0.004**
Group * Language	762.37	808.01	−370.19	740.37	6.2317	1	**0.013**
Group * Sentence constituent * Language	763.33	817.26	−368.66	737.33	3.0448	2	0.218

Note: * signals interaction.

Table A3. Estimates of the GLMM model predicting the frequency of non-targeted CS by Constituent within the Temporal phrase.

Fixed Factor	AIC	BIC	logLik	Deviance	Chisq	Df	*p*
Constituent	660.86	683.04	−325.43	650.86	83.654	3	**<0.001**
Language	625.04	651.66	−306.52	613.04	37.818	1	**<0.001**
Constituent * Language	627.67	667.6	−304.83	609.67	3.3716	3	0.338
Group	625.37	656.42	−305.68	611.37	1.676	1	0.196
Group * Constituent	626.47	670.83	−303.23	606.47	6.5763	4	0.160
Group * Language	620.58	656.07	−302.29	604.58	8.466	2	**0.015**
Group * Language * Constituent	628.75	704.17	−297.38	594.75	9.8221	9	0.365

Notes

1 These terms are used interchangeably in the present paper. We also use 'non-targeted CS'.

References

Alloway, Tracy Packiam, and Susan Gathercole. 2005. Working memory and short-term sentence recall in young children. *European Journal of Cognitive Psychology* 17: 207–20. [CrossRef]

Altman, Carmit, Efrat Harel, Natalia Meir, Perry Iluz-Cohen, Joel Walters, and Sharon Armon-Lotem. 2021. Using a Monolingual Screening Test for Assessing Bilingual Children. *Clinical Linguistics & Phonetics*. Online ahead of print. [CrossRef]

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, 5th ed. Arlington: American Psychiatric Association.

Armon-Lotem, Sharon, and Natalia Meir. 2016. Diagnostic accuracy of repetition tasks for identification of Specific Language Impairment (SLI) in bilingual children: Evidence from Russian and Hebrew. *International Journal of Language & Communication Disorders* 51: 715–31.

Armon-Lotem, Sharon, and Joel Walters. 2010. An approach to differentiating bilingualism and language impairment. In *The Handbook of Psycholinguistic & Cognitive Processes: Perspectives in Communication Disorders*. Edited by Jackie Guendouzi, Filip Loncke and Mandy Williams. London: Routledge/Taylor & Francis, pp. 465–89.

Armon-Lotem, Sharon, Gabi Danon, and Joel Walters. 2008. The use of prepositions by bilingual SLI children: The relative contribution of representation and processing. In *Proceedings of GALA (Generative Approaches to Language Acquisition*. Edited by Anna Gavarró and Maria João Freitas. Cambridge: Cambridge Scholars Publishing, pp. 41–46.

Azuma, Shoji, and Richard P. Meier. 1997. Open class and closed class: Sentence-imitation experiments on intrasentential code-switching. *Applied Psycholinguistics* 18: 257–76. [CrossRef]

Basnight-Brown, Dana M., and Jeanette Altarriba. 2007. Code-switching and code-mixing in bilinguals: Cognitive, developmental, and empirical approaches. In *Speech and Language Disorders in Bilinguals*. Edited by Alfredo Ardila and Eliane Ramos. New York: Nova Science Publishers, Inc., pp. 69–89.

Bedore, Lisa, and Elizabeth Peña. 2008. Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism* 11: 1–29. [CrossRef]

Bobb, Susan C., and Zofia Wodniecka. 2013. Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology* 25: 568–85. [CrossRef]

Bobb, Susan C., Laila Y. Drummond Nauck, Nicole Altvater-Mackensen, Katie Von Holzen, and Nivedita Mani. 2016. Listening with your cohort: Do bilingual toddlers co-activate cohorts from both languages when hearing words in one language alone? In *Cognitive Control and Consequences of Multilingualism*. Edited by John W. Schwieter. Amsterdam: John Benjamins, pp. 47–69.

Boerma, Tessel, and Elma Blom. 2020. Effects of developmental language disorder and bilingualism on children's executive functioning: A longitudinal study. *Research in Developmental Disabilities* 107: 103782. [CrossRef] [PubMed]

Branzi, Francesca, Marco Calabria, Maria Lucrezia Boscarino, and Albert Costa. 2016. On the overlap between bilingual language control and domain-general executive control. *Acta Psychologica* 166: 21–30. [CrossRef] [PubMed]

Butterworth, Brian, Ruth Campbell, and David Howard. 1986. The uses of short-term memory: A case study. *The Quarterly Journal of Experimental Psychology* 38: 705–37. [CrossRef] [PubMed]

Coady, Jeffry A., Julia L. Evans, and Keith R. Kluender. 2010. The Role of phonotactic frequency in sentence repetition by children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* 53: 1401–16. [CrossRef]

Costa, Albert. 2005. Lexical Access in Bilingual Production. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Anette M. B. de Groot. Oxford: Oxford University Press, pp. 308–25.

Costa, Albert, and Mikel Santesteban. 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. [CrossRef]

Declerck, Mathieu, and Andrea M. Philipp. 2015. A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review* 22: 1630–45.

Declerck, Mathieu, Kristin Lemhöfer, and Jonathan Grainger. 2017. Bilingual language interference initiates error detection: Evidence from language intrusions. *Bilingualism: Language and Cognition* 20: 1010–16. [CrossRef]

Degani, Tamar, Varda Kreiser, and Rama Novogrodsky. 2019. The joint effects of bilingualism, DLD and item frequency on children's lexical-retrieval performance. *International Journal of Language & Communication Disorders* 54: 485–98.

Ebbinghaus, Hermann. 1913. *On Memory: A Contribution to Experimental Psychology*. New York: Teachers College Press.ade.

Ebert, Kerry Danahy. 2014. Role of auditory non-verbal working memory in sentence repetition for bilingual children with primary language impairment. *International Journal of Language & Communication Disorders* 49: 631–36.

Ebert, Kerry Danahy, and Kathryn Kohnert. 2016. Language learning impairment in sequential bilingual children. *Language Teaching* 49: 301–38. [CrossRef]

Festman, Julia, Antoni Rodriguez-Fornells, and Thomas F. Münte. 2010. Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behavioral and Brain Functions* 6: 5. [CrossRef]

Fricke, Melinda, and Gerrit Jan Kootstra. 2016. Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language* 91: 181–201. [CrossRef]

Friedmann, Naama, and Rama Novogrodsky. 2008. Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI. In *Language Acquisition and Development*. Edited by Anna Gavarró and Maria João Freitas. Cambridge: Cambridge Scholars Press, pp. 205–17.

Genesee, Fred, and Elena Nicoladis. 2009. Bilingual first language acquisition. In *Blackwell Handbook of Language Development*. Edited by Erika Hoff and Marilyn Shatz. Oxford: Blackwell Publishers, pp. 324–42.

Gollan, Tamar H., and Matthew Goldrick. 2016. Grammatical constraints on language switching: Language control is not just executive control. *Journal of Memory and Language* 90: 177–99. [CrossRef]

Gollan, Tamar H., Elizabeth R. Schotter, Joanne Gomez, Mayra Murillo, and Keith Rayner. 2014. Multiple levels of bilingual language control: Evidence from language intrusions in reading aloud. *Psychological Science* 25: 585–95. [CrossRef]

Gollan, Tamar H., Alena Stasenko, Chuchu Li, and David P. Salmon. 2017. Bilingual language intrusions and other speech errors in Alzheimer's disease. *Brain and Cognition* 118: 27–44. [CrossRef] [PubMed]

Goralnik, Esther. 1995. *Goralnik Diagnostic Test*. Even Yehuda: Matan. (In Hebrew)

Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef] [PubMed]
Green, David W., and Li Wei. 2014. A control process model of code-switching. *Language, Cognition and Neuroscience* 29: 499–511. [CrossRef]
Greene, Kai J., Elizabeth D. Peña, and Lisa M. Bedore. 2012. Lexical choice and language selection in bilingual preschoolers. *Child Language Teaching and Therapy* 29: 27–39. [CrossRef]
Grosjean, François. 1982. *Life with Two Languages: An Introduction to Bilingualism*. Cambridge: Harvard University Press.
Grosjean, François. 2001. The bilingual's language modes. In *One Mind, Two Languages: Bilingual Language Processing*. Edited by Janet L. Nicol. Malden: Blackwell, pp. 1–22.
Gross, Megan, and Margarita Kaushanskaya. 2018. Contributions of nonlinguistic task-shifting to language control in bilingual children. *Bilingualism: Language and Cognition* 21: 181–94. [CrossRef]
Gross, Megan, and Margarita Kaushanskaya. 2020. Cognitive and Linguistic predictors of language control in bilingual children. *Frontiers in Psychology* 11: 968–68. [CrossRef]
Gross, Megan, and Margarita Kaushanskaya. 2022. Language control and code-switching in Bilingual Children with Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research* 65: 1104–27. [CrossRef]
Gumperz, John. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.
Gutiérrez-Clellen, Vera F., Gabriela Simon-Cereijido, and Angela Erickson Leone. 2009. Code-switching in bilingual children with specific language impairment. *International Journal of Bilingualism* 12: 91–109. [CrossRef]
Hage, Simone Vasconcelos Rocha, Lidiane Sawasaki, Yvette Hyter, and Fernanda Dreux Miranda Fernandes. 2021. Social Communication and pragmatic skills of children with Autism Spectrum Disorder and Developmental Language Disorder. *CoDAS* 34: e20210075. [CrossRef]
Halmari, Helena. 2005. "I'm forgetting both": L1 maintenance and codeswitching in Finnish–English language contact. *The International Journal of Bilingualism* 9: 397–433. [CrossRef]
Hart, Betty, and Todd Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H. Brookes Publishing.
Håkansson, Gisela, Eva-Kristina Salameh, and Ulrika Nettelbladt. 2003. Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics* 41: 255–88. [CrossRef]
Henry, Lucy A., David J. Messer, and Gilly Nash. 2012. Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry* 53: 37–45. [CrossRef]
Heredia, Roberto R., and Jeanette Altarriba. 2001. Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science* 10: 164–68. [CrossRef]
Hofweber, Julia, Theodoros Marinis, and Jeanine Treffers-Daller. 2016. Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism* 6: 648–68. [CrossRef]
Iluz-Cohen, Perry, and Sharon Armon-Lotem. 2013. Language proficiency and executive control in bilingual children. *Bilingualism: Language and Cognition* 16: 884–899. [CrossRef]
Iluz-Cohen, Perry, and Joel Walters. 2012. Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism:Language and Cognition* 15: 1–17. [CrossRef]
Im-Bolter, Nancie, Janice Johnson, and Juan Pascual-Leone. 2006. Processing limitations in children with specific language impairment: The role of executive function. *Child Development* 77: 1822–41. [CrossRef]
Isham, William, and Harlan Lane. 1993. Simultaneous interpretation and the recall of source-language sentences. *Language and Cognitive Processes* 8: 241–64. [CrossRef]
Jia, Gisela, Kathryn Kohnert, Julissa Collado, and Francia Aquino-Garcia. 2006. Action naming in Spanish and English by sequential bilingual children and adolescents. *Journal of Speech, Language, and Hearing Research* 49: 588–602. [CrossRef]
Kapantzoglou, Maria, Julie Esparza Brown, Lauren M. Cycyk, and Gerasimos Fergadiotis. 2021. Code-Switching and Language Proficiency in Bilingual Children with and without Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research* 64: 1605–20. [CrossRef]
Klavans, Judith L. 1985. The syntax of code-switching: Spanish and English. In *Selected Papers from the XIIIth Linguistic Symposium on Romance Languages*. Edited by Larry D. King and Catherine A. Maley. Amsterdam: John Benjamins, pp. 213–31.
Kohnert, Kathryn. 2002. Picture naming in early sequential bilinguals: A 1-year follow-up. *Journal of Speech, Language, and Hearing Research* 45: 759–71. [CrossRef]
Kohnert, Kathryn. 2010. Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders* 43: 456–73. [CrossRef]
Kohnert, Kathryn, Elizabeth Bates, and Arturo E. Hernandez. 1999. Balancing bilinguals: Lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research* 42: 1400–13. [CrossRef] [PubMed]
Kolers, Paul A. 1966. Reading and talking bilingually. *Journal of Psychology* 79: 357–76. [CrossRef]
Komeili, Mariam, and Cloë R. Marshall. 2013. Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. *Clinical Linguistics & Phonetics* 27: 152–61.

Kootstra, Gerrit Jan, Janet G. van Hell, and Ton Dijkstra. 2010. Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language* 63: 210–31. [CrossRef]

Kootstra Gerrit Jan, Ton Dijkstra, and Janet G. van Hell. 2020. Interactive alignment and lexical triggering of code-switching in bilingual dialogue. *Frontiers in Psychology* 11: 1747. [CrossRef]

Kroll, Judith F., Susan C. Bobb, and Noriko Hoshino. 2014. Two Languages in Mind: Bilingualism as a Tool to Investigate Language, Cognition, and the Brain. *Current Directions in Psychological Science: A Journal of the American Psychological Society* 23: 159–63. [CrossRef]

Kuzyk, Olivia, Margaret Friend, Vivianne Severdija, Pascal Zesiger, and Diane Poulin-Dubois. 2020. Are there cognitive benefits of code-switching in bilingual children? A longitudinal study. *Bilingualism: Language and Cognition* 23: 542–53. [CrossRef]

Lahey, Margaret, and Jan Edwards. 1999. Naming errors of children with specific language impairment. *Journal of Speech and Hearing Research* 42: 195–205. [CrossRef]

Lenth, Russel, Paul Buerkner, Maxime Herve, Maarten Jung, Jonathon Love, Fernando Miguez, Hannes Riebl, and Henrik Singmann. 2022. Emmeans: Estimated Marginal Means, Aka Least-Squares Means [Computer Software]. Available online: https://github.com/rvlenth/emmeans (accessed on 9 September 2022).

Leonard, Laurence B. 2014. *Children with Specific Language Impairment*. Cambridge: MIT Press.

Lipski, John M. 2016. The role of unintentional/involuntary codeswitching: Did I really say that? In *Spanish-English Codeswitching in the Caribbean and the US*. Edited by Rosa E. Guzzardo, Catherine M. Tamargo Mazak and M. Carmen Parafita Couto. Amsterdam: John Benjamins, pp. 139–68.

Lipski, John M. 2017. Language switching constraints: More than syntax? Data from Media Lengua. *Bilingualism: Language and Cognition* 20: 722–46. [CrossRef]

Litcofsky, Kaitlyn A., and Janet G. Van Hell. 2017. Switching direction affects switching costs: Behavioral, ERP and time frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97: 112–39. [CrossRef] [PubMed]

Lüdecke, Daniel, Alexander Bartel, Carsten Schwemmer, Chuck Powell, Amir Djalovski, and Johannes Titz. 2022. sjPlot: Data Visualization for Statistics in Social Science. R Package Version 2.8.11 [Computer Software]. Available online: https://CRAN.R-project.org/package=sjPlot (accessed on 1 September 2022).

Mammolito, Angela P. 2015. Code-Switching and Grammatical Error Rates of Spanish-English Bilingual School-Aged Children with a Language Disorder. Master's thesis, Florida State University–Tallahassee, Tallahassee, FL, USA.

Marinis, Theodoros, and Sharon Armon-Lotem. 2015. Sentence repetition. In *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*. Edited by Sharon Armon-Lotem, Jan de Jong and Natalia Meir. Bristol: Multilingual Matters, pp. 95–124.

Marton, Klara, Brocha Abramoff, and Shari Rosenzweig. 2005. Social cognition and language in children with specific language impairment (SLI). *Journal of Communication Disorders* 38: 143–62. [CrossRef] [PubMed]

Marton, Klara, Luydmyla Kelmenson, and Milana Pinkhasova. 2007. Inhibition control and working memory capacity in children with SLI. *Psychologia* 50: 110–21. [CrossRef] [PubMed]

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

McClure, Erica. 1981. Formal and functional aspects of the codeswitched discourse of bilingual children. In *Latino Language and Communicative Behavior*. Edited by Richard P. Durán. Norwood: Ablex, pp. 69–94.

McMillen, Stephanie, Zenzi M. Griffin, Elizabeth D. Peña, Lisa M. Bedore, and Gary M. Oppenheim. 2020. "Did I Say Cherry?" Error patterns on a blocked cyclic naming task for bilingual children with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research* 63: 1148–64. [CrossRef]

Meir, Natalia, and Sharon Armon-Lotem. 2015. Disentangling bilingualism from SLI in Heritage Russian: The impact of L2 properties and length of exposure to the L2. In *Language Acquisition and Development: Proceedings of GALA 2013*. Edited by Cornelia Hamann and Esther Ruigendijk. Cambridge: Cambridge Scholars Publishing, pp. 299–314.

Meir, Natalia, Joel Walters, and Sharon Armon-Lotem. 2015. Disentangling SLI and bilingualism using sentence repetition tasks: The impact of L1 and L2 properties. *International Journal of Bilingualism* 20: 421–52. [CrossRef]

Modigliani, Vito, and Donald G. Hedges. 1987. Distributed rehearsals and the primacy effect in single-trial free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13: 426–36. [CrossRef]

Montanari, Simona, Wendy Ochoa, and Kaveri Subrahmanyam. 2019. A longitudinal investigation of language mixing in Spanish-English dual language learners: The role of language proficiency, variability, and sociolinguistic factors. *Journal of Child Language* 46: 913–37. [CrossRef]

Montgomery, James W. 2002. Examining the nature of lexical processing in children with specific language impairment: Temporal processing or processing capacity deficit? *Applied Psycholinguistics* 23: 447–70. [CrossRef]

Myers-Scotton, Carol. 1993a. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, Carol. 1993b. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Oxford University Press.

Norbury, Courtenay F., Tracey Gemmell, and Rhea Paul. 2014. Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language* 41: 485–510. [CrossRef]

Novogrodsky, Rama, and Natalia Meir. 2020. Multilingual children with special needs in early education: Communication is the key. In *The handbook of Early Language Education*. Edited by Mila Schwartz and Danijela Prošić-Santovac. Berlin: Springer, pp. 1–29.

O'Toole, Ciara, and Tina M. Hickey. 2012. Diagnosing language impairment in bilinguals: Professional experience and perception. *Child Language Teaching and Therapy* 29: 91–109. [CrossRef]

Paradis, Johanne. 2010. The interface between bilingual development and specific language impairment. *Applied Psycholinguistics* 31: 3–28. [CrossRef]

Peynircioğlu, Zehra F., and Aydyn Y. Durgunoğlu. 2002. Code switching in bilingual preschool children. In *Bilingual Sentence Processing*. Edited by Roberto R. Heredia and Jeanetter Altarriba. Amsterdam: Elsevier, pp. 339–58.

Peña, Elizabeth D., Lisa M. Bedore, Mirza J. Lugo-Neris, and Nahar Albudoor. 2020. Identifying developmental language disorder in school age bilinguals: Semantics, grammar, and narratives. *Language Assessment Quarterly* 17: 541–58. [CrossRef] [PubMed]

Poulisse, Nanda. 2000. Slips of the tongue in first and second language production. *Studia Linguistica* 54: 136–49. [CrossRef]

Poulisse, Nanda, and Theo Bongaerts. 1994. First language use in second language production. *Applied Linguistics* 15: 36–57. [CrossRef]

Pratt, Amy, Elizabeth Peña, and Lisa Bedore. 2021. Sentence repetition with bilinguals with and without DLD: Differential effects of memory, vocabulary, and exposure. *Bilingualism: Language and Cognition* 24: 305–18. [CrossRef]

Prior, Anat, and Tamar H. Gollan. 2013. The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology* 25: 622–45. [CrossRef]

R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raichlin, Rina, Joel Walters, and Carmit Altman. 2019. Some wheres and whys in bilingual codeswitching: Directionality, motivation and locus of codeswitching in Russian-Hebrew bilingual children. *The International Journal of Bilingualism* 23: 629–50. [CrossRef]

Rast, Rebekah, and Jean-Yves Dommergues. 2003. Towards a characterisation of saliency on first exposure to a second language. In *EUROSLA Yearbook: Papers from the Annual Conference of the European Second Language Acquisition Association*. Edited by Susan Foster-Cohen and Simona Pekarek-Doehler. Amsterdam: John Benjamins, vol. 3, pp. 131–56.

Raven, John C. 1998. *Raven's Coloured Progressive Matrices*. New York: Kluwer Academic/Plenum Publishers.

Ribot, Krystal M., and Erika Hoff. 2014. "¿Cómo estas?" "I'm good." Conversational code-switching is related to profiles of expressive and receptive proficiency in Spanish-English bilingual toddlers. *International Journal of Behavioral Development* 38: 333–41. [CrossRef]

Rice, Mabel L., Janna B. Oetting, Janet Marquis, John Bode, and Soyeong Pae. 1994. Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research* 37: 106–22. [CrossRef]

Riches, Nick G. 2012. Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language and Communication Disorders* 47: 499–510. [CrossRef] [PubMed]

Rothweiler, Monika. 2007. Multilingualism and language impairments. In *Handbook of Multilingualism and Multilingual Communication*. Edited by Peter Auer and Li Wei. Berlin: de Gruyter Mouton, pp. 229–46.

Ruigendijk, Esther, Gerd Hentschel, and Jan Patrick G. Zeller. 2016. How L2-learners' brains react to code-switches: An ERP study with Russian learners of German. *Second Language Research* 32: 197–223. [CrossRef]

Schotter, Elizabeth R., Chuchu Li, and Tamar H. Gollan. 2019. What reading aloud reveals about speaking: Regressive saccades implicate a failure to monitor, not inattention, in the prevalence of intrusion errors on function words. *Quarterly Journal of Experimental Psychology* 72: 2032–45. [CrossRef] [PubMed]

Sheng, Li, Elizabeth D. Peña, Lisa M. Bedore, and Christine E. Fiestas. 2012. Semantic deficits in Spanish-English bilingual children with language impairment. *Journal of Speech, Language, and Hearing Research* 55: 1–15. [CrossRef]

Smolak, Erin, Stephanie de Anda, Bianka Enriquez, Diane Poulin-Dubois, and Margaret Friend. 2020. Code-switching in young bilingual toddlers: A longitudinal, cross-language investigation. *Bilingualism: Language and Cognition* 23: 500–18. [CrossRef]

Soesman, Aviva. 2018. Codeswitching in English-Hebrew Bilingual Children with Typical Development and Specific Language Impairment (SLI): Effects of Syntactic Category, Sentence Position and Directionality. Doctoral Dissertation, Bar-Ilan University, Ramat Gan, Israel.

Soesman, Aviva, and Joel Walters. 2021. Codeswitching within prepositional phrases: Effects of switch site and directionality. *The International Journal of Bilingualism* 25: 747–71. [CrossRef]

Tallal, Paula, and Rachel Stark. 1981. Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America* 69: 568–79. [CrossRef]

Tan, Lydia, and Geoff Ward. 2000. A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 1589–625. [CrossRef]

Thomas, Michael S. C., and Alan Allport. 2000. Language switching costs in bilingual visual word recognition. *Journal of Memory and Language* 43: 44–66. [CrossRef]

Thordardottir, Elin. 2015. The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology* 17: 97–114. [CrossRef]

Tomblin, J. Bruce, Nancy L. Records, Paula Buckwalter, Xuyang Zhang, Elaine Smith, and Marlea O'Brien. 1997. Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research* 40: 1245–60. [CrossRef] [PubMed]

Treffers-Daller, Jeanine. 1993. *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective (Topics in Sociolinguistics, 9)*. Berlin: Mouton de Gruyter.

Van Hell, Janet G., Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2015. Intra-sentential code-switching: Cognitive and neural approaches. In *The Cambridge Handbook of Bilingual Processing*. Edited by John W. Schwieter. Cambridge: Cambridge University Press, pp. 459–82.

Vinther, Thora. 2002. Elicited imitation: A brief overview. *International Journal of Applied Linguistics* 12: 54–73. [CrossRef]

Von Studnitz, Roswitha E., and David W. Green. 2002. The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition* 5: 241–51. [CrossRef]
Vu, Jennifer A., Alison L. Bailey, and Carollee Howes. 2010. Early cases of code-switching in Mexican-heritage children: Linguistic and sociopragmatic considerations. *Bilingual Research Journal* 33: 200–19. [CrossRef]
Walters, Joel. 2005. *Bilingualism: The Sociopragmatic-Psycholinguistic Interface*. Mahwah: Erlbaum.
Wei, Li. 1998. The 'why' and 'how' questions in the analysis of conversational code-switching. In *Code-Switching in Conversation*. Edited by Peter Auer. London: Routledge, pp. 165–76.
Wiig, Elisabeth H., Wayne A. Secord, and Mara E. Semel. 2004. *Clinical Evaluation of Language Fundamentals–Preschool 2*. Toronto: Harcourt/Psych Corporation.
Yim, Dongsun, and John Rudoy. 2013. Implicit statistical learning and language skills in bilingual children. *Journal of Speech, Language, and Hearing Research* 56: 310–22. [CrossRef]
Zentella, Ana C. 1997. *Growing Up Bilingual: Puerto Rican Children in New York*. Hoboken: Blackwell Publishing.
Zheng, Xiaochen, Ardi Roelofs, and Kristin Lemhöfer. 2018. Language selection errors in switching: Language priming or cognitive control? *Language, Cognition and Neuroscience* 33: 139–47. [CrossRef]

languages

Article

Is There an Effect of Diglossia on Executive Functions? An Investigation among Adult Diglossic Speakers of Arabic

Najla Alrwaita [1], Lotte Meteyard [1], Carmel Houston-Price [1] and Christos Pliatsikas [1,2,*]

[1] School of Psychology and Clinical Language Sciences, University of Reading, Reading RG6 6AL, UK
[2] Centro de Investigación Nebrija en Cognición, Universidad Nebrija, 28015 Madrid, Spain
* Correspondence: c.pliatsikas@reading.ac.uk

Abstract: Recent studies investigating whether bilingualism has effects on cognitive abilities beyond language have produced mixed results, with evidence from young adults typically showing no effects. These inconclusive patterns have been attributed to many uncontrolled factors, including linguistic similarity and the conversational contexts the bilinguals find themselves in, including the opportunities they get to switch between their languages. In this study, we focus on the effects on cognition of diglossia, a linguistic situation where two varieties of the same language are spoken in different and clearly separable contexts. We used linear mixed models to compare 32 Arabic diglossic young adults and 38 English monolinguals on cognitive tasks assessing the executive function domains of inhibition, and switching. Results revealed that, despite both groups performing as expected on all tasks, there were no effects of diglossia in any of these domains. These results are discussed in relation to the Adaptive Control Hypothesis. We propose that any effects on executive functions that could be attributed to the use of more than one language or language variety may not be readily expected in contexts with limited opportunities for switching between them, especially in younger adults.

Keywords: diglossia; bilingualism; cognition; executive functions

Citation: Alrwaita, Najla, Lotte Meteyard, Carmel Houston-Price, and Christos Pliatsikas. 2022. Is There an Effect of Diglossia on Executive Functions? An Investigation among Adult Diglossic Speakers of Arabic. *Languages* 7: 312. https://doi.org/10.3390/languages7040312

Academic Editors: John W. Schwieter, Juana M. Liceras and Raquel Fernández Fuertes

Received: 1 April 2022
Accepted: 28 November 2022
Published: 16 December 2022

1. Introduction

The effects of using two languages on cognition have been widely discussed in the psycholinguistic literature on bilingualism. It is widely accepted that the two languages of a bilingual are constantly active (Kroll and Bialystok 2013; Marian and Spivey 2003). In order to select the language that is appropriate to a given context, a bilingual must access domain general mechanisms that are not restricted to language processing (Green and Abutalebi 2013). These mechanisms are thought to include executive functions such as inhibition, switching and updating (Miyake et al. 2000). Significantly better performance for bilinguals compared to monolinguals in non-linguistic tasks tapping into these functions have led several researchers to claim cognitive advantages in bilinguals, which stem from their long-term greater reliance on these executive functions; however, this claim remains controversial, especially as a far as young bilinguals are concerned (for a comprehensive review, see Valian (2015)). Much less is known about the effects on cognition of speaking two *varieties* of a single language, in linguistic situations such as bidialectalism and diglossia; indeed, the limited available evidence has only added to the controversy in the field (Alrwaita et al. 2022). Many explanations have been put forward to explain the contradictions in the evidence, including some relating to the type and frequency of opportunities for switching between languages/varieties. One of the most prominent proposals is the Adaptive Control Hypothesis (Green and Abutalebi 2013), which predicts that different bilingual conversational contexts would require different degrees of reliance on executive functions, and as a result, modulate the extent of the relevant effects on domain general cognition (Bialystok et al. 2006; Costa et al. 2009). One of the contexts that the Adaptive Control Hypothesis describes is the Single Language Context (SLC), where limited opportunities for language switching occur, and as a result minimal, if any, effects on cognition should be expected compared to the Dual Language Context which allows for more switching (for details, see Section 1.3). In the present study, we focus on Arabic diglossia as a case for studying the effects of using two varieties of the same language on cognition. Arabic diglossia offers a good example of an SLC which provides limited opportunities for switching between the two varieties, allowing to test the predictions of the Adaptive Control Hypothesis for SLCs. The remaining sections of this introduction define diglossia, summarize the evidence on the effects of two languages, or two varieties of a single language, on cognition in young adults, before outlining the aims of the present study.

1.1. Defining Diglossia

According to Ferguson (1959), diglossia refers to the coexistence of two or more related varieties of the same language in one community. Ferguson described four linguistic situations as prototypes of diglossia, including Standard/Colloquial Arabic (in the Arabic-speaking world), Katharevousa/Dhimotiki (in Greece), Standard/Swiss German (in German), and Standard/Creole French (in Haiti). In such situations one variety functions as the “standard” formal language, while the other is typically a regional dialect. The two varieties are very divergent: the standard or “high” (H) variety is highly codified, more grammatically complex, often the language used for written literature or for educational or religious reasons, and learned through formal education. As a result, a person speaking the high variety is often seen as more educated or of higher status. The “low” (L) variety tends to be less grammatically complex, used for everyday conversations in informal settings, and is learnt at home.

As already mentioned, one of the prototypical diglossic situations is diglossia in Arabic-speaking countries. Modern Standard Arabic (MSA) is the common high variety which is used across Arabic-speaking nations in formal contexts and found in the literature, governance, religious discourse and formal speeches. Low varieties of Arabic, known as colloquial dialects, vary between countries and are used in informal communication, in music, films, sport and everyday conversation (Albirini 2016). These two varieties also have different levels of language complexity. For example, H includes an elaborate system of inflections, such as number, gender, person, case and definiteness. This results in a more complex grammatical phrases in H compared to L. L, on the other hand, is known to have a simplified morphological system which lacks inflections (Albirini 2016). From a syntactic perspective, H is known to have a SVO word order while L has an SVO order (Boudelaa and Marslen-Wilson 2013).

In contrast to the disappearing diglossic situations in other countries (e.g., Greece) (Frangoudaki 1992), diglossia remains a defining characteristic of the Arab world. This serves to maintain the Islamic heritage and the language of the Qura’an (Amara and Mar’I 2002), enforces a sense of nationalism and functions as a unifying force amongst Arab countries (Palmer 2008). It is worth noting that Arabic diglossia takes different forms in different countries, and several factors affect the local Arabic variety that is spoken. For example, Rosenhouse and Goral (2006) divided Arabic dialects on the basis of four categories: First, geographical dimension, separating Eastern dialects, those spoken in Saudi Arabia, Yemen, Kuwait, Oman and the United Arab Emirates, from Western dialects, including those spoken in Morocco, Algeria, Tunisia, and Libya; second, sociolinguistic background, including Bedouin vs. al-Hadar backgrounds; third, religious background (Islam, Christianity or Judaism); and last, gender, age and education (Alsahafi 2016). It is important to note that mutual intelligibility often exist between local varieties, and it has been suggested that the more geographically close countries are, the more intelligibility is found between the speakers of these countries’ varieties. It has also been reported that Arabic speaking individuals often modify their variety to approach another individual’s variety (Giles 1973; S’hiri 2002). This usually happens during intralingual contact and the process is referred to as convergent accommodation (Alsahafi 2016). For instance, dialects of Western Arab countries, such as Tunisia and Morocco, have diverged from Standard Arabic more noticeably than dialects of Eastern Arab countries, such as Jordanian or Palestinian. Thus, it is expected that Moroccans, for instance, would modify their dialect when speaking to Jordanians (Alsahafi 2016). Diglossia in the Arab world, therefore, constitutes an ideal candidate for studying the effects of using two varieties of a language on cognition, a phenomenon that remains under- researched (Alrwaita et al. 2022).

1.2. The Effects of Bilingualism on Executive Functions

Before considering the effects of diglossia on cognition, it is useful to review the evidence relating to bilingualism, which has received greater attention from researchers. As discussed earlier, the juggling of two languages in the bilingual mind is thought to activate domain general cognitive mechanisms that support executive functions. Thus, bilinguals are "trained" to prevent interference from the non-target language in order to use the target language (Bialystok et al. 2012; Carlson and Meltzoff 2008). The need to constantly inhibit interference is thought to be the origin of the reported enhanced inhibition abilities in bilinguals compared to monolinguals, which are seen even in non-linguistic tasks (Calvo and Bialystok 2014; Costa et al. 2008; Kroll and Bialystok 2013; Ross and Melinger 2017). Inhibition, the ability to suppress attention to, or ignore, misleading information in order to attend to an appropriate target, along with switching, the ability to switch from one task to another, and updating, the ability to temporarily hold and update information in working memory for processing, are fundamental to executive functions (Miyake et al. 2000) and the most commonly investigated domains in bilingual studies. However, some controversy still exists on whether these effects are reliable, especially as far as young adults are concerned. For example, while several studies have reported benefits for bilinguals across a variety of tasks tapping different executive function domains (Bialystok 2011; Costa et al. 2008, 2009; Emmorey et al. 2008; Soveri et al. 2011; Treffers-Daller et al. 2020), a good number of studies has also failed to find such effects in some or all of executive function domains (Bialystok et al. 2012; Hernández et al. 2013; Kousaie and Phillips 2012; Paap and Greenberg 2013; Scaltritti et al. 2017); with some authors suggesting that benefits are most commonly found in older bilingual adults (Hilchey and Klein 2011). The lack of a consistent advantage for young adults has been explained in terms of the peak performance hypothesis, which states that, in contrast to children and older adults, young adults are at the peak of their cognitive performance (Bialystok et al. 2005), making it difficult to find evidence of advantages in one group over another (Paap et al. 2015).

However, other suggestions have also been put forward regarding the discrepancies of the findings in young adults, including how bilinguals are defined in each of these studies (Surrain and Luk 2019), as well as the quantity and quality of bilingual language use (Yang et al. 2016), which is the main focus of the present study. Indeed, it is perhaps surprising that the inconsistent findings in the literature have not been systematically linked to differences in linguistic context or amount of language switching of the bilingual participants studied. It has recently been suggested that the extent to which bilinguals engage their executive functions to resolve conflict between their two languages may be key to whether any domain general cognitive effects should be expected in bilinguals. According to the Adaptive Control Hypothesis (ACH) (Green and Abutalebi 2013), there are three conversational contexts that vary in the amount of switching between a bilingual's two languages, each impacting on executive functions differently: (a) the Single Language Context, where speakers use each language in a different context (e.g., one language at home and another at work); (b) the Dual Language Context, where speakers use the two languages in the same context, and switching occurs between sentences but not within sentences; (c) the Dense Code-Switching context, where speakers freely switch between the two languages even within the same sentence. According to the ACH, Dual Language Contexts require the most inhibitory control, followed by Single Language Contexts and the Dense Code-Switching contexts; therefore, Dual Language Contexts should result in particularly enhanced executive functions in bilinguals (Green and Abutalebi 2013). Following the Adaptive Control Hypothesis categorization, diglossia, and Arabic diglossia in particular, is closest to a Single Language Context, considering the exclusive contexts in which the varieties are used and the minimal switching between the two varieties. Therefore, Diglossia constitutes the appropriate test case to study the effects on cognition of such linguistic contexts.

1.3. The Effects of Speaking Two Varieties of One Language on Executive Functions

Only a few studies have investigated the effects of speaking two varieties of a single language on executive functions of young adults. We first review the evidence from bidialectalism, a linguistic situation that is very similar to diglossia but where the typical H and L varieties are better represented as a standard variety and a local dialect (Papapavlou and Pavlou 1998). Scaltritti et al. (2017) investigated whether the amount of exposure to the two varieties of a language relates to bidialectals' performance in inhibitory control tasks. In two experiments, Scaltritti et al. used a questionnaire about the participants' exposure to standard Italian and a local Venetian dialect, which they used to calculate a dialect familiarity score. Scaltritti and colleagues reported no significant correlation between this score and participants' performance in the Flanker and Simon tasks; moreover, they reported no benefits in these tasks for bidialectals when compared to Italian monolinguals, in line with studies that reported no benefits in bilingual young adults (Kousaie and Phillips 2012; Paap and Greenberg 2013). Scaltritti et al. (2017) argued that the limited opportunities for switching (which however was not measured directly) between the different language varieties in the Italian-Venetian context explains the lack of a bidialectal advantage in this study. In a different study, Poarch et al. (2019) tested the performance of a group of Swabian-German bidialectals aged 18–26 in the Simon and Flanker tasks. Poarch and colleagues used a dialect use questionnaire to place their participants in a continuum of dialect dominance, from Swabian-dominant to balanced Swabian-German speakers. Poarch et al. reported that dialect dominance negatively correlated with both Simon and Flanker effects, meaning in that the more dominant the participants were in Swabian, the better the cognitive control they demonstrated, which went against their original predictions. This finding was interpreted as evidence that using the non-standard dialect more than the standard one enhances inhibitory control more than equal use of the two dialects may do, directly contradicting the suggestions by Scaltritti et al. (2017). It is also worth noting that in Scaltritti et al. (2017) the lack of inhibitory control was linked to less switching, relating to less usage of the Venetian dialect.

To date, only one study has investigated the effects of diglossia on executive functions in young adults (Antoniou and Spanoudis 2020). This study was conducted in Cyprus, where Modern Standard Greek functions as the high variety and Cypriot Greek as the low variety (Antoniou et al. 2016). Diglossic and multilingual benefits were explored across the executive functions domains of inhibition, switching and updating (Miyake et al. 2000) by comparing diglossics (speakers of Cypriot Greek and Modern Standard Greek), multilingual participants (speakers of Cypriot Greek, Modern Standard Greek and another language) and monolingual speakers of Modern Standard Greek on the Stroop, Flankers, Color-Shape, N-back and Corsi block tasks. Antoniou and Spanoudis reported both multilingual and diglossic benefits across all executive functions components, but stopped short of reporting how these relate to opportunities for switching between languages. Notably, while these findings contradict the results of studies in bilinguals and bidialectals of a similar age, they corroborate those of Antoniou et al. (2016), who found diglossic benefits in Cypriot-Greek children. Some additional corroborating evidence also comes from a recent study comparing Arab diglossic older adults to bilinguals and monolinguals, which reported significant effects of diglossia, including a diglossic benefit in tasks tapping inhibitory control and switching (Alrwaita et al. n.d.). The authors interpreted their findings as the outcome of long-term experience in inhibiting the non-target language at SLC contexts.

The discrepancy between the evidence from diglossia and bidialectalism, two very similar linguistic situations, is intriguing, and potentially related to the different usage of the two language varieties in bidialectal and diglossic settings. Most importantly, the evidence from the few studies on diglossia contradicts the suggestion that executive control demands, and the related cognitive benefits, should be low in diglossic environments, given the clear separation between the contexts in which each variety is used.

1.4. The Role of Context and Switching Opportunities

As mentioned, both diglossia and bidialectalism can be considered to constitute SLCs in the Adaptive Control Hypothesis terms (Green and Abutalebi 2013); both offer limited opportunities for switching between the two varieties, which are used in a specific context. For these reasons, the Adaptive Control Hypothesis predicts that SLCs have minimal effects on executive functions compared to dual language contexts, which feature greater language switching requirements. The predictions of the Adaptive Control Hypothesis are supported by studies highlighting the role of language switching in modulating performances in executive function tasks and by the observation that benefits of bilingualism are more likely to be found in tasks that require more switching (Bialystok et al. 2006; Costa et al. 2009). It is therefore important to consider why, in conditions where we would expect to see minimal or no benefits (diglossic and bidialectal environments), contradictory results have been found in the only available study of diglossia (Antoniou and Spanoudis 2020).

While the bidialectal and diglossic situations appear to be similar, it is worthwhile considering the differences in how the H and L varieties function in each case, and how often speakers switch between them. First, bidialectalism develops in diglossic situations when H takes over L, meaning that the high variety is used for both formal and informal purposes (Rowe and Grohmann 2013). Second, diglossic native speakers are raised in homes where only L is used (Keller 1973). Finally, in bidialectalism H is seen as the more prestigious variety and learning it is considered more important than learning L, while in diglossia both varieties are expected to be learned (Rowe and Grohmann 2013). Linguistically, in bidialectalism H and L are similar in complexity, while in diglossia H is syntactically and grammatically more complex than L (Ferguson 1959).

There are also some differences in the functional uses of bidialectalism and diglossia, which suggest that diglossia better corresponds to an SLC than bidialectalism does. In diglossia, switching between the two dialects depends entirely on the activity, as all users understand both varieties, although the level of understating of the H variety can vary with education. However, in bidialectalism the use of the two varieties is overlapping (Shapiro et al. 1989); which means that both H and L can be used for formal purposes. It follows that we should be less likely to observe cognitive benefits in diglossia than in bidialectalism. As described above, the limited available evidence points in the opposite direction, suggesting that broad brush categorizations of environments as diglossic or bidialectal might not necessarily describe the important characteristics of those environments. For example, in diglossia, exposure to and usage of the two varieties might vary considerably across different environments (Kaye 2001). Given that lower exposure to one variety results in less switching between the two varieties, and therefore to less enhancement of executive control abilities (Kirk et al. 2014; Scaltritti et al. 2017), relative use of the two varieties in diglossia is likely an important factor to consider. Interestingly, the diglossic situation in Cyprus has been recently characterized as diglossic transitioning into type B diglossia or "diaglossia", where the use of the standard or H dialect is no longer restricted to written purposes, but also extends to oral communication. In such cases, the extent of code-switching depends more on the situation than on the activity (speaking/writing) (Auer 2011; Rowe and Grohmann 2013), a situation that is more similar to bidialectalism than true diglossia. This issue highlights the need for studies of diglossia in environments other than Cyprus.

Arabic offers the ideal test case for investigating the effects of diglossia on executive functions. First, in typical diglossia there is a clear separation between the contexts in which each variety is used, which should limit enhancement to executive functions (Costa et al. 2009). This separation is strongly enforced in Arabic diglossia, where there is rigid use of the high variety for formal purposes and the low variety for informal purposes (Albirini 2016), ruling out the possibility of an overlap in their use (Kaye 2001). According to Albirini (2016), even the most educated Arabs do not use the formal dialect (H) for informal purposes, or vice versa; doing so would be a clear violation of sociolinguistic norms. In contrast, switching between the high and the low dialects is very frequent in Cyprus, even when the speaker does not intend to switch (Pavlou 2004). Second, some Arabic speakers are rarely exposed to the high variety; even educated Arabs would find it difficult to hold conversations entirely in the high variety, and understanding it is even more difficult for uneducated Arabs (Kaye 2001). Again, this implies less switching between the two dialects and, as a result, limited effects on executive functions, if any at all. Arabic therefore offers a clear example of diglossia in which two language varieties are used in an SLC.

2. This Study

To investigate whether Arabic diglossia has an effect on executive functions, we compared young Arab diglossics (in their early to late twenties on average) to English-speaking monolinguals in a series of tasks that have been reported to demonstrate effects of diglossia in older adults (Alrwaita et al. n.d.), namely the Flanker, Stroop, and Color-shape tasks. If our results replicate previous evidence for diglossic benefits (Alrwaita et al. n.d.; Antoniou and Spanoudis 2020), this would suggest that diglossic situations, in general, enhance performance in tasks tapping executive functions, irrespective of the amount of language switching (which, as discussed, differs between Cyprus and the Arab world) or the age of the participants. Conversely, if we fail to report any effects of diglossia, and taking into account the findings from older Arab diglossics, this would constitute evidence for the peak performance hypothesis even for young diglossics in strict SLCs as they materialise in the Arab world.

3. Methods

3.1. Participants

70 young adults participated, 32 Arabic speaking diglossics (22 females, mean age = 29.6, SD = 5.7), and 38 English speaking monolinguals (34 females, mean age = 21.6; SD = 5.6)). Arabic speaking adults were recruited from the Prince Sultan University in Riyadh, Saudi Arabia. English speaking adults were recruited from the School of Psychology and Clinical Language Sciences at the University of Reading. One English-speaking individual was excluded for having bilingual parents and scoring significantly higher than other monolinguals in the Language and Social Background Questionnaire (LSBQ) (Anderson et al. 2018) (see below). Our final monolingual group consisted of 37 participants (34 females, mean age = 21.7; SD = 5.7) and was significantly younger than the diglossic group ($F(1,67) = 33.8$, $p < 0.001$).

3.2. Language Background and Proficiency Measures

Prior to the study, our monolingual participants were assessed using the Language and Social Background Questionnaire (LSBQ) (Anderson et al. 2018). The LSBQ examines the degree to which, and the domains in which, participants use English, and other languages (if applicable), in their daily lives. To measure proficiency, the questionnaire includes self- rating questions about reading, writing and listening skills in English/other languages. Factor scores were calculated for each participant for three domains: English proficiency, non- English social use, non-English home use. A composite 'bilingualism score' was computed by summing the factor scores weighted by the variance in each factor. According to Anderson et al. (2018), those with composite scores below −3.13 should be classified as monolinguals, while composite scores above 1.23 indicate bilinguals. Our monolingual young adults had an average composite score of −8.29 (SD = 1.2; range: −9.66 to −4.91), meaning they could all be safely classified as monolinguals according to the LSBQ. Moreover, monolinguals scored low in all three domains of switching found in the LSBQ, including switching with family (mean: −0.93, SD: 0.53, range: −1.12 to 0.96), switching with friends (mean: −0.44, SD: 0, range: −0.44 to −0.44), and switching in social media (mean: −0.45, SD: 0.9, range: −0.47 to −0.08).

To ensure that our diglossic group were indeed diglossics (knowing both the standard and the spoken variety), we adapted the LSBQ to investigate the degree of variety use, and the domains in which each variety was used (see Appendix A). Given that Standard Arabic is the language used for writing and reading, the Arabic version of LSBQ was conducted in Standard Arabic. Diglossics achieved an average composite score of −0.22 (SD = 2.6; range −6.62–4.92). According to Anderson et al. (2018), this score lies in the grey area between bilingualism and monolingualism. Crucially, results from the LSBQ revealed low scores in the three domains of switching, including with family (mean: −0.57; SD: 0.60; range −1.12 to 0.96), with friends (mean: −0.025; SD: 0.04; range (−0.44 to 1.09), and in social media (mean: 0.12, SD: 0.55; range −0.47 to 1.21). This shows that diglossics belong to an environment with limited amounts of switching (SLC); however, they significantly engage to switching more than monolinguals do in all three domains (all ps < 0.05).

Because the level of knowledge of Standard Arabic differs according to the level of education and exposure to the standard dialect (Kaye 2001), proficiency in Standard Arabic was measured using a vocabulary test designed by Masrai and Milton (2019). This pen-and-paper test comprises a checklist of 100 of the 50,000 most common Arabic words (generated by Kilgarriff et al. (2014), from the web-based corpus of Arabic (Sharoff 2006)). The test is divided into two parts, each including 20 words and 20 non-words intermixed. Participants indicated if they recognized each word by responding either "yes" or "no". The duration of the test is between 10–15 min. To estimate participant's vocabulary size, measured as words known out of 50,000 words, all "yes" answers to real words are given a score of 500 to form an unadjusted vocabulary score, and each "yes" answer to a false word deducts 2500 points from the unadjusted score to form an adjusted vocabulary score. The final adjusted score gives the participants total vocabulary knowledge. Based on this, our diglossic group had an average of 87.59% (SD = 12.17) of "yes" responses to real words, an average of 5.25% (SD = 3.69) to "yes" responses to non-words, and an estimated average of 30,672 (SD = 7329; median = 30,375; range 15,000–42,000) known words. No one was excluded from the diglossic group based on this score, as they all demonstrated at least basic competency in standard Arabic (Masrai and Milton 2019).

3.3. Executive Function Tasks

Our tasks tapped on the executive function domains of inhibition (Flanker and Stroop) and switching (Color-shape). The tasks were delivered using E- prime 2.0 software (Psychology Software Tools, Sharpsburg, PA, USA). All tasks were presented in a 15.6-inch computer screen, and all participants were tested in private rooms.

3.3.1. Inhibition-Flanker

This task had three conditions: Congruent, Incongruent and Neutral. In all conditions, a central arrow appeared on the screen and participants were asked to indicate if the arrow pointed to the right or left by pressing either (<) or (>) buttons on the keyboard. Participants were instructed to respond as fast and accurately as possible. In the Congruent condition, there were surrounding (flanking) arrows, which pointed in the same direction as the target central arrow (<<<<<). In the Neutral condition, the target central arrow appeared with dashes on the left and the right sides (–<–). In the Incongruent condition, there were surrounding arrows pointing in the opposite direction to the target central arrow (>><>>). Forty-four trials of each type were distributed across four blocks, each including 33 trials, 11 per condition, presented in a random order. Each trial began with a 250 ms fixation cross, followed by a stimulus lasting for 5000 ms, or until a response was provided. Trials were separated by a blank screen which appeared for 250 ms.

3.3.2. Inhibition-Stroop

In this task, a single word appeared on the screen and participants were asked to decide the color in which the word was written. The task consisted of three conditions: Congruent, Incongruent and Neutral. In Congruent trials, color words were presented that were consistent with the ink color; for example the word "green" was presented in green ink. In Neutral trials, non-color words (e.g., "dry") were presented in different ink colors (red, green or blue). In Incongruent trials, color words were presented in inconsistent ink color; for example, the word "green" was presented in red ink. The three different colors (red, green and blue) were assigned to three adjacent keyboard buttons and participants responded by clicking the corresponding button.

48 trials of each type were distributed across three blocks, each consisting of 48 trials in random order, with a total of 144 trials. Each trial began with a 250 ms fixation cross, followed by the presentation of the word for a maximum of 5000 ms, or until a response was provided. For Arabic-speaking participants, the same task was administrated translated in Arabic.

3.3.3. Switching–Color-Shape

In this task, participants were presented with three possible patterns in either blue, green or red. There were three blocks. First, in the color block, two patterns appeared on the screen and participants had to decide whether the patterns were in the same or different colors. Second, in the pattern block participants were presented with two patterns and were asked to decide whether they were the same or different. Third, in the switching block participants had to switch between attending to the color or the pattern. In each trial, a word (color/pattern) was presented at the top of the screen, indicating whether color or pattern should be responded to. To respond, participants had to press the S button on the keyboard for "same", or the D button for "different".

There was a total of 124 trials. The blocked color task and blocked pattern task each included 31 trials. The switching block included 62 randomized trials consisting of 31 'Stay' trials, where participants were asked to attend to the same property (color/pattern) as the previous trial, and 31 "Change" trials, where participants had to switch to attending to the other property. Each trial began with a 100 ms central fixation cross, followed by a stimulus presentation until a response was detected. There was a 250 ms blank screen between trials.

4. Results

Analyses of accuracy and reaction time (RT) data were run using generalized linear mixed-effects models (with a binomial link function for Accuracy data) and linear mixed effect models (for RT data) from the statistical package *lme4* (Bates et al. 2015), in R studio (version 4.0.3). Reaction times in the outlier data were removed by using a cut-off of 2 standard deviations above or below the task median. Because reaction times were not normally distributed, they were log- transformed prior to the analysis.

Model fitting followed the same procedure for all data. Age was entered as a covariate. Task Condition (e.g., congruent, incongruent, neutral; stay or change trials) were entered as main effects. Group (monolingual or diglossic) was entered as a main effect. We also tested the interaction between Group and Condition (i.e., whether monolingual or diglossic participants responded to the task conditions differently). First, a maximal model was fit with random intercepts for participants and correlated random slopes for within-subject conditions. If there were convergence issues with the maximal model, the first step was removing correlations between random effects. If convergence issues persisted (as they did for all models that initially had convergence issues for the maximal model) we then removed random slopes and ran a model with participant intercepts only. We compared model fit for the maximal model and the final simplified model, and not any differences in the model output where the simplified model did not fit the data as well as the maximal model. Where a factor had more than two levels (Flanker and Stroop tasks with incongruent, congruent and neutral conditions) we relevelled the reference factor to extract all possible comparisons between conditions, and used emmeans (Lenth et al. 2021) to report the relevant comparisons for the interaction between group (diglossics vs. monolinguals) and condition (incongruent, congruent, neutral).

4.1. Flanker

4.1.1. Accuracy

A maximal model with random intercepts for participants and slopes for condition varying over participant did not converge. We removed correlations for the random effects and re-ran the model, but convergence issues were again present. A model with participant intercepts only converged. The maximal model was a better fit to the data than the intercepts only model (X^2 = 12.985, df = 5, $p < 0.05$). Between the maximal and intercept only model, the only difference in significant results was that the maximal model did not find a significant difference between the Incongruent and Neutral conditions (Est = 0.646, SE = 0.424, Z = 1.525, $p = 0.13$). We report the results from the intercept only model (see Table 1), and interpret the main effect difference between Incongruent and Neutral conditions with caution.

Table 1. Accuracy results for the Flanker task.

Fixed Effects				
	Est/Beta	SE	z	*p* value
Intercept	2.83	1.01	2.797	$p < 0.01$ *
Age	0.074	0.033	2.24	$p < 0.05$ *
Condition:				
Congruent vs. Incongruent	−0.84	0.288	−2.909	$p < 0.005$ *
Congruent vs. Neutral	0.146	0.323	−0.454	$p = 0.65$
Incongruent vs. Neutral	0.69275	0.2759	2.511	$p < 0.05$ *
Group: Monolingual—Diglossic	0.26004	0.522	0.499	$p = 0.62$
Group x Condition interaction:				
Incongruent vs. Congruent: Monolingual vs. Diglossic	−0.663	0.37865	−1.753	$p = 0.08$
Congruent vs. Neutral: Monolingual vs. Diglossic	−0.21838	0.43023	−0.508	$p = 0.61$
Incongruent vs. Neutral: Monolingual vs. Diglossic	0.44522	0.35084	1.269	$p = 0.20$
Random Effects		Variance	S.D.	
		1.478	1.216	
Subject (Intercept)				
Model fit		Marginal		Conditional
R2		0.103		
				0.381

Note: * denotes a significant effect.

Results showed a main effect of Age (Beta = 0.074 (0.033), z = 2.24, $p < 0.05$) with older participants having higher accuracy. There was a difference between Congruent and Incongruent conditions (Beta = −0.84 (0.29), Z = −2.909, $p < 0.005$) with higher accuracy in the Congruent condition. There were no other significant main effects or interactions.

4.1.2. Reaction Times

For the RTs analysis, all incorrect trials were removed, affecting 7% of the trials pf the Diglossics and 4% of the trials of the Monolinguals.

Figure 1A illustrates RTs data for the Flanker task. The maximal model for Flanker RTs gave a convergence warning, which was repeated when the model was run without correlated slopes. An intercept only model converged. The maximal model was a better fit to the data (X^2 = 152.36, df = 5, $p < 0.001$) and the intercept only model showed marginally significant interactions where the maximal model did not. We therefore report the maximal model here, rather than report results from the intercept only model that may be spurious. The reported maximal model shows difficulties with convergence hence the large dfs reported for some comparisons (Table 2); note also the high correlations between intercepts and slopes, indicating that the slopes do not add substantial information over and above the intercepts for some conditions.

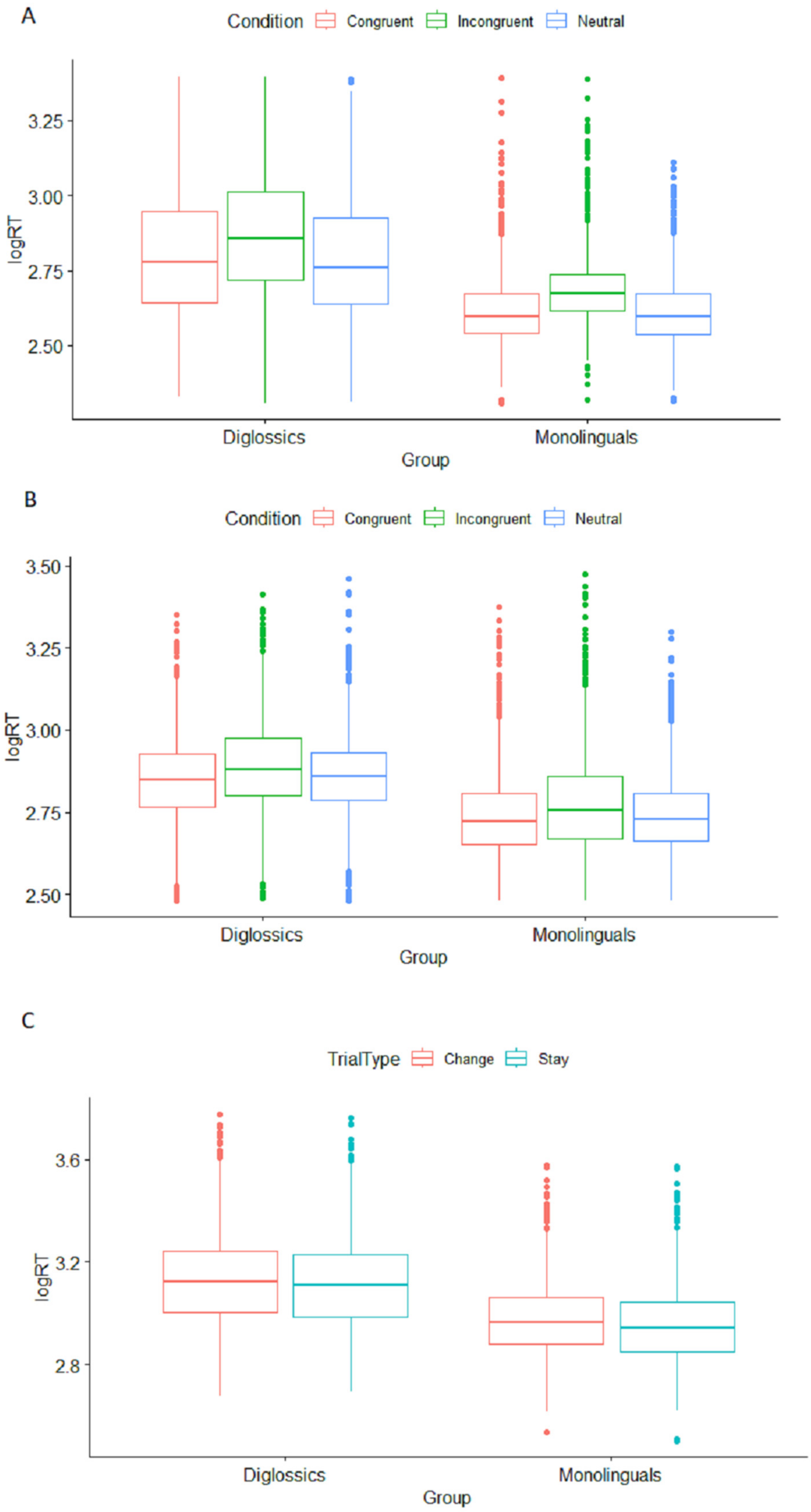

Figure 1. Reaction times for Diglossics and Monolinguals across conditions for (**A**) Flanker, (**B**) Stroop and (**C**) Color-shape tasks.

Table 2. Reaction time results for the Flanker task.

Fixed Effects					
	Est/Beta	SE	t value	df	*p* value
Intercept	2.713	0.073	37.27	68.49	$p < 0.001$ *
Age	0.003	0.002	1.071	66.12	$p = 0.288$
Condition:					
Congruent vs. Incongruent	0.08	0.008	9.781	68.1	$p < 0.001$ *
Congruent vs. Neutral	−0.012	0.005	−2.51	221.8	$p = 0.013$ *
Incongruent vs. Neutral	−0.09	0.009	−10.44	65.94	$p < 0.001$ *
Group: Monolingual—Diglossic	−0.15	0.035	−4.32	70.91	$p < 0.001$ *
Group x Condition interaction:					
Incongruent vs. Congruent: Monolingual vs. Diglossic	−0.01	0.011	−0.896	68.13	$p = 0.373$
Congruent vs. Neutral: Monolingual vs. Diglossic	0.009	0.006	1.443	222.87	$p = 0.151$
Incongruent vs. Neutral: Monolingual vs. Diglossic	0.019	0.012	1.6	65.93	$p = 0.115$
Random Effects					
		Variance	S.D.	Correlation	
Subject (Intercept)		0.015	0.13		
Condition (Slope)					
Congruent vs. Incongruent		0.0015	0.039	−0.09	
Congruent vs. Neutral		0.0001	0.01	−0.94	−0.26
Incongruent vs. Neutral		0.001	0.043	−0.45	0.97
Residual		0.0115	0.107		
Model fit					
R2		Marginal		Conditional	
		0.256			0.669

Note: * denotes a significant effect.

All task conditions showed significant differences, with RTs in the congruent condition being faster than those in the incongruent condition (Beta = 0.08 (0.008), t (68.1) = 9.781, $p < 0.001$), but slightly longer than those in the Neutral condition (Beta = −0.012 (0.005), t (221.8) = −2.51, $p < 0.05$). RTs in the incongruent condition were significantly longer than those in both the congruent and neutral conditions ((Beta = 0.074 (0.033), z = 2.24, $p < 0.05$)). There was a main effect of group, with Monolinguals having faster RTs than Diglossics (Beta = −0.15 (0.035), t (70.91) = −4.32, $p < 0.001$). There was no significant interaction between Group and task Condition.

4.2. Stroop

4.2.1. Accuracy

A maximal model with random intercepts for participants and slopes for condition varying over participant did not converge. We removed correlations for the random effects and re-ran the model, convergence issues were again present. A model with participant intercepts only converged. The intercept only model and the maximal model did not differ in how well they fit the data (X2 = 2.6328, df = 5, $p = 0.76$). We report results from the intercept only model. There were no significant effects (see Table 3).

4.2.2. Reaction Times

All incorrect trials were removed. This affected 2% of the trials of the Diglossic group, and 3% of the trials of the Monolingual group.

The maximal model with random intercepts and slopes for both subject and condition, converged without errors (see Table 4). Figure 1B illustrates RT data for the Stroop task.

There were significant main effects for all task Condition comparisons. RTs in the congruent condition were faster than RTs in the Incongruent condition (Beta = 0.04 (0.005), t (67.3) = 7.385, $p < 0.001$), and faster than RTs in the Neutral condition (Beta = 0.03 (0.005), t (67.7) = 2.74, $p < 0.01$). RTs in the Incongruent condition were slower than RTs in the Neutral condition (Beta = −0.03 (0.006), t (52.89) = −4.486, $p < 0.001$). There was a main effect of group, with Monolinguals having faster RTs than Diglossic participants (Beta = −0.18 (0.006), t (73) = −5.265, $p < 0.001$). There were no significant interactions between Group and Condition.

Table 3. Accuracy results for the Stroop task.

Fixed Effects					
	Est/Beta	SE	z value	*p* value	
Intercept	4.19	0.71	5.946	$p < 0.001$ *	
Age	0.032	0.02197	1.454	$p = 0.15$	
Condition:					
Congruent vs. Incongruent	−0.79	0.352	−2.243	$p < 0.05$	
Congruent vs. Neutral	−0.288	0.38474	−0.749	$p = 0.45$	
Incongruent vs. Neutral	0.50238	0.32119	1.564	$p = 0.12$	
Group: Monolingual—Diglossic					
Group x Condition interaction:					
Incongruent vs. Congruent: Monolingual vs. Diglossic	−0.055	0.419	−0.131	$p = 0.90$	
Congruent vs. Neutral: Monolingual vs. Diglossic	−0.29	0.451	−0.639	$p = 0.52$	
Incongruent vs. Neutral: Monolingual vs. Diglossic	−0.234	0.373	−0.626	$p = 0.53$	
Random Effects					
		Variance	S.D.		
Subject (Intercept)		0.4717	0.6868		
Model fit					
R2		Marginal		Conditional	
		0.088			0.202

Note: * denotes a significant effect.

Table 4. Reaction time results for the Stroop task.

Fixed Effects					
	Est/Beta	SE	t value	df	*p* value
Intercept	2.85	0.04	66.543	69.2	$p < 0.001$ *
Age	0.0001	0.001	−0.044	67.3	$p = 0.965$
Condition:					
Congruent vs. Incongruent	0.0396	0.005	7.385	67.3	$p < 0.001$ *
Congruent vs. Neutral	0.03	0.005	2.74	67.7	$p < 0.01$ *
Incongruent vs. Neutral	−0.0266	0.006	−4.486	52.89	$p < 0.001$ *
Group: Monolingual—Diglossic	−0.18	0.02		73	$p < 0.001$ *
Group x Condition interaction:					
Incongruent vs. Congruent: Monolingual vs. Diglossic	−0.00335	0.007	−0.459	67.8	$p = 0.65$
Congruent vs. Neutral: Monolingual vs. Diglossic	−0.009	0.0065	−1.43	68.4	$p = 0.16$
Incongruent vs. Neutral: Monolingual vs. Diglossic	−0.0059	0.008	−0.729	53.3	$p = 0.469$
Random Effects					
		Variance	S.D.	Correlation	
Subject (Intercept)		0.005	0.07		
Condition (Slope)					
Congruent vs Incongruent		0.0004	0.02	0.62	
Congruent vs Neutral		0.0002	0.013	−0.33	−0.36
Incongruent vs Neutral		0.0006	0.024	−0.76	0.76
Residual		0.013	0.14		
Model fit					
R2		Marginal		Conditional	
		0.154			0.405

Note: * denotes a significant effect.

4.3. Color-shape

4.3.1. Accuracy

The maximal model for Accuracy gave a convergence warning, which was repeated when the model was run with un-correlated slopes. When the model was run with intercepts only, no convergence warnings were given. The maximal model fit the data better than the model with intercepts only (X^2 = 76.441, df = 55, $p < 0.05$). However, the maximal model was clearly unreliable as all predictors were reported as highly significant with very large z values. Therefore, we report the model with intercepts only (see Table 5).

Table 5. Accuracy results for the Color-shape task.

Fixed Effects				
	Est/Beta	SE	z value	*p* value
Intercept	3.058	0.61	5.015	$p < 0.001$ *
Age	0.01	0.02	0.531	$p = 0.07$
Condition: Change vs. Stay	0.84	0.26	3.196	$p < 0.005$ *
Group: Monolingual—Diglossic	−0.522	0.287	−1.815	$p = 0.07$
Group x Condition interaction:				
Change vs. Stay: Monolingual vs. Diglossic	−0.322	0.31	−1.038	0.299
Random Effects				
		Variance	S.D.	
Subject (Intercept)		0.4182	0.6466	
Model fit *(for intercept only model, as conditional R2 was unavailable for the maximal model)*				
R^2		Marginal	Conditional	
		0.066	0.172	

Note: * denotes a significant effect.

There was a significant main effect of Condition, with Change trials having lower accuracy than Stay trials (Beta = 0.84 (0.26), z = 3.196, $p < 0.005$). There were no other significant effects.

4.3.2. Reaction Times

Along with extreme values, all incorrect trials were removed affecting 3%of the trials of the Diglossic group, and 6% of the trials of the Monolingual group.

Figure 1C illustrates RTs data for the Switching task. The maximal model for Switch RTs gave a convergence warning, which was repeated when the model was run with un-correlated slopes. We ran the model with intercepts for participants only, and this gave no convergence warnings. The intercept only model fit the data as well as a model with uncorrelated slopes (X^2 = 0.1, df = 2, $p = 0.9512$). We therefore report the results from the model with intercepts only (see Table 6).

There was a significant main effect of condition (Beta = −0.013 (0.005), z = −2.459, $p < 0.05$), with RTs in the Change condition being slower than those in the Stay condition. There was a significant main effect of Group (Beta = −0.132 (0.031), z = −4.2, $p < 0.001$), with Monolinguals having faster RTs as compared to Diglossics.

Table 6. Reaction time results for the Color-shape task.

Fixed Effects					
	Est/Beta	SE	t value	df	*p* value
Intercept	3.07	0.07	43.832	67.23	$p < 0.001$ *
Age	0.002	0.002	0.81	67.05	$p = 0.421$
Condition: Change vs. Stay	−0.013	0.005	−2.459	4804	$p < 0.05$ *
Group: Monolingual—Diglossic	−0.132	0.031	−4.2	68.86	$p < 0.001$ *
Group x Condition interaction					
Change vs. Stay: Monolingual vs. Diglossic	−0.008	0.007	−1.09	4804	$p = 0.276$
Random Effects					
		Variance	S.D.		
Subject (Intercept)		0.0111	0.1053		
Residual		0.0157	0.1253		
Model fit					
R^2		Marginal	Conditional		
		0.179	0.519		

Note: * denotes a significant effect.

4.4. Continuous Experience-Based Factors as Predictors of Diglossics' Performance

Following up from studies that used measures quantifying the bidialectal experiences as continuous predictors of their participants' performance (Poarch et al. 2019; Scaltritti et al. 2017), we ran two sets of regression analyses on the diglossics' data using our own measures as predictors of the Flanker effect (Incongruent-Congruent), the Stroop effect (Incongruent-Congruent), and the Switch effect (Change-Stay), by adding age and sex as covariates. The first measure was diglossics' proficiency in Modern Standard Arabic as measured by the Arabic vocabulary test. These analyses did not yield significant results for the Flanker effect ($p = 0.63$), the Stroop effect ($p = 0.593$), and the Switch effect (Change-stay) ($p = 0.63$). The second measure was the diglossic LSBQ composite score. These analyses did not yield significant results for the Flanker effect ($p = 0.77$) or the Switch effect ($p = 0.63$); however, the model was significant for the Stroop effect ($R^2 = 0.301$, $F(3, 28) = 4.02$, $p = 0.017$), with the LSBQ composite score demonstrating a positive correlation with the Stroop effect ($p = 0.04$), suggesting that the higher the composite score, the higher the Stroop effect. In order to explore this further, we ran separate models to investigate the effect of the three switching domains that the LSBQ measures. The models were not significant for switching with friends ($p = 0.628$) and for switching in social media ($p = 0.606$), but the effect was significant when switching with family was used as the predictor ($R^2 = 0.265$, $F(3, 28) = 3.37$, $p = 0.033$), which positively correlated with the Stroop effect ($p = 0.009$). This may have driven the significant effect of the composite score.

5. Discussion

Following suggestions that conversational contexts requiring different levels of switching affect executive functions differently (Green and Abutalebi 2013) this study examined whether diglossic participants (who, with minimal opportunities for code-switching between their two language varieties, correspond to Green & Abutalebi's SLC) perform similarly to monolinguals. We compared the performance of Arabic speaking young adults to English speaking monolinguals of a similar age on three tasks, Flanker, Stroop and Color-Shape, which tap the executive function domains of inhibition and switching (Miyake et al. 2000). In each task we compared performance between cognitively challenging and control conditions, namely between Congruent and Incongruent trials for the Flanker and Stroop tasks, and between Stay and Change trials for the Color-shape task. While all of our tasks yielded the expected pattern (i.e., lower accuracy and slower RTs for the more cognitively challenging conditions), the difference in performance across task levels did not differ between our diglossic and monolingual groups in any of the tasks. Specifically, our results did not reveal smaller cognitive costs for diglossics compared to monolinguals, a pattern that has been typically interpreted as a benefit for bilingual groups when compared to monolinguals (Valian 2015); in other words, our results suggest no diglossic benefits in any task. These findings will be discussed with reference to previous studies and theoretical proposals for the effect of using two languages on cognition.

It is worth noting that the monolingual group had shorter overall RTs than the diglossic group across all tasks (all ps < 0.001), while no such pattern was observed with respect to overall accuracy. While a bilingual benefits in global RTs have been often linked to the high monitoring skills of bilinguals compared to monolinguals (Hilchey and Klein 2011), our study fails to replicate these effects in a diglossic population; rather, our findings point toward a global disadvantage for the diglossic group. It is important to note that bilingual advantages in overall reaction times have not been replicated in all studies (Antón et al. 2019; Kirk et al. 2014). In order to explain this discrepancy, some have rejected bilingualism as a factor leading to faster overall RTs, and attributed such effects to unmatched external factors such as socioeconomic status (Antón et al. 2019). Since monolinguals in our study demonstrated faster RTs than diglossics, we cannot readily attribute this seeming diglossic "disadvantage" to the two groups' language experiences; indeed, we had no grounds to predict it, nor has it been reported previously. As suggested by Antón et al. (2019), this finding might be related to external factors which warrant further investigation. One candidate is task familiarity; our monolingual group were psychology students recruited for course credit who may have been familiar with the classic cognitive tasks we employed. In contrast, our diglossics were not recruited from a psychology department. Prior training has been shown to result in better performance in inhibition, switching and updating tasks (Chevalier et al. 2012; Hughes et al. 2009; Salminen et al. 2012) and to increased processing speed (Dux et al. 2009), and this suggestion might explain our overall RTs patterns. Another possible explanation could be that, while both groups still lie within the grey area of young adults as defined by bilingual studies (Donnelly et al. 2019), diglossics were slightly but significantly older than monolinguals, and it has been suggested that processing speed and reaction times decline with age (Ferreira et al. 2015). Nevertheless, the inclusion of age as a covariate in our models may this explanation less likely as a candidate.

An unexpected finding in our study was the positive correlation between the Stroop effect and measures of diglossic experience, namely the LSBQ composite score and the "switching with family" subscale of the same test. If anything, we would expect the opposite pattern, that is, the more experience the participants have in switching between languages, the easier it would be to process the incongruent trials, i.e., the smaller the Stroop effect would be. However, this prediction is usually drawn from findings from bilinguals not necessarily in SLCs, so it may not entirely apply to our diglossic environment and the particularities it encompasses. Taken at face value, this pattern suggests that people who switch in one context only may actually face difficulties in inhibitory control as it is measured by the Stroop task. Nevertheless, such effects that have occasionally be characterised as bilingual *disadvantages* (but see Luk (2022), for a criticism on this term), have been reported very rarely in the literature, and have been attributed to Type I errors (De Bruin et al. 2015). Based on this, our relatively small sample size, and the fact that such an effect was not observed for the related Flanker effect, we treat this finding with caution.

To our knowledge, this is the first study that reports no benefits in executive functions across the board in diglossics. Our findings contradict recent evidence suggesting some benefits in inhibition in older diglossics from the same linguistic environment (Alrwaita et al. n.d.), lending support to the peak performance hypothesis (Bialystok et al. 2005). However, our findings also contradict those from previous studies conducted in Cyprus by Antoniou et al. (2016), who found diglossic benefits in children (Antoniou et al. 2016) and young adults (Antoniou and Spanoudis 2020). The contradiction between our and those findings might be explained in light of the amount of switching required by our diglossic group. In Arabic the standard dialect and regional varieties are clearly separated by context, and orally switching between the two varieties is very unlikely (Albirini 2016). In this respect Arabic is a good example of an SLC as described by the Adaptive Control Hypothesis

(Green and Abutalebi 2013), where each language/dialect is used in a specific context and switching between them rarely occurs. This contrasts with the diglossic situation in Cyprus, where switching between the two varieties occurs frequently (Pavlou 2004), as the two varieties are not restricted to a specific context (Rowe and Grohmann 2013). Therefore, we suggest that the diglossic benefits found in those previous studies (Antoniou et al. 2016; Antoniou and Spanoudis 2020) are due to the diglossic situation in Cyprus allowing for switching between the two varieties. Importantly, our results are in line with Scaltritti et al. (2017) who reported no bidialectal benefits in executive functions amongst young adults. This pattern of findings across studies calls for careful attention to the context in which each language is used in each environment, irrespective of whether the case in hand is diglossia, bidialectalism and bilingualism. As discussed previously, the type and amount of code-switching cannot be assumed to be constant in each of these situations. Together, these findings suggest that it is overly simplistic to label a diglossic environment as an SLC, or a bilingual environment as a Dual Language Context, based on narrow linguistic criteria. Strong predictions about the effects of multiple language use on cognition should not be drawn based on these crude categories alone; instead, the everyday linguistic experiences of the user should be central to the formulation of predictions.

This study aimed to investigate the role of context solely as a factor in enhancing EFs. The results of the study, however, suggest that the relationship between language and cognition is a complex one, involving heterogenous populations, sociocultural contexts and individual experiences. One limitation of this study is the lack of consideration for other factors alongside context. For instance, many studies have highlighted the importance of controlling for factors such as the socio-economic status (Engel de Abreu et al. 2012; Morton and Harper 2007), age (Bialystok et al. 2005), education (Luo et al. 2013), age of second language acquisition (Kapa and Colombo 2013; Pelham and Abrams 2014) and cultural background (Samuel et al. 2018). In the case of our study, most of these factors (with the exception of educational level) were not systematically controlled, and they could potentially explain the overall difference in reaction times between our groups, but not necessarily the *absence* of effects that would suggest a benefit for diglossics. Nevertheless, future studies on diglossic populations should aim to account for such factors by using tools that carefully measure them.

6. Conclusions

Despite the wealth of studies into the effect of speaking two or more languages on executive functions, little research has investigated similar effects of speaking two language *varieties*. In this study we assessed diglossic and monolingual young adults on the executive function domains of inhibition and switching (Miyake et al. 2000), which have been shown to be affected in older participants from the same linguistic environment (Alrwaita et al. n.d.). Our study found no diglossic benefits in any on these tasks, suggesting that, even if single language contexts are able to confer some effects on aspects of domain general cognition, these may not be observable in younger individuals. We argue that examining the language context, in terms of the amount of code-switching employed, is essential to understanding the relationship between using multiple languages/language varieties and cognition. Further, we argue that no contextual assumptions should be ascribed to all bilingual or diglossic situations. Rather, careful attention should be paid to the specific code- switching requirements of each language environment.

Author Contributions: Conceptualization, C.P. and N.A.; methodology, C.P. and N.A.; validation, C.P., L.M. and N.A.; formal analysis, N.A. and L.M.; investigation, N.A.; resources, C.P.; data curation, N.A. and L.M.; writing—original draft preparation, N.A.; writing—review and editing, C.P. and C.H.-P.; visualization, L.M.; supervision, C.P. and C.H.-P.; project administration, N.A.; funding acquisition, N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded Saudi Government Scholarship awarded to N.A. The APC was funded by the University of Reading Open Access Fund.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Research Ethics Committee of the University of Reading.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are openly available in OSF at https://osf.io/z6g2p, accessed on 7 December 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Language and Social Background Questionnaire (Translated to Arabic)

استبيان عن الخلفية اللغوية والاجتماعية

العمر ______

- الجنس: ذكر أنثى
- الوظيفة / المرحلة الدراسية (مثال: موظف / غير موظف / المرحلة الدراسية) ______
- تستخدم اليد اليمنى اليسرى

تاريخ الولادة ______

هل تمارس ألعاب الفيديو الأكشن والتي تتضمن ألعاب إطلاق

النار من منظور الشخص الأول؟ نعم لا

كم ساعة في الأسبوع تمارس تقريبا هذه المجموعة من ألعاب الفيديو؟ ______ إذا كان الجواب

- هل تعاني من أي مشاكل سمعية؟ نعم لا
- إذا كان الجواب نعم، فهل تستخدم أي سماعات خاصة؟ نعم لا
- هل تعاني من أي مشاكل في النظر؟ نعم لا
- إذا كان الجواب نعم، هل تستخدم نظارات طبية أو عدسات؟ نعم لا
- هل تم تصحيح نظرك إلى الحد الطبيعي بالنظارات أو العدسات؟ نعم لا
- هل لديك عمى الألوان؟ نعم، فأي نوع؟ نعم لا
- إذا كان الجواب ______
- هل تعرضت إلى إصابة في الرأس؟ نعم، هل من الممكن أن نعم لا
- تشرح لنا ______ إذا كان الجواب
- هل تعاني من أي مشاكل عصبية؟ (مثال: الصرع) (الرجاء تسميتها ______ إذا نعم لا
- كان الجواب نعم)
- هل تتعاطى حاليا أي علاجات نفسية؟ نعم لا

ارجو تحديد أعلى مرحلة دراسية ووظيفية لكل من الاب والام:

األم:	األب:
ثانوية ___ثانوية___ شهادة شهادة ما بعد المرحلة الثانوية___ ___دراسات عليا ___ الوظيفة:___ الأولى:___ الثانية:___ اللغة أخرى:___	ثانوية ___ثانوية___ شهادة شهادة ما بعد المرحلة الثانوية___ ___دراسات عليا ___ الوظيفة:___ الأولى:___ الثانية:___ اللغة أخرى:___

- هل ولدت في السعودية؟   نعم   لا
- اذا كان الجواب لا اين ولدت؟ ___
- متى قدمت للسعودية؟ ___

- هل عشت مؤقتًا في أي دولة لا تستخدم اللغة العربية كلغة سائدة؟   نعم   لا

اذا كان الجواب نعم اين و وكم كانت المدة؟	١/___	من:	إلى:
	٢/___	من:	إلى:
	٣/___	من:	إلى:

الخلفية اللغوية

الرجاء ذكر جميع اللغات واللهجات التي تتحدثها وتفهمها مع ترتيبها حسب إتقانك إياها مع ذكر اللغة العربية●

اللغة:	أين تعلمتها :	متى التعلم (ضع تاريخ تعلمها تقريباً أو السن)	كم من الوقت تقضي في التحدث باللغة يومياً؟ الرجاء ذكر النسبة.
١/______	□ المدرسة □ المجتمع □ مكان آخر: ___		
٢/______	□ المدرسة □ المجتمع □ مكان آخر: ___		
٣/______	□ المدرسة □ المجتمع □ مكان آخر: ___		
٤/______	□ المدرسة □ المجتمع □ مكان آخر: ___		

تقييم لدرجة مهاراتك اللغوية من ٠ الى ١٠ في اللغة العربية واللغات الأخرى التي تتكلمها

- العربية الفصحى:

٠ غير متقن ٥ ١٠ متقن

في المواقف التالية، كم من الوقت تستخدم فيه اللغة العربية الفصحى: نادرا بعض الوقت

غالبا جميع الوقت لا إطلاقا

المحادثة	☐	☐	☐	☐	☐
الاستماع	☐	☐	☐	☐	☐
الكتابة	☐	☐	☐	☐	☐
القراءة	☐	☐	☐	☐	☐

- العربية العامية: فهم

غير متقن ٠ ... ٥ ... متقن ١

التحدث ●——|——|——|——●

الفهم/القراءة ●——|——|——|——●

الكتابة ●——|——|——|——●

●——|——|——|——●

-

في الوقت الحالي، كم من الوقت تستخدم اللهجة العامية: ... نادرا ... بعض الوقت

غالبا ... جميع الوقت تقريبا

التحدث	☐	☐	☐	☐	☐
الاستماع	☐	☐	☐	☐	☐
الكتابة	☐	☐	☐	☐	☐
القراءة	☐	☐	☐	☐	☐

- في اللغة الإنجليزية:

٠ غير متقن — ٥ — ١٠ متقن

التحدث

الفهم/القراءة

الكتابة

-

في الوقت الحالي، كم من الوقت تستخدم فيه اللغة الإنجليزية: نادرا — بعض الوقت — غالبا — جميع الوقت تقريبا

التحدث	☐	☐	☐	☐	☐
الاستماع	☐	☐	☐	☐	☐
الكتابة/القراءة	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐

اللغة المستخدمة في المجتمع

في أي لغة كنت تسمعها أو تستخدمها بشكل متكرر داخل أو خارج المنزل في المراحل العمرية التالية

	جميعه عربي	معظمه عربي	نصف عربي ونصف لغة أخرى	معظمها لغة أخرى	لغة أخرى فقط
رضيع					
الروضة					
المرحلة الابتدائية					
المرحلة الثانوية					

الرجاء تحديد اللغة المستخدمة بشكل عام عند التحدث مع الأشخاص التاليين :

	اللغة العربية	اللغة العربية نص حي	[illegible] اللغة الأخرى	[illegible] اللغة الأخرى	ومعظمها اللغة الأخرى	اللغة الأخرى فقط
الوالدين						
الأخوة						
الأجداد						
الأقارب الآخرون						
الزوج/الزوجة						
الزملاء						
الجيران						
الأصدقاء						

الرجاء تحديد أي اللغات تستخدم بشكل عام في الأوضاع التالية :

	العامية اللغة العربية	اللغة العربية فصحى	عامية أكثر من فصحى اللغة الأخرى	فصحى أكثر من اللغة الأخرى	معظمها اللغة الأخرى	اللغة الأخرى فقط
المنزل						
المدرسة						
العمل						
النشاطات الاجتماعية (الخروج مع الأصدقاء مثلا)						
النشاطات الدينية						
النشاطات الترفيهية (الرياضة، الهوايات، الألعاب)						
التسوق، المطاعم						
المرافق الصحية، المرافق الحكومية، المكاتب، البنوك.						

الرجاء تحديد أي اللغات تستخدم بشكل عام في النشاطات التالية :

	اللغة العربية العامية	اللغة العربية الفصحى	العامية والفصحى واللغة الأخرى	الفصحى واللغة الأخرى	معظمها اللغة الأخرى	اللغة الأخرى فقط
القراءة						
البريد الإلكتروني						
الرسائل النصية						
التواصل الاجتماعي (فيس بوك، تويتر، ..)						
الكتابة (الواجبات، المشتريات، الملاحظات)						
مشاهدة الأفلام						
تصفح الانترنت						
الصلاة						

في بعض الأوقات نقوم بالتحويل من اللغة الأم إلى أخرى، مثال نتحدث اللغة العربية العامية ثم نحول الكلام إلى بعض اللغة العربية الفصحى، تعرف هذه الظاهرة بـ "التحويل اللغوي".
الرجاء تحديد مقدار تحويلك بين اللغات في الحالات التالية :

التحويل بين اللغة العربية الفصحى والعربية العامية :

	أبدا	قليل	أحيانا	غالبا	دائما
مع الوالدين والعائلة					
مع الأصدقاء					
على التواصل الاجتماعي (فيس بوك، تويتر ...)					

التحويل بين اللغة العربية والإنجليزية :

	أبدا	قليل	أحيانا	غالبا	دائما
مع الوالدين والعائلة					
مع الأصدقاء					
على التواصل الاجتماعي (فيس بوك، تويتر ...)					

References

Albirini, Abdulkafi. 2016. *Modern Arabic Sociolinguistics: Diglossia, Variation, Codeswitching, Attitudes and Identity*. London and New York: Routledge, Taylor & Francis Group.

Alrwaita, Najla, Carmel Houston-Price, and Christos Pliatsikas. 2022. The effects of using two variants of one language on cognition: Evidence from bidialectalism and diglossia. *Linguistic Approaches to Bilingualism.*. [CrossRef]

Alrwaita, Najla, Lotte Meteyard, Toms Voits, Carmel Houston-Price, and Christos Pliatsikas. n.d. *Executive Functions are Modulated by the Context of Dual Language Use: Comparing Diglossic and Bilingual Older Adults*. Bilingualism: Language and Cognition.

Alsahafi, Morad. 2016. Diglossia: An overview of the Arabic situation. *International Journal of English Language and Linguistics Research* 4: 1–11.

Amara, Muhammad Hasan, and Abd Al-Rahman Mar'I. 2002. Policy and Teaching Arabic as a Mother Tongue. In *Language Education Policy: The Arab Minority in Israel*. Dordrecht: Springer, pp. 61–86. [CrossRef]

Anderson, John A. E. Anderson, Lorinda Mak, Aram Keyvani Chahi, and Ellen Bialystok. 2018. The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods* 50: 250–63. [CrossRef]

Antón, Eneko, Manuel Carreiras, and Jon Andoni Duñabeitia. 2019. The impact of bilingualism on executive functions and working memory in young adults. *PLoS ONE* 14: e0206770. [CrossRef]

Antoniou, Kyriakos, and George Spanoudis. 2020. An Investigation of the Multilingual and Bi-dialectal Advantage in Executive Control. *The Cognitive Science Society 2020* 8: 2050–56.

Antoniou, Kyriakos, Kleanthes Grohmann, Maria Kambanaros, and Napoleon Katsos. 2016. The effect of childhood bilectalism and multilingualism on executive control. *Cognition* 149: 18–30. [CrossRef]

Auer, Peter. 2011. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In *Perspectives on Variation*. Edited by Nicole Delbecque, Johan van der Auwera and Dirk Geeraerts. Berlin and New York: De Gruyter Mouton. [CrossRef]

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

Bialystok, Ellen. 2011. Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale* 65: 229–35. [CrossRef]

Bialystok, Ellen, Fergus I. M. Craik, and Gigi Luk. 2012. Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences* 16: 240–50. [CrossRef]

Bialystok, Ellen, Fergus I. M. Craik, and Jennifer Ryan. 2006. Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32: 1341–54. [CrossRef]

Bialystok, Ellen, Michelle M. Martin, and Mythili Viswanathan. 2005. Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism* 9: 103–19. [CrossRef]

Boudelaa, Sami, and William D. Marslen-Wilson. 2013. Morphological structure in the Arabic mental lexicon: Parallels between standard and dialectal Arabic. *Language and Cognitive Processes* 28: 1453–73. [CrossRef]

Calvo, Amelia, and Ellen Bialystok. 2014. Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition* 130: 278–88. [CrossRef]

Carlson, Stephanie M., and Andrew N. Meltzoff. 2008. Bilingual experience and executive functioning in young children. *Developmental Science* 11: 282–98. [CrossRef]

Chevalier, Nicolas, Tifanny D. Sheffield, Jennifer MizeNelson, Caron A. C. Clark, Sandra A. Wiebe, and Kimberly Andrews Espy. 2012. Underpinnings of the Costs of Flexibility in Preschool Children: The Roles of Inhibition and Working Memory. *Developmental Neuropsychology* 37: 99–118. [CrossRef]

Costa, Albert, Mireia Hernández, and Nuria Sebastián-Gallés. 2008. Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106: 59–86. [CrossRef]

Costa, Albert, Mireia Hernández, Jordi Costa-Faidella, and Nuria Sebastián-Gallés. 2009. On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113: 135–49. [CrossRef]

De Bruin, Angela, Barbara Treccani, and Sergio Della Sala. 2015. Cognitive Advantage in Bilingualism: An Example of Publication Bias? *Psychological Science* 26: 99–107. [CrossRef]

Donnelly, Seamus, Patricia J. Brooks, and Bruce D. Homer. 2019. Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review* 26: 1122–47. [CrossRef]

Dux, Paul E., Michael N. Tombu, Stephenie Harrison, Baxter P. Rogers, Frank Tong, and Rene Marois. 2009. Training Improves Multitasking Performance by Increasing the Speed of Information Processing in Human Prefrontal Cortex. *Neuron* 63: 127–38. [CrossRef]

Emmorey, Karen, Gigi Luk, Jennie. E. Pyers, and Ellen Bialystok. 2008. The Source of Enhanced Cognitive Control in Bilinguals. *Psychological Science* 19: 1201–6. [CrossRef]

Engel de Abreu, Pascale M. J., Anabela Cruz-Santos, Carlos J. Tourinho, Romain Martin, and Ellen Bialystok. 2012. Bilingualism Enriches the Poor: Enhanced Cognitive Control in Low-Income Minority Children. *Psychological Science* 23: 1364–71. [CrossRef]

Ferguson, Charles A. 1959. Diglossia. *WORD* 15: 325–40. [CrossRef]

Ferreira, Daniel, Rat Correia, Antonieta Nieto, Alejandra Machado, Yaiza Molina, and José Barroso1. 2015. Cognitive decline before the age of 50 can be detected with sensitive cognitive measures. *Psicothema* 27: 216–22. [CrossRef]

Frangoudaki, Anna. 1992. Diglossia and the present language situation in Greece: A sociological approach to the interpretation of diglossia and some hypotheses on today's linguistic reality. *Language in Society* 21: 365–81. [CrossRef]

Giles, Howard. 1973. Accent Mobility: A Model and Some Data. *Anthropological Linguistics* 15: 87–105.

Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The Adaptive Control Hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef]

Hernández, Mireia, Clara D. Martin, Francisco Barceló, and Albert Costa. 2013. Where is the bilingual advantage in task-switching? *Journal of Memory and Language* 69: 257–76. [CrossRef]

Hilchey, Mathew D., and Raymond M. Klein. 2011. Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review* 18: 625–58. [CrossRef]

Hughes, Claire, Rosie Ensor, Anji Wilson, and Andrew Graham. 2009. Tracking executive function Across the Transition to School: A Latent Variable Approach. *Developmental Neuropsychology* 35: 20–36. [CrossRef]

Kapa, Leah L., and John Colombo. 2013. Attentional control in early and later bilingual children. *Cognitive Development* 28: 233–46. [CrossRef]

Kaye, Alan S. 2001. Diglossia: The state of the art. *International Journal of the Sociology of Language* 2001: 117–29. [CrossRef]

Keller, Rudolf E. 1973. Diglossia in German-speaking Switzerland. *Bulletin of the John Rylands Library* 56: 130–49. [CrossRef]

Kilgarriff, Adam, Frieda Charalabopoulou, Maria Gavrilidou, Janne Bondi Johannessen, Saussan Khalil, Sofie Johansson Kokkinakis, Robert Lew, Serge Sharoff, Ravikiran Vadlapudi, and Elena Volodina. 2014. Corpus-based vocabulary lists for language learners for nine languages. *Language Resources and Evaluation* 48: 121–63. [CrossRef]

Kirk, Neil W., Linda Fiala, Kenneth C. Scott-Brown, and Vera Kempe. 2014. No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology* 26: 640–48. [CrossRef]

Kousaie, Shanna, and Natalie A. Phillips. 2012. Ageing and bilingualism: Absence of a "bilingual advantage" in Stroop interference in a nonimmigrant sample. *Quarterly Journal of Experimental Psychology* 65: 356–69. [CrossRef]

Kroll, Judith F., and Ellen Bialystok. 2013. Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology* 25: 497–514. [CrossRef]

Lenth, Russel V., Paul Buerkner, Maxime Herve, Jonathon Love, Hannes Riebl, and Henrik Singmann. 2021. emmeans: Estimated Marginal Means, aka Least-Squares Means (Version 1.5.4). Available online: https://CRAN.R-project.org/package=emmeans (accessed on 8 December 2022).

Luk, Gigi Justice. 2022. Examining the assumptions and framing of researching bilingual (dis)advantage. In *Applied Psycholinguistics*. Edited by Rachel Hayes-Harb. Cambridge: Cambridge University Press, pp. 1–15.

Luo, Lin, Fergus I. M. Craik, Sylvain Moreno, and Ellen Bialystok. 2013. Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging* 28: 28–34. [CrossRef]

Marian, Viorica, and Michael Spivey. 2003. Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* 6: 97–115. [CrossRef]

Masrai, Ahmed, and James Milton. 2019. How many words do you need to speak Arabic? An Arabic vocabulary size test. *The Language Learning Journal* 47: 519–36. [CrossRef]

Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter, and Tor D. Wager. 2000. The Unity and Diversity of executive functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology* 41: 49–100. [CrossRef]

Morton, J. Bruce, and Sarah N. Harper. 2007. What did Simon say? Revisiting the bilingual advantage. *Developmental Science* 10: 719–26. [CrossRef]

Paap, Kenneth R., and Zachary I. Greenberg. 2013. There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* 66: 232–58. [CrossRef]

Paap, Kenneth R., Hunter A. Johnson, and Oliver Sawi. 2015. Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69: 265–78. [CrossRef]

Palmer, Jeremy. 2008. Arabic Diglossia: Student Perception of spoken Arabic after living in the Arabic-speaking world. *Journal of Second Language Acquisition and Teaching* 15: 81–95.

Papapavlou, Andreas N., and Pavlos Pavlou. 1998. A Review of the Sociolinguistic Aspects of the Greek Cypriot Dialect. *Journal of Multilingual and Multicultural Development* 19: 212–20. [CrossRef]

Pavlou, Pavlos. 2004. Greek dialect use in the mass media in Cyprus. *International Journal of the Sociology of Language* 2004: 101–18. [CrossRef]

Pelham, Sabra D., and Lise Abrams. 2014. Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40: 313–25. [CrossRef]

Poarch, Gregory J., Jan Vanhove, and Raphael Berthele. 2019. The effect of bidialectalism on executive function. *International Journal of Bilingualism* 23: 612–28. [CrossRef]

Rosenhouse, Judith, and Mira Goral. 2006. Bilingualism in the Middle East and North Africa: A Focus on the Arabic-Speaking World. In *The Handbook of Bilingualism*. New York: John Wiley & Sons, Ltd., pp. 835–68. [CrossRef]

Ross, Josephine, and Alissa Melinger. 2017. Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Developmental Science* 20: e12405. [CrossRef]

Rowe, Charley, and Kleanthes K. Grohmann. 2013. Discrete bilectalism: Towards co-overt prestige and diglossic shift in Cyprus. *International Journal of the Sociology of Language* 2013: 119–42. [CrossRef]

S'hiri, Sonia. 2002. Speak Arabic Please!: Tunisian Arabic Speakers' Linguistic Accommodation to Middle Easterners. In *Language Contact and Language Conflict in Arabic*. Milton Park: Routledge.

Salminen, Tiina, Tilo Strobach, and Torsten Schubert. 2012. On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience* 6: 166. [CrossRef]

Samuel, Steven, Karen Roehr-Brackin, Hyensou Pak, and Hyunji Kim. 2018. Cultural Effects Rather Than a Bilingual Advantage in Cognition: A Review and an Empirical Study. *Cognitive Science* 42: 2313–41. [CrossRef]

Scaltritti, Michele, Francesca Peressotti, and Michelle Miozzo. 2017. Bilingual advantage and language switch: What's the linkage? *Bilingualism: Language and Cognition* 20: 80–97. [CrossRef]

Shapiro, Michael C., Bhadriraju Krishnamurti, Colin P. Masica, and Anjanji K Sinha. 1989. South Asian Languages: Structure, Convergence and Diglossia. *Journal of the American Oriental Society* 109: 475. [CrossRef]

Sharoff, Serge. 2006. Open-source Corpora: Using the net to fish for linguistic data. *International Journal of Corpus Linguistics* 11: 435–62. [CrossRef]

Soveri, Anna, Matti Laine, Heikki Hämäläinen, and Kenneth Hugdahl. 2011. Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm*. *Bilingualism: Language and Cognition* 14: 371–78. [CrossRef]

Surrain, Sarah, and Gigi Luk. 2019. Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–15. *Bilingualism: Language and Cognition* 22: 401–15. [CrossRef]

Treffers-Daller, Jeanine, Zehra Ongun, Julia Hofweber, and Michal Korenar. 2020. Explaining Individual Differences in Executive Functions Performance in Multilinguals: The Impact of Code-Switching and Alternating Between Multicultural Identity Styles. *Frontiers in Psychology* 11: 561088. [CrossRef]

Valian, Virginia. 2015. Bilingualism and cognition. *Bilingualism: Language and Cognition* 18: 3–24. [CrossRef]

Yang, Hwajin, Andree Hartanto, and Sujin Yang. 2016. The importance of bilingual experience in assessing bilingual advantages in executive functions. *Cortex* 75: 237. [CrossRef] [PubMed]

languages

Review

Bilingualism, Culture, and Executive Functions: Is There a Relationship?

Wenhan Xie [1], Jeanette Altarriba [2] and Bee Chin Ng [1,*]

[1] Department of Linguistics and Multilingual Studies, Nanyang Technical University, Singapore 639798, Singapore
[2] Department of Psychology, University at Albany, State University of New York, Albany, NY 12222, USA
* Correspondence: mbcng@ntu.edu.sg

Abstract: The relationship between executive functions (EF) and bilingualism has dominated debate in the field. This debate was characterised by optimism for a bilingual advantage until the last decade, when a steady stream of articles reported failure to find a consistently positive effect for bilingualism. In addition to addressing concerns about study quality, this turn of events has spurred research into other variables that may explain the conflicting findings. While recent studies have focused on sociodemographic variables and interactional contexts such as age, code-switching frequency, and socioeconomic class to account for various group and individual differences, the impact of culture is seldom scrutinised. This paper examines the possible effect of culture among bilingual studies on EF by first contextualising how bilingual EF are studied and outlining the absence of culture as a macro variable, followed by a discussion on how culture and language are often conflated. This paper directs attention to the small but emerging research that tracks the importance of culture as a separate variable from language. This review discusses why macro culture and individual monoculturalism or biculturalism need to be carefully elucidated as a factor that can interact with the bilingual experience in shaping EF.

Keywords: culture; bilingualism; executive function; biculturalism; bilingual advantage

Citation: Xie, Wenhan, Jeanette Altarriba, and Bee Chin Ng. 2022. Bilingualism, Culture, and Executive Functions: Is There a Relationship? *Languages* 7: 247. https://doi.org/10.3390/languages7040247

Academic Editor: John W. Schwieter

Received: 15 May 2022
Accepted: 7 September 2022
Published: 23 September 2022

1. Introduction

The last century has seen an acceleration in research interest on the impact of bilingualism on cognition. Depending on the period which we draw our sample of research from, we are likely to get a very different picture of this relationship. Multiple studies in the literature have charted the course of the withering (the 1920s), waxing (1960 onwards), and waning (post 2013) enthusiasm about the bilingual advantage (cf. Antoniou 2019; Lehtonen et al. 2018; Bialystok and Craik 2022 for a recent review on this topic). This review differs from earlier reviews examining the presence or absence of a bilingual advantage by focusing on the variable of culture in studies on bilingualism and executive functions (EF) in the last two decades. We evaluate the small but growing evidence of cultural effects in recent bilingualism studies and argue that there is compelling evidence for culture to be carefully scrutinised as a critical variable that relates to bilingual people's EF.

2. The Bilingual Advantage in EF Debate

Ever since the seminal study by Peal and Lambert (1962), several studies comparing monolinguals and bilinguals seem to have yielded favourable findings regarding bilingual effects, with reports of cognitive advantages for bilinguals across their lifespans. Among the elderly, scholars argue that bilingualism may contribute to a "cognitive reserve" that may delay neurodegenerative diseases such as Alzheimer's and dementia by approximately four years compared with monolinguals (Bak et al. 2014; Bialystok et al. 2007; Calvo et al. 2016). Some researchers have proposed that bilinguals might have an advantage over their monolingual peers due to constant monitoring, selection, and inhibition of languages. In particular, they argue that bilinguals have an edge even in domain-general, non-linguistic tasks that tap into these cognitive processes (Green 1998; Green and Abutalebi 2013; Bialystok 2017).

While the exact definition and components of EF are still the subject of much debate (see Miyake and Friedman 2012 for a discussion), the consensus is that EF are higher-order cognitive functions that can be categorised into three core groups: inhibition and interference control (e.g., selective attention and behavioural and cognitive inhibition), cognitive flexibility (i.e., mental flexibility in task and set shifting), and working memory (Diamond 2013; Miyake et al. 2000). EF are believed to be the foundation upon which essential skills such as planning (Collins and Koechlin 2012) and critical reasoning (Lunt et al. 2012) are built. Moreover, they are also found to predict theory of mind in developing children (Sabbagh et al. 2006) and have significant implications for an individual's personal, social, and academic development (Best et al. 2011).

Proponents of a bilingual advantage in domain-general EF argue that the advantage stems from a bilingual's need to direct attention in monitoring speech input and selecting the appropriate language to respond while inhibiting the production of words and phrases in the other language(s). Some studies have suggested that all languages in a bilingual's repertoire are jointly activated (Costa et al. 1999; Marian and Spivey 2003; Thierry and Wu 2007) during listening (Spivey and Marian 1999), reading (Dijkstra and Kroll 2005), and speech (Kroll et al. 2006). The act of "juggling" multiple activated schemas is said to hone domain-general interference and inhibition control (Kroll et al. 2015).

In the recent literature, we can track two dominant views explaining the roots of the bilingual advantage. The first model, put forth by Green (1998), has sometimes been termed the "Inhibition Control Model" proposes that bilinguals' experience in inhibiting the co-activated non-relevant language is present even in monolingual speech contexts. This experience enhances their ability to inhibit distracting or irrelevant stimuli better than monolinguals. Some key studies supporting this view include those of Bialystok et al. (2004) and Kroll et al. (2008). Expanding on this model, Green and Abutalebi's (2013) adaptive control hypothesis argues that different control processes, including goal maintenance, conflict monitoring, interference suppression, cue detection, response inhibition, task (dis)engagement, and opportunistic planning, are activated depending on the demands of different interactional contexts (single language, dual language, and dense code-switching). Emerging evidence suggests depending on the level of cognitive control, these different contexts could result in varying enhancements of cognitive control (Beatty-Martínez et al. 2020; Ooi et al. 2018). Other researchers have suggested that code-switching between languages intersententially and intrasententially might share similarities with domain-general task switching (Prior and MacWhinney 2009; Hartanto and Yang 2016). Supporting evidence for this hypothesis comes from brain imaging studies which showed an overlap in neural connections and architecture for language switching and domain-general non-linguistic switching tasks (de Baene et al. 2015; Weissberger et al. 2019). Proponents of this model theorise that these "cognitive exercises" augment domain-general EF and confer a bilingual advantage that is evident in children as young as 7 months old (Kovács and Mehler 2009).

The second model takes the view that inhibition and control alone are insufficient to explain the enhancement of bilinguals' cognitive abilities. Bialystok et al. (2010; see also Bialystok 2015, 2017) discussed the possibility that the bilingual advantage in EF may stem from how bilinguals monitor, control, and direct attention and that bilingual exposure may alter how individuals manage these resources (cf. Bialystok et al. 2012; Colzato et al. 2008; Costa et al. 2009). This concept, put forth as attentional control, is thought of as being broader than any single component in Miyake's (Miyake et al. 2000) three-component model of EF (inhibition and interference control, shifting, and working memory). It involves the ability to monitor, suppress, or ignore irrelevant stimuli and direct cognitive resources to either maintain or switch depending on the relevant information (Bialystok 2017). Indeed, Bialystok and Craik (2022) suggested that bilinguals' ability to direct attention has a far-reaching prowess which is enough to "enhance processes of both facilitation and inhibition, as well as processes underlying cognitive flexibility and resource allocation".

In the ensuing years since the early 2000s, numerous studies across the field have reported on how bilinguals excel in tasks requiring participants to pay attention to relevant information while suppressing distracting changes to rules or irrelevant cues. These advantages reportedly extend to adults, as evidenced by their greater accuracy and faster resolving of incongruent stimuli in the Simon (Antoniou et al. 2016; Bialystok et al. 2004; Tse and Altarriba 2014), Flanker (Costa et al. 2009), and Stroop tasks (Nayak et al. 2020; Poulin-Dubois et al. 2011). They were also observed among the elderly (above 60 years of age), who showed lower processing costs, suggesting greater efficiency even after considering the typical age-related decrease in performance. This led some scholars to suggest that lifelong bilingualism may mitigate specific age-related declines in cognitive performance when concerning inhibition and control (Bialystok 2021).

Bialystok (2011) argues that bilinguals' ability to switch between languages or language varieties fluidly may enhance their cognitive flexibility, especially in switching and shifting tasks. A seminal study by Prior and MacWhinney (2009), which showed how proficient bilingual college students outperformed monolinguals in reaction time during switch trials but not in non-switch trials, points to a bilingual advantage in cognitive shifting. Similar studies, such as that by Wiseheart et al. (2014), have also reported how bilinguals were better able to resolve ambiguity in stimulus–response associations, resulting in lower global switching costs. Bilinguals are also thought to hone cognitive control as a response to their sociolinguistic environment, with varying levels and types of code switching (e.g., intersentential vs. intrasentential), resulting in more efficient cognitive processing faculties shown, for example, in reduced switching costs (Yang et al. 2016).

Up till 2010, the evidence for a bilingual advantage in the literature is robust. However, in more recent years, the bilingual superiority effect in EF has not been unanimously reported. In the last decade, many studies have either failed to replicate findings reporting a bilingual advantage or did not come to that conclusion when comparing bilingual and monolingual populations.

3. Conflicting Findings in EF Performance

3.1. Methodological Concerns

In recent years, challenges to the firm conclusion of a bilingual advantage have been highlighted from various perspectives and by different research groups (Hernández et al. 2013; Paap and Greenberg 2013; Paap and Sawi 2014; Paap et al. 2015; Antón et al. 2014; de Bruin et al. 2015; Von Bastian et al. 2016). A vital issue of concern that has been widely discussed is publication bias, where journal publications favour the publishing of significant effects in support of a bilingual advantage (de Bruin et al. 2015). Lehtonen et al.'s (2018) meta-analytic review is, by far, the most comprehensive attempt to synthesise existing studies relating to bilingualism and executive functions in adults. The review involved 152 studies covering a range of 6 executive function domains (inhibitory control, monitoring, shifting, working memory, attention, and verbal fluency). The authors included unpublished doctoral and masters' theses in addition to journal articles while accounting for the various effect sizes of each paper to mitigate the effects of publication bias. They reported that no apparent advantage in any of the six executive function domains provided evidence that bilingual adults were at an advantage compared with their monolingual peers after correcting for publication bias.

Several studies have also been unable to replicate the findings of the original bilingual advantage found in seminal studies for the Stroop (Paap et al. 2018, 2019), Simon (Gathercole et al. 2014), and Flanker tasks (Paap et al. 2019) and switching ability (Goriot et al. 2018), even after matching monolingual and bilingual participants to a number of linguistic and sociodemographic variables. Paap's (2019) comparison of mean reaction time differences in interference-control tasks across 177 studies for "benchmark" tests of interference-control, such as the Simon, Stroop, Flanker and Attention Network Tasks (ANT) also found that more than 80% of studies returned null results. Indeed, such studies raise important questions regarding both study quality and replicability. Furthermore, a recent meta-analysis by Lowe et al. (2021) synthesised data from a profusion of published studies and unpublished data sets to examine the effect of language status (bilingual vs. monolingual) on EF among children. After comparing more than a thousand effect sizes in studies with monolingual and bilingual participants between 3 and 17 years, the researchers detected a small overall effect of bilingual language status on children's EF that was completely attenuated when corrected for publication bias.

Another extensive experimental study with more than 4500 participants by Dick et al. (2019) also failed to find evidence demonstrating better performance among bilinguals in the Flanker, Dimension Change Card Sort, and Stop-Signal Reaction Time tests, which measure common EF indicators including inhibition, attention, and set-shifting abilities among 9–10-year-old children in the United States. Bilingual participants (n = 1740) were grouped based on three different definitions of bilingualism: bilingual status (if children identified as speaking any language other than English), bilingual degree (ratio of non-English language use to English use), and bilingual use (continuous measure of non-English language frequency with different interlocutors). After controlling for demographic covariates such as socioeconomic status, age, education level, and intelligence, the regression analyses revealed a bilingual disadvantage in terms of poorer English vocabulary, consistent with other studies in the literature (e.g., Martin-Rhee and Bialystok 2008). While the researchers reported a bilingual disadvantage in the Stop-Signal Reaction Time test, no other significant differences were found between the monolingual and bilingual participants. Thus, despite the report of a bilingual advantage in the early 2000s, much of the current discussion is hotly contesting the validity and replicability of such claims.

3.2. Hidden Confounds

Other than the methodological issues raised, some researchers argue that differences in experimental designs and confounds in participant selection have led to discrepancies in findings. Variables such as ethnicity, social background, and social-economic status (Bak 2016; Blom et al. 2017; Morton and Harper 2007), different definitions of early or late bilingualism (Yang et al. 2016), a participant's cultural upbringing (Tran et al. 2015), or immigrant status (Kousaie and Phillips 2012) have not always been well controlled or considered in many studies, showing a bilingual advantage in EF. Interestingly, much of the discussion is very similar to the discussion raised in Peal and Lambert's (1962) watershed report on the effect of bilingualism on cognition. Researchers such as de Bruin (2019) have called for more nuanced studies that consider the complex social variables that influence bilinguals as part of different groups and communities. In the following section, we examine some possible reasons for the contradictory findings and put forth a case for why culture as a variable should be considered in earnest by researchers in the field.

The lack of a bilingual advantage in the abovementioned studies signals the need for closer scrutiny of the variables that may contribute to bilinguals' seemingly variable performance. As Luk and Bialystok (2013) pointed out, the cognitive consequences of bilingualism need to be considered within the broader, multi-dimensional bilingual experience. This view is further echoed by Valian (2014) and de Bruin (2019), who argue that diverse differences between bilinguals' environmental contexts, social interaction habits, and various sociolinguistic variables may influence the findings for EF.

First, the age group of the participants may play a role in detecting a significant bilingual advantage. Some articles suggest that the bilingual advantage seems to be most pronounced among young children and the elderly (Bak et al. 2014; Luo et al. 2010), while the differences between monolingual and bilingual young adults may sometimes show insignificant or no group differences, possibly because young adults are at the zenith of their cognitive functioning (Bialystok et al. 2004). Hence, bilingual effects may not be as pronounced in young adults compared with the elderly (for a counter perspective, see Samuel et al. 2018), as ageing has been found to take a toll on cognitive and executive function processing, and differences in performance among older bilinguals could be attributed to the cognitive reserve built up from lifelong bilingualism (Abutalebi et al. 2015; Antoniou and Wright 2017; Nickels et al. 2019).

Another possible source of variation involves the definition and classification of bilinguals based on their age of acquisition and fluency. Researchers in the field do not all agree on how "early" or "late" bilingualism may be unambiguously classified due to disagreement on the exact age range of the critical period for language acquisition (Byers-Heinlein and Lew-Williams 2013). Consequently, these different yardsticks can result in different conclusions. For example, Martin et al. (2013) suggested that early bilinguals (AoA: between 2 and 4 years) had the edge over late bilinguals. In contrast, Pelham and Abrams (2014) reported that while bilingualism had a facilitative effect on conflict resolution in incongruent trials over the monolinguals in their study, no differences in executive control were found between early (AoA: 7 years or younger) and late bilinguals (AoA: age 13 and above). When adopting the age of acquisition as a continuous variable in a mixed model regression, Von Bastian et al. (2016) could not detect a bilingual advantage in EF between participants for a broad range of tasks measuring inhibitory control, shifting, conflict monitoring, and general cognitive abilities. Indeed, such methodological differences can lead to dissonant findings and be an important source of variation in whether a difference in EF tests is found.

Additionally, concerns have also been raised regarding self-rated language proficiency when it is used to compare groups from different cultural or geographical backgrounds. Tomoschuk et al. (2018) found that self-rated proficiencies and picture-naming test scores were not consistently correlated among different bilingual populations, with Chinese-English bilinguals performing better in the picture-naming task compared with Spanish-English bilinguals who had identical self-rated proficiency scores. Notably, even grouping participants by language dominance (other language vs. English) did not remove these discrepancies. These findings indicate potential confounds, as the subjective, self-rated scores in one population may be conceptualised differently than another group of participants due to different social environments and upbringings, ages of acquisition, and language dominance, even if they are taken from the same community (e.g., college students).

In fact, there is evidence that it is not only language proficiency but how balanced the bilingual is in terms of proficiency and use in both languages that can impact performance in tests of EF. A seminal study by Tse and Altarriba (2012) involved administering a Stroop task on 110 bilingual adults to determine the effects of language proficiency on their attentional control system. The researchers reported that responses to the Stroop task were positively correlated to second-language proficiency and that bilinguals who were more balanced in both language proficiencies were more efficient at conflict resolution in incongruent trials. In particular, they argued that a bilingual's language proficiencies honed aspects of attentional control (conflict resolution and goal maintenance) that explained the different performance in the Stroop task. Similarly, Yow and Li (2015) found that bilinguals who were both proficient and balanced in their use of both languages outperformed those who were less balanced in terms of use and proficiency in tasks of inhibition and set-shifting.

These studies, we think, provide compelling arguments to consider the issue of balance when assessing bilingual proficiency. While most studies offer some measure of proficiency, problems with validity and comparability can still arise as a result of methodological differences that can potentially mask the effects of a bilingual advantage in EF.

4. Culture: An Often-Overlooked Factor

Another factor that has seldom been discussed is the effect of culture as a potential confounding variable. Culture is undoubtedly a complex construct that can be defined at many levels. It must encompass the depth and breadth of interactions, behaviours, emotions, mindsets, and ways of being, both tangible and intangible. For the purpose of this article, culture will be taken to refer to the learned and shared system of beliefs, values, preferences, and social norms that are spread by shared activities (Altarriba and Basnight-Brown 2022; Arshad and Chung 2022; Bezin and Moizeau 2017).

Over the years, various operationalisations have been proposed to allow researchers to distinguish between cultures. One highly influential model frequently used to delineate the differences between cultures in cross-cultural psychology is Hofstede's cultural dimensions (Hofstede 1980, 2001, 2011). These cultural dimensions—power distance, individualism vs. collectivism, uncertainty avoidance, masculinity vs. femineity, long-term orientation vs. short-term orientation, and indulgence vs. restraint—allow researchers to make broad-brush generalisations of the cultural traits typical to the residents of different countries and geographical regions (see Varnum et al. 2010 for a review). While this model uses nationality as a proxy for culture and is subject to the ecological fallacy of overgeneralisation on the individual level, scholars have used it extensively over the years to generalise traits for comparisons between participant samples. We felt it necessary to raise this issue at the outset, as nearly all instances of cultural comparisons in the literature used nationality as a proxy, and the majority made reference to Eastern (East Asian) compared with Western culture (Table 1).

Table 1. Some generalised traits of "Eastern" compared with "Western" cultures.

	"Eastern"	**"Western"**
Power distance	Large	Small
Uncertainty avoidance	High	Low
Individualism-collectivism	Collectivism	Individualism
Orientation	Long term	Short term
Indugence-restraint	Restrained	Indulgence

Zooming in on research at the intersection of bilingualism, cultural differences, and EF, only a dozen or so studies in the past decade have explicitly investigated the impact of culture on bilingualism and EF (see Appendix A for an overview of the key studies). Some of the earliest studies of this nature were a response to Morton and Harper's (2007) argument for the possibility of one's ethnic (and cultural) background accounting for the bilingual advantage previously found in early childhood bilinguals. In the years that followed, several studies attempted to control for the effect of culture on EF by intentionally selecting samples with diverse language and cultural backgrounds based on nationality (e.g., Bialystok and Viswanathan 2009). Others sought to isolate bilingual effects by maintaining cultural homogeneity through sampling bilinguals and monolinguals with similar cultural backgrounds (Yang and Yang 2016). In the following section, we review select articles in the existing literature that have included cultural comparisons in studies of language status (mono- or bilingualism) and EF.

4.1. Studies in Young Children

As indicated earlier, over the years, multiple studies have emerged where bilinguals were found to have a global advantage in tasks of executive control (Bialystok 1999; Bialystok and Viswanathan 2009; Bialystok et al. 2010; Barac and Bialystok 2012; Tran et al. 2015, 2018; Yang et al. 2011; Yang and Yang 2016). Nevertheless, an effect of culture was found despite it not overriding the effect of bilingualism in many of these articles (see Appendix A for a breakdown of the bilingual or cultural effects reported among key studies). For example, while Yang et al.'s study in 2011 reported an overall bilingual advantage in terms of accuracy and reaction time, as well as in conflict resolution for the Attention Network Task (ANT), they also found that the overall accuracy of Korean monolinguals from Korea was higher than that of Korean or English monolinguals from the USA. Interestingly, both Korean and English monolinguals from the USA performed similarly. This suggests that an effect of culture, unrelated to the languages spoken, could be evident even in 4-year-old children.

Similarly, Tran et al. (2015) also found a similar effect in East Asian children who outperformed Western and Latin American children in terms of reaction time and accuracy, despite the bilingual advantage over monolinguals still prevailing overall. Another longitudinal study of Vietnamese, Argentinian, and American children by Tran et al. (2018) uncovered a global, bilingual cognitive advantage in their longitudinal study of 96 3-year-old children as well as a cultural effect where Eastern (Vietnamese) children outperformed Western (USA) and Latin American (Argentinian) children on the day/night task, which measures verbal response inhibition.

Some studies also suggest that cultural effects may modulate advantages in EF over bilingualism. Recent research seeking to disentangle the effects of language and culture among preschoolers matched participants on a measure of "country of origin" as a proxy for culture. Cho et al. (2021) tested Korean monolinguals (Korea), Korean-English bilinguals (Canada), and English monolinguals (Canada) on a modified colour and word Stroop task as a measure of inhibition control. The researchers found that while Korean-English bilinguals outperformed English monolingual children in terms of accuracy in incongruent trials, Korean-English bilinguals performed no differently than Korean monolinguals after controlling for age and SES. Critically, they found that the country of origin was the key modulating variable predicting accuracy in incongruent trials after controlling for demographic variables and performance in congruent trials.

Taken together, it seems plausible that cultural upbringing could play an important role in early EF development. In particular, cultural differences seem to be most evident in executive control tasks, possibly interacting with bilingual experience in shaping EF. One conjecture that could explain the inconsistencies reported could stem from differences in the cultural expectations of children regarding obedience and following directions. These expectations could be ingrained in children in different cultural contexts at a much earlier age. For example, children in East Asian cultures are typically expected to follow the rules more closely and practice response inhibition from a younger age than children from America (Kelkar et al. 2013; Lan et al. 2011). These could provide possible explanations for why the country of origin (as a loose proxy for "culture") could help to explain the variability reported in the literature. Nevertheless, with the current dearth in this vein of research, the exact nature of the interaction between culture and language with EF among young children remains an open question. Further research is needed to elucidate what conditions might reveal cultural effects (i.e., beyond "East vs. West" distinctions, integrating other measures of culture such as long-term orientation vs. short-term orientation instead of only individualism vs. collectivism) that could have knock-on effects on EF components such as attention, especially since most existing studies have only examined the effect of culture and language on inhibition control.

4.2. Studies in Adults

Contrary to studies conducted with children, examining the effect of cultural differences in the EF of bilingual adults seems to paint a different but clearer picture. In a study disambiguating the effect of cultural background and language, Samuel et al. (2018) tested inhibition control using a Simon task on 211 adult participants from three cultural backgrounds: British, Korean, and mixed nationalities (drawn from 33 countries). Bilingualism was taken as a continuous measure of three factors: L2 proficiency, language dominance, and code-switching frequency. Analysis using linear mixed-effects regression revealed that Koreans outperformed the British group in every measure (RT, accuracy, and smaller Simon effect) and in every model while performing faster overall than the mixed group in two out of three models. The mixed nationality group also outpaced the British participants in nearly all measures across every model. This provides critical evidence that even macro-level cultural effects can possibly account for the different levels of performance on common tasks of inhibition and control, especially among adult participants.

In one of the few studies that explicitly separated multicultural identity and its effect on language and executive function, Treffers-Daller et al. (2020) found a bilingual advantage in a reduced Flanker conflict effect when comparing bilinguals and monolinguals. Notably, they reported that multicultural identity styles (Ward et al. 2018) were the key explanatory variable in explaining EF variance among bilingual subjects at an individual level in their model. Similarly, Xie and Ng (in preparation) also found a significant effect in resolving conflict in a Flanker task among high-proficiency bicultural bilinguals who differ in their frequency of cultural switching in their daily lives.

Preliminary evidence also suggests that activating different cultural frames could be associated with different performance in inhibition control tasks. Ye et al. (2017) reported a bilingual advantage in incongruent trials for a Flanker task in mixed cultural contexts. High-proficiency Mandarin-English bilinguals outperformed participants with high proficiency in Mandarin but low English proficiency when filler slides showed both Western (British and American) and Eastern (Chinese) cultural icons. Interestingly, the bilingual advantage was not replicated in single cultural contexts (e.g., fillers with only Eastern cultural icons) or in congruent trials. The authors speculated that the tasks may not be challenging enough to elucidate an advantage in conflict resolution. Further analyses also showed that bicultural contexts attenuate proficient bilinguals' cognitive performance significantly when examining the results of both the mixed cultural and single cultural conditions for proficient bilinguals. Indeed, findings such as these beg the question as to whether existing inconsistencies reported in the literature about bilinguals' (dis)advantage in EF could be explained by differences in individual participants' cultural milieus, or if it is their cultural switching habits that have shrouded the "true" performance in tasks relating to EF.

In summary, while many of these studies on young children yielded mixed results, when the critical studies were examined (Appendix A), only four reported a complete absence of cultural effects. Indeed, Bialystok and Craik (2022) acknowledged that culture could be a possible confound to the bilingual advantage in tests on EF, although they highlighted several studies that have shown language effects overriding cultural effects. Yet, there seems to be a growing body of research suggesting that the effect of culture may be more pronounced when examining adult populations. We hypothesise that cultural effects might be more evident in adulthood due to the prolonged experience of honing their cultural selves and may also result from their exposure to multiple cultures and the acquisition of bicultural or multicultural identities. Indeed, this growing body of research suggests that culture as a variable of interest should be given a second look (e.g., Altarriba 2008).

5. Is There a Need to Disambiguate Cultural Effects from Language Effects in Studies on Executive Functions?

It has been standard practice that individuals recruited for bilingualism studies are matched by the languages they profess to know. In fact, it is not uncommon to include individuals from different language backgrounds, nationalities, races, and ethnicities so long as they broadly speak the same language, broadly because most studies typically report large language "families" instead of specific language varieties or dialects (e.g., English speakers may refer to speakers of any English dialect (American English, British English, Singapore English, Indian English, etc.)). However, this does not capture a sufficiently nuanced perspective of the cultural diversity among bilingual populations.

In recent years, scholars such as Grosjean (2015) have argued that bilingualism and biculturalism are often conflated in research on bilingualism. This is not a new idea, as Soffietti (1960) made a case for the reality that culture and language status are distinct, where bilinguals could be monocultural or bicultural and monolinguals could be monocultural or bicultural. Indeed, the reality is somewhat complex as just as there is a whole spectrum of bilinguals from dominant bilinguals to balanced bilinguals, simultaneous bilinguals, receptive bilinguals, etc., an entire range of bi- or even monocultural bilinguals exists. Grosjean (1992) argues that some Europeans may be multilingual, having studied two or more languages in a school setting. However, they are monocultural as they only work and stay in a country and a single cultural setting. On the other hand, individuals can be monolingual but multicultural by immersing themselves in different cultures across their social contacts but choosing not to learn their languages (Schwartz et al. 2017; Shih and Sanchez 2005).

While Grosjean (2015) limited his discussion of biculturalism and bilingualism to identity and personality, we believe it is critical to consider the implications of culture and biculturalism on cognition. Up until now, research has suggested that cross-cultural differences influence the development of EF even among preschool children as early as the age of three, where cultural variation in the development of EF is seen in various tasks of inhibition and cognitive flexibility (Legare et al. 2018; Ling et al. 2018; Norenzayan et al. 2002; Imada et al. 2012). To fully appreciate the distinction between bicultural and monocultural bilinguals, we first need to account for culture as a variable associated with cognition and particularly EF. We will first unpack the relevance of cross-cultural differences in shaping EF and then situate these differences in bicultural individuals.

5.1. Macro-Scale Cultural Differences and EF

An area of cross-cultural comparison that has received much attention over the last two decades pertains to the so-called "East-West difference". Individuals from East Asian countries who are of East Asian descent or have been exposed to East Asian culture and values from a young age are commonly considered to embody Eastern cultural values. They are viewed as being more collectivist, having higher power distances in relationships with authority, being more likely to avoid uncertainty, and having a stronger sense of interdependence (Markus and Kitayama 1991, 2010; Oyserman and Lee 2008) compared with Americans who epitomise Western culture in their desire for individual autonomy, egalitarian relationships, and willingness to take risks (Triandis 1989, 2001; Harkness et al. 2000). The differences in Eastern and Western cultures have been well documented and shown to influence parents' thoughts on how to raise their children (Chen et al. 1998; Bornstein 2013), socialisation and personality (Varnum et al. 2010), self-evaluation (Kim et al. 2009), individual's ethics and moral values (Garcia et al. 2014; Jia and Krettenauer 2017), as well as self-regulation (Imada et al. 2012; Jaramillo et al. 2017; Krassner et al. 2016) and emotion (Kitayama et al. 2006; Masuda et al. 2008; Pavlenko 2005). Clearly, these distinctions may characterise many individuals. However, it is important to maintain that there is a broad range of these variables throughout any given group or culture; that is, they always reside as a distribution within and among these populations.

These very cultural differences have been theorised to influence cognition. Cultural value systems such as individualism vs. collectivism, tolerance towards uncertainty, sociolinguistic factors about when formal schooling should begin, as well as culture-specific parenting attitudes related to child autonomy, expected discipline, level of parental control, and emotional socialisation are critical determinants in shaping EF (Sarma and Thomas 2020).

Over the years, various studies have sought to distinguish the effects of culture on EF. In a seminal study, Masuda and Nisbett (2001) reported differences in memory recall for East Asians compared with Americans due to different cultural upbringings. The researchers reported that East Asian adults could remember more background information in the visual stimuli and became less accurate when background cues were changed compared with their American counterparts. The authors hypothesised that these differences could result from different modes of attentional direction and focus emphasised by Eastern and Western cultures. Another study investigating how different cultural backgrounds would influence cognition by Sabbagh et al. (2006) examined the interaction between EF and theory of mind in 5-year-old children from America and China. The researchers found that the Chinese children outperformed the American children in all EF tests, such as the Dimensional Change Card Sort and Stroop tasks. In the same vein, Oh and Lewis (2008) found that Korean preschoolers performed significantly better at inhibition and task switching than British children of a similar age, while no differences were found in the measures of working memory. More recently, Imada et al. (2012) presented further evidence relating differences in EF to cultural differences. They compared 175 children from America and Japan and found that sensitivity to contextual cues was highly correlated with performance in EF tasks. Specifically, the children from Japan outperformed American children in the set-shifting Dimension Change Card Sort Test but were also more likely to peek on impulse during a delayed gratification task, indicating poorer impulse control.

One influential theory that has emerged at the forefront to explain the cultural effect on EF proposes that differences between cultures could stem from the disparity in cultural values and upbringing (including family and societal education strategies) that influence what individuals from contrasting cultures focus on (Nisbett and Masuda 2003, 2006). According to this theory, the subdivision between Eastern and Western cultures in cross-cultural studies has been correlated with different types of attentional focus and cognitive processing. Individuals brought up with Eastern, collectivistic cultural influences tend to have attentional processing tuned to focus on contextual cues (e.g., non-verbal cues, more holistic, and distributed attention), while Western cultures direct attention to process-specific, individualised information (Nisbett et al. 2001; Nisbett and Masuda 2003). In a series of four studies examining how culture mediates statistical learning, Kiyokawa et al. (2012) found differences in local and global perceptual biases in their British and Japanese participants, respectively, even after manipulating participants' familiarity with the sequence elements (strings of large (global) letters made out of small (local) letters). Their results provide evidence that cultural differences can influence the type of unconscious knowledge being learned (see also Kiyokawa et al. 2010 and Ling et al. 2018 for additional studies). Thus, scholars have hypothesised that with the emphasis on greater "general sensitivity", Eastern cultures may have an advantage in recognising and reacting to stimuli that are displayed when reaction time is measured (Nisbett and Miyamoto 2005; Miyamoto et al. 2006; Kuwabara and Smith 2012).

Consequently, the evidence above suggests that even broad-brush cultural differences in nationality or ethnicity are associated with different performances in EF tasks in cross-cultural comparisons. Extending this to our discussion on bilingualism and EF, this indicates that comparisons of participants based on language status (bilingual or monolingual) may not be directly comparable around the world, even if factors such as socioeconomic status and other sociolinguistic variables are controlled for. It is perhaps even more important to distinguish between society-level culture and the cultural identities and values an individual adopts. This will be discussed in the following section on micro-level cultural factors.

5.2. Micro-Level Cultural Factors: Individual Acculturation and Bilingualism

Individual biculturalism describes an individual's exposure and internalisation of two or more cultures as part of their identity (Nguyen and Benet-Martínez 2007, 2012). According to LaFromboise et al. (1993), multicultural individuals typically possess a certain degree of multicultural competence, be it in terms of explicit cultural knowledge, a tacit understanding of cultural scripts (culturally appropriate etiquette) and values, or friends and family within the other culture, and are often able to communicate in languages associated with the different cultures. In this paper, we define bicultural individuals more narrowly, as individuals who possess more than one cultural identity in their repertoire and can shift between cultural mindsets (frames) in their repertoire, choosing the appropriate actions, values, and norms when they interact with individuals and groups from other cultures (Benet-Martínez et al. 2002; Hong et al. 2000).

Early studies on biculturalism often focused on migrants' and immigrants' acculturation processes (Berry 1997). These studies have generated significant interest among cultural psychologists as they delineate how acculturation to a new culture can occur. One of the most influential models, Berry's Acculturation Model (Berry 1997, 2003), investigates how individuals acculturate to the host society. Specifically, the model proposes that individuals acculturate through two main processes: cultural maintenance and contact and participation. Individuals typically adopt one of four distinct acculturation strategies that result in either cultural assimilation, integration, separation, or marginalisation (Table 2). Of relevance to our discussion is how this model delineates various acculturation outcomes and shows a distribution of outcomes ranging from monocultural to bicultural (e.g., Schwartz et al. 2017; Meca et al. 2017). This has an important bearing on a bilingual's profile, as those who are more culturally integrated (i.e., more balanced) are likely to consider themselves more bicultural compared with bilinguals who do not identify as integrating (separating, assimilating, or marginalising), thus considering themselves more monocultural.

Table 2. Four possible outcomes from Berry's Acculturation Model (Berry 1997, 2003).

	+ ← Maintenance of heritage culture → −	
− ← Cultural Adaptation → +	Integration	Assimilation
	Separation	Marginalisation

Research by social psychologists has drawn a theoretical link between how acculturating to different cultures can hone an individual's cognitive abilities. The acculturation complexity model by Tadmor and Tetlock (2006) proposes that the level of immersion in new cultures (indexed by the willingness to acculturate) can hone an individual's cognitive skills differently. Specifically, cognitive abilities such as selective attention and inhibition can be developed, resulting from the pressures of resolving diverse cultural complexities and the necessity of being able to behave "appropriately" among interactants of different cultures. These differences in an individual's cultural preferences are thought to influence such cognitive gains, as individuals who prefer to use only one of the two cultures more frequently are less likely to experience and have fewer opportunities to resolve conflicts arising from cultural differences compared with someone with an equal preference for both cultures.

Over the years, research has suggested that bicultural people who integrate both cultures are able to provide more complex descriptions of each culture compared with monocultural people or bicultural people with a distinct preference for one culture (Benet-Martínez et al. 2006) and that a bicultural person's acculturation strategies impact cognition through instances of conflict mitigation and behavioural inhibition (Crisp and Turner 2011). Recently, Spiegler and Leyendecker (2017) showed how Turkish-German immigrant children with a balanced view of both cultures outperformed their peers, who favoured one culture over the other in terms of cognitive flexibility. Indeed, converging evidence suggests that individuals who identify as bicultural frequently integrate different sets of cultural knowledge, or as Cheng et al. (2014) described it, "bicultural individuals possess 'two cultural minds'—two sets of cultural knowledge, use two cultural schemas to guide their thoughts and behaviour, and can activate these two cultural frames of references" (p. 279).

Research building on this hypothesis posits a theory of cultural frame switching relevant to bilingualism studies and EF. Cultural frame switching is a term coined from the observation that individuals with multiple cultural identities are able to switch across cultural frameworks (or their various cultural minds) depending on the cultural cues being presented. In seminal studies by Hong et al. (2000) examining this phenomenon, the researchers found that bicultural individuals behaved differently depending on the cultural primes used. For instance, when bicultural individuals (Chinese Americans) were primed with icons representing the Chinese or American cultures (e.g., a dragon as an icon of Chinese culture or the American flag as an icon of American culture), the participants' appraisal of an ambiguous situation tended to embody specific cultural values of the culture associated with the prime. In these studies, the participants were asked if they thought an animated video of a fish swimming in front of a school of fish was being chased by the other fish (an external push factor) or if it was leading other fish (an internal factor). When primed with Chinese cultural primes, bicultural participants were comparatively more confident that the fish was being chased (an external attribution). When primed with American primes, they were more confident that the fish was leading the school (internal attribution).

The researchers argued that bicultural individuals could tap into different systems and schemas of cultural meaning and switch between them depending on the environment and context. Building on this, further research supported bicultural individuals' ability to switch unconsciously and seamlessly (Benet-Martínez et al. 2002) and even change identities based on cultural cues (Luna et al. 2008). Cultural frame switching often occurs when there is a significant difference in the interactant's environment, such as when moving between the public and private spheres. For instance, Suárez-Orozco et al. (2008) brought attention to the situation of immigrants in Boston and San Francisco in their longitudinal study of 470 immigrant children from countries such as Mexico, Central America, and China. Using parental interviews, test scores, and case studies, the authors exemplify the different ways in which children cope with the disparity between their heritage culture and the broader American culture and how intentional switching between cultural frames may be observed when moving between these different domains.

Extending the theory of cultural frame switching to bilingualism, Ramírez-Esparza et al. (2006) reported that Spanish-English bilinguals presented different personality traits depending on the language in which they answered self-reporting personality questionnaires. In particular, they showed increased extraversion, agreeableness, and conscientiousness when answering in English compared with Spanish. Another study by Chen and Bond (2010) provided further evidence for bilingual personality switching in examining how the interviewer's ethnicity can influence the way interlocutors present themselves. Bilingual interviewers (two Caucasian and two Chinese) interviewed 76 Chinese-English bilinguals, and observers were asked to rate participants' personalities by the traits of extraversion and openness. Each participant was interviewed by a Chinese and a Caucasian interviewer, both of whom used English and Chinese separately. The researchers found that the participants displayed increased extraversion and were more willing to speak about experiences with Caucasian interviewers compared with Chinese interviewers, regardless of the language mode used for the interviews. This highlights how bilingual and bicultural people have the resources to switch between cultural frames implicitly, often without conscious effort, depending on the appraisal of the social speech context and cultural environment.

In summary, these studies on bilingualism and culture support the idea that bicultural individuals have different cultural systems that are selected and inhibited fluidly and automatically, depending on the situational context. In addition, bicultural individuals may need to direct cognitive resources to monitor the situational context and choose the appropriate cultural values, attitudes, and ideologies which are relevant while inhibiting inappropriate behaviour when switching between cultures. These processes seem to mirror how a bilingual person's language system is hypothesised to function in terms of language inhibition and control and in terms of directing attention for monitoring and language selection. If this is indeed the case, the cultural frame switching that bicultural bilingual people participate in could influence the development of executive functions in a similar fashion to how bilingualism hones the executive functions system.

This is a crucial issue to resolve as to date, bilingual research has frequently conflated bilingualism and biculturalism (Grosjean 2015). Moreover, the reality is somewhat complex. Just as there is a whole spectrum of bilinguals from dominant bilinguals to balanced bilinguals, simultaneous bilinguals, receptive bilinguals, etc., an entire spectrum of bi- or even monocultural bilinguals exists. According to Grosjean (1992), some Europeans may be multilingual, having studied two or more languages in a school setting. However, many are predominantly monocultural, as they only work and stay in one country and a single cultural setting. On the other hand, individuals can be monolingual but multicultural by immersing themselves in other cultures in their social contacts but choosing not to learn their languages (Padilla 2006). This nuanced view that forms a bilingual-bicultural separation is an aspect that was absent in nearly all of the literature we reviewed.

6. Future Directions

Evidence within the literature suggests that a comparison of cultural effects, be it generalisations based on nationality or a more individual scale (individual mono- or biculturalism), is associated with differences in EF. Among scholars interested in the intersection of languages and EF, it is essential that we consider how cultural variables can be a mediating factor when describing bilinguals' EF.

Here, we would like to propose a few directions that the field could take to examine the effect of culture on bilinguals' EF in greater detail. Most existing studies incorporating an aspect of culture within experiments looking at bilinguals' EF have typically used a general "East vs. West" distinction for contrast. Future studies should move beyond the assumption that an individual's culture is based on their nationality or that they belong to a particular culture simply due to being born and raised in his or her country of origin. With most existing studies matching participants based on geography or where they currently reside as an earmark for culture, we may be making assumptions about macro-level culture based on citizenship, nationality, and ethnicity, all of which may not necessarily reflect the individuals' cultural affiliations or identities.

Similarly, research distinguishing the effects of bilingualism on different populations should elucidate bilinguals' cultural allegiances, examine if they are bicultural, and test if they can switch between various cultural frames. Most studies examining the cognitive effects of bilingualism do not mention their participants' cultural or bicultural affiliations. Thus far, studies have conflated bilingualism and biculturalism, with biculturalism subsumed within the construct of bilingualism or not considered a variable. This situation poses a problem when looking at the effects of language status and EF, as they may be separate constructs. Critically, researchers in the field need to be aware of and differentiate the cultural statuses of their participants, namely whether they are monocultural or bicultural and whether they have frequent practice in switching between cultural identities. In particular, bicultural switching effects may explain the discrepancies in the results among certain participant samples who adopted significantly different cultural behaviours at home and at work or school. One such example is that of participants who are second- or third-generation immigrants. While previous studies have addressed the importance of distinguishing between immigrant and non-immigrant populations (Mezzacappa 2004; Fuller-Thomson and Kuh 2014), most bilingual studies do not typically consider the "carryover" cultural effects for second- or even third-generation immigrants. Although these second-generation immigrants may be citizens of the host country, those whose parents (first-generation immigrants) come from a very different culture may find it difficult to reconcile the differences between the cultures within and outside the home. In fact, social and cultural psychologists have long been studying the different acculturation struggles of second-generation immigrants and the influence on their multicultural identity.

For example, Stroink and Lalonde (2009) reported on how some second-generation Asian-Canadian immigrants found it difficult to integrate the differences between their Eastern family culture with the general Western culture in the larger society, resulting in a conflicting bicultural identity. Lee and Kim's (2014) qualitative study on second-generation Korean immigrants in Germany detailed different coping mechanisms and shifting strategies for bridging Eastern and Western cultures in their daily lives such that individuals could "blend into both Korean and German societies, similar to a chameleon" (p. 97). As such, we hypothesise that if switching between cultural frames occurs both commonly and frequently (e.g., in day-to-day life and interactions), bicultural switching could hone domain-general inhibition control and task-switching. The impact of bicultural switching might thus show similar EF gains reported among code-switching bilingual speakers, such as in the work of Hartanto and Yang (2016). Consequently, the level of bicultural switching (or non-switching) could be distinguished among samples of bilingual and monolingual participants to examine if bicultural switching might interact with the bilingual experience in shaping EF.

As bicultural individuals can select, inhibit, and switch from one cultural mindset (frame) to another depending on their interactional context, we hypothesise that similar cognitive processes may be involved, and its effects on executive control may be analogous to code-switching in bilinguals. According to the adaptive control hypothesis (Green and Abutalebi 2013), it is proposed that "language control processes themselves adapt to the recurrent demands placed on them by the interactional context" (p. 515). In particular, the researchers hypothesised that different interactional contexts (single language, dual language, and code-switching) will impose varying demands on the cognitive control system. For instance, using two languages separately, such as in school and at home (single-language context), is believed to require inhibition of a bilingual person's other language. However, using two languages in the same context with different speakers (dual-language contexts) requires the most stringent cognitive control processes, such as monitoring, interference, and response inhibition. Finally, frequent code-switching in the same interactional context is hypothesised to require less cognitive suppression and control, as it allows for "opportunistic planning" to freely use lexical items from either language.

Similarly, we hypothesise that the process of selecting, juggling, and switching between multiple cultural mindsets may engage—and thus enhance—cognitive control mechanisms, including inhibitory control, monitoring, and shifting (Spiegler and Leyendecker 2017). While some aspects of EF have been compared in the existing literature on bilingualism, the few studies that have discussed the effects of culture on language and EF have mainly examined response inhibition control (e.g., the Simon and Stroop tasks) and attention-related tasks (e.g., the Attention Network Task). This leaves room for other components, including memory and cognitive flexibility (switching), to be further explored.

7. Conclusions

Although a plethora of evidence contradicting a bilingual advantage in EF exists, some papers still suggest that bilinguals might have the edge over monolinguals in specific contexts. We believe the disparity motivates questions about why there is so little consensus surrounding this complex set of topics, as well as establishing that it is unlikely that a single variable can fully explain the wide variability in findings. Hence, possible confounding variables such as an individual's cultural affiliation and level of multiculturalism should be carefully considered. While culture's exact implications and complex effects on EF in bilingual populations are still being examined, growing evidence suggests that cultural variables may mediate bilingual individuals' cognitive abilities. We believe that greater emphasis needs to be placed on understanding the effects of broad, macro-level cultures, the influences of bicultural switching, and the broader impacts of biculturalism in future studies of a bilingual person's cognitive advantage. Only by first considering the individual differences and sociolinguistic factors and broadening the current boundaries of what constitutes bilingual people's interaction context will we be able to arrive at a better understanding of the unique contribution of bilingualism to EF.

In this paper, we argue that cultural variation and biculturalism deserve greater attention from scholars interested in the cognitive effects of bilingualism. Here, we suggest that a more nuanced view of culture needs to be considered, as cultural effects are seldom accounted for or explicitly manipulated in studies of bilingualism. As recent studies comparing important demographic and sociolinguistic factors are making headway in unravelling the puzzle that is bilingual people's cognition, we hope that this paper takes a small step towards a clearer picture of the exact role that language plays in relation to EF.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of key publications explicitly discussing culture in studies on bilingual EF. Asterisk (*) indicates a note about the participant sample.

Publication	EF Task(s)	Type of Participants (Country or Nationality, M_{age})		Presence of Bilingual Advantage	Cultural Effects
		Bilinguals	Monolinguals		
(Sabbagh et al. 2006)	Stroop (day/night and grass/snow), bear/dragon, Dimension Change Cart Sort Task (DCCS), tower building, whisper, and Kansas Reflection-Impulsivity Scale for Preschoolers (KRISP)	-	109 Chinese (China, 4 y.o.) 107 English (USA, 4 y.o.)	-	Among the monolingual children tested, Eastern preschoolers (China) outperformed Western preschoolers (USA) in all measures of executive functioning
(Morton and Harper 2007)	Simon Task	17 English-French (Canada, 6.8 y.o.)	17 English (Canada, 6.8 y.o.)	No BL advantage or disadvantage in either congruent or incongruent trials	Authors suggested that matching by ethnicity (cultural background) and socioeconomic variables mitigated bilingual advantage
(Bialystok and Viswanathan 2009)	Anti-saccade task (faces task)	30 English-combination (Canada, 8.5 y.o) 30 English-Tamil or Telugu (India, 8.6 y.o.)	30 English (Canada, 8.5 y.o.)	BL advantage in conditions based on inhibitory control and cognitive flexibility, but there was no significant difference between groups in response suppression	No group differences between bilingual groups (Canadians vs. Indians)
(Bialystok et al. 2010)	Luria's tapping task, opposite worlds, reverse categorisation, ANT (Flanker), and mutual exclusivity	27 English-combination (Canada, 3.5 y.o.)	40 English (Canada, 3 y.o) 20 French (France, 3.5 y.o.)	BL advantage in all executive control (Conflict) tasks	No significant cultural effects reported
		29 English-combination (Canada, 4.5 y.o.)	29 English (Canada, 4.5 y.o) 17 French (France, 4.5 y.o.)		
(Yang et al. 2011)	Attention Network Task (ANT)	15 Korean-English (USA, 4 y.o.)	15 English (USA, 4 y.o.) 13 Korean (USA, 4 y.o.) 13 Korean (Korea, 4 y.o.)	Bilingual advantage over monolingual groups in accuracy and RT, as well as in conflict resolution (Executive control)	Overall accuracy of Korean monolinguals from Korea higher than Korean or English monolinguals in the USA, though RT was slower. (Korean MLs in the USA performed similarly to English MLs in the USA.)
Barac and Bialystok (2012)	Colour-shape task switching	30 Chinese-English (5.9 y.o.) 28 French-English (6.2 y.o.) 20 Spanish-English (6.2 y.o.)	26 English (5.9 y.o.)	Bilingual advantage with smaller global task-switching costs	No significant cultural effects reported
		* Note: Participant location not specified. Only "large multicultural city" was mentioned (p. 416).			

Table A1. *Cont.*

Publication	EF Task(s)	Type of Participants (Country or Nationality, M_{age})		Presence of Bilingual Advantage	Cultural Effects
		Bilinguals	Monolinguals		
(Tran et al. 2015)	ANT	13 Spanish-English (USA) 15 Vietnamese-English (USA) 16 Vietnamese-Cantonese (Vietnam)	14 English (USA) 19 Spanish (Argentina) 20 Vietnamese (Vietnam)	BL advantage in accuracy and RT over monolinguals	Eastern children have faster RT and greater accuracy than Western or Latin American children. Significant culture and ANT network interaction. Main effect of culture on task performance. Cultural background plays a vital role in the development of the alerting and executive control networks.
		* Longitudinal study. Children were initially M_{age} 38.8 months old (3 y.o.) and tested at 5 time points 6 months apart.			
Yang and Yang (2016)	ANT	32 Korean-English BL (USA, second-generation immigrants, 5–6 y.o.)	31 English (USA, 5–6 y.o.)	BL advantage in global performance, accuracy, and RT ANT: No BL effects on network efficiency scores	No specific cultural effects seen. Cultural differences controlled by studying culturally homogenous children and adults.
		20 Korean-English (Korean undergraduates who arrived at 10 y.o., currently 19.9 y.o)	19 English (USA, 20.7 y.o.)	BL advantage in global performance + reaction time ANT: BL advantage in + orienting + executive control	
(Ye et al. 2017)	Flanker Task	18 Mandarin-English (China, 21 y.o.,)	18 Mandarin (China, 21 y.o.) * Not a completely monolingual: English exposure present, but without passing college English tests	Some evidence that high-proficiency bilinguals outperformed low-proficiency bilinguals	More demanding mixed cultural context cues bring out advantage in incongruent trials for high-proficiency bilingual participants.
(Samuel et al. 2018)	Simon Task	78 British (21 y.o.) 64 Korean (23 y.o.) 69 mixed nationalities (23 y.o.) * Level of bilingualism taken as continuous variables based on L2 proficiency, dominance, and code-switching frequency		No BL advantage	East Asian (Korean) participants outperformed Western (British) participants on RT and accuracy regardless of monolingual or bilingual status.

Table A1. *Cont.*

Publication	EF Task(s)	Type of Participants (Country or Nationality, M_{age})		Presence of Bilingual Advantage	Cultural Effects
		Bilinguals	Monolinguals		
(Tran et al. 2018)	DCCS, day/night Stroop, bear dragon, and gift delay	13 Spanish-English (USA) 15 Vietnamese-English (USA) 20 Vietnamese-Cantonese (Vietnam)	13 English (USA) 19 Spanish (Argentina) 20 Vietnamese (Vietnam)	BL advantage in DCCS, day/night, and gift delay task and advantage in inhibition and shifting	Eastern children (Vietnamese) outperformed Western and Argentinian children in the day/night task. Cultural effect in response inhibition.
		* Longitudinal study. Children were initially M_{age} 38.7 months old and tested every 6 months.			
(Treffers-Daller et al. 2020)	Flanker Task	29 Turkish-English (Turkey, 32.5 y.o.) 28 Turkish-English (Cyprus, 25.25 y.o.) * All BL participants were immigrants	30 English (UK, 32.3 y.o.)	BL advantage in inhibition (reduced conflict effect)	Among BLs, multicultural switching style (alternating or hybrid) was the key explanatory variable for variance in EF performance.
(Cho et al. 2021)	Colour and Word Stroop Task	33 Korean-English (Canada, 4.7 y.o.)	36 English (Canada, 4.4 y.o) 43 Korean (Korea, 4.3 y.o)	No BL advantage (no disadvantage)	BL East-Asians show higher accuracy on inhibitory control than ML Canadian children.

References

Abutalebi, Jubin, Lucia Guidi, Virginia Borsa, Matteo Canini, Pasquale A. Della Rosa, Ben A. Parris, and Brendan S. Weekes. 2015. Bilingualism provides a neural reserve for aging populations. *Neuropsychologia* 69: 201–10. [CrossRef] [PubMed]

Altarriba, Jeanette. 2008. Expressions of emotion as mediated by context. *Bilingualism: Language and Cognition* 11: 165–67. [CrossRef]

Altarriba, Jeanette, and Dana Basnight-Brown. 2022. The psychology of communication: The interplay between language and culture through time. *Journal of Cross-Cultural Psychology* 53: 860–74. [CrossRef]

Antón, Eneko, Jon A. Duñabeitia, Adelina Estévez, Juan A. Hernández, Alejandro Castillo, Luis J. Fuentes, Douglas J. Davidson, and Manuel Carreiras. 2014. Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology* 5: 398. [CrossRef]

Antoniou, Mark. 2019. The advantages of bilingualism debate. *Annual Review of Linguistics* 5: 395–415. [CrossRef]

Antoniou, Mark, and Sarah M. Wright. 2017. Uncovering the mechanisms Responsible for why language learning may promote healthy cognitive aging. *Frontiers in Psychology* 8: 2217. [CrossRef] [PubMed]

Antoniou, Kyriakos, Kleanthes K. Grohmann, Maria Kambanaros, and Napoleon Katsos. 2016. The effect of childhood bilectalism and multilingualism on executive control. *Cognition* 149: 18–30. [CrossRef]

Arshad, Memoona, and Joanne M. Chung. 2022. Practical recommendations for considering culture, race, and ethnicity in personality psychology. *Social and Personality Psychology Compass* 16: e12656. [CrossRef]

Bak, Thomas H. 2016. The impact of bilingualism on cognitive ageing and dementia. *Linguistic Approaches to Bilingualism* 6: 205–26. [CrossRef]

Bak, Thomas H., Jack J. Nissan, Michael M. Allerhand, and Ian J. Deary. 2014. Does bilingualism influence cognitive aging? *Annals of Neurology* 75: 959–63. [CrossRef]

Barac, Raluca, and Ellen Bialystok. 2012. Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development* 83: 413–22. [CrossRef] [PubMed]

Beatty-Martínez, Anne L., Christian A. Navarro-Torres, Paola E. Dussias, María Teresa Bajo, Rosa E. Guzzardo Tamargo, and Judith F. Kroll. 2020. Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 1022–47. [CrossRef]

Benet-Martínez, Verónica, Fiona Lee, and Janxin Leu. 2006. Biculturalism and cognitive complexity: Expertise in cultural representations. *Journal of Cross-Cultural Psychology* 37: 386–407. [CrossRef]

Benet-Martínez, Verónica, Janxin Leu, Fiona Lee, and Michael W. Morris. 2002. Negotiating biculturalism. *Journal of Cross-Cultural Psychology* 33: 492–516. [CrossRef]

Berry, John W. 1997. Immigration, acculturation, and adaptation. *Applied Psychology* 46: 5–34. [CrossRef]

Berry, John W. 2003. Conceptual approaches to acculturation. In *Acculturation: Advances in Theory, Measurement, and Applied Research*. Edited by Kevin Chun, Pamela Balls Organista and Gerardo Marín. Washington, DC: American Psychological Association, pp. 17–37.

Best, John R., Patricia H. Miller, and Jack A. Naglieri. 2011. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences* 21: 327–36. [CrossRef]

Bezin, Emeline, and Fabien Moizeau. 2017. Cultural dynamics, social mobility, and urban segregation. *Journal of Urban Economics* 99: 173–87. [CrossRef]

Bialystok, Ellen. 1999. Cognitive complexity and attentional control in the bilingual mind. *Child Development* 70: 636–44. [CrossRef]

Bialystok, Ellen. 2011. Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale* 65: 229–35. [CrossRef]

Bialystok, Ellen. 2015. Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives* 9: 117–21. [CrossRef]

Bialystok, Ellen. 2017. The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin* 143: 233–62. [CrossRef] [PubMed]

Bialystok, Ellen. 2021. Bilingualism: pathway to cognitive reserve. *Trends in Cognitive Sciences* 25: 355–64. [CrossRef] [PubMed]

Bialystok, Ellen, and Fergus I. M. Craik. 2022. How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin & Review* 29: 1246–69. [CrossRef]

Bialystok, Ellen, and Mythili Viswanathan. 2009. Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112: 494–500. [CrossRef] [PubMed]

Bialystok, Ellen, Fergus I. M. Craik, and Gigi Luk. 2012. Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences* 16: 240–50. [CrossRef] [PubMed]

Bialystok, Ellen, Fergus I. M. Craik, and Morris Freedman. 2007. Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45: 459–64. [CrossRef]

Bialystok, Ellen, Fergus I. M. Craik, Raymond Klein, and Mythili Viswanathan. 2004. Bilingualism, aging, and cognitive control: Evidence from the simon task. *Psychology and Aging* 19: 290–303. [CrossRef]

Bialystok, Ellen, Raluca Barac, Agnes Blaye, and Diane Poulin-Dubois. 2010. Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development* 11: 485–508. [CrossRef]

Blom, Elma, Tessel Boerma, Evelyn Bosma, Leonie Cornips, and Emma Everaert. 2017. Cognitive advantages of bilingual children in different sociolinguistic contexts. *Frontiers in Psychology* 8: 552. [CrossRef]

Bornstein, Marc H. 2013. Parenting and child mental health: A cross-cultural perspective. *World Psychiatry* 12: 258–65. [CrossRef]
Byers-Heinlein, Krista, and Casey Lew-Williams. 2013. Bilingualism in the early years: What the science says. *Learning Landscapes* 7: 95–112. [CrossRef] [PubMed]
Calvo, Noelia, Adolfo M. García, Laura Manoiloff, and Agustín Ibáñez. 2016. Bilingualism and cognitive reserve: A critical overview and a plea for methodological innovations. *Frontiers in Aging Neuroscience* 7: 249. [CrossRef] [PubMed]
Chen, Sylvia Xiaohua, and Michael Harris Bond. 2010. Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. *Personality and Social Psychology Bulletin* 36: 1514–28. [CrossRef] [PubMed]
Chen, Xinyin, Paul D. Hastings, Kenneth H. Rubin, Huichang Chen, Guozhen Cen, and Shannon L. Stewart. 1998. Child-rearing attitudes and behavioral inhibition in Chinese and Canadian toddlers: A cross-cultural study. *Developmental Psychology* 34: 677–86. [CrossRef]
Cheng, Chi-Ying, Fiona Lee, Verónica Benet-Martínez, and Que-Lam Huynh. 2014. Variations in multicultural experience: Influence of bicultural identity integration on socio-cognitive processes and outcomes. In *The Oxford Handbook of Multicultural Identity*. Edited by Veronica Benet-Martínez and Ying-Yi Hong. New York: Oxford University Press, pp. 276–99.
Cho, Isu, Jewan Park, Hyun-joo Song, and J. Bruce Morton. 2021. Disentangling language status and country-of-origin explanations of the bilingual advantage in preschoolers. *Journal of Experimental Child Psychology* 212: 105235. [CrossRef]
Collins, Anne, and Etienne Koechlin. 2012. Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology* 10: e1001293. [CrossRef]
Colzato, Lorenza S., Maria Teresa Bajo, Wery Van den Wildenberg, Daniela Paolieri, Sander Nieuwenhuis, Wido La Heij, and Bernhard Hommel. 2008. How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34: 302–12. [CrossRef]
Costa, Albert, Michele Miozzo, and Alfonso Caramazza. 1999. Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language* 41: 365–97. [CrossRef]
Costa, Albert, Mireia Hernández, Jordi Costa-Faidella, and Núria Sebastián-Gallés. 2009. On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113: 135–49. [CrossRef]
Crisp, Richard J., and Rhiannon N. Turner. 2011. Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin* 137: 242–66. [CrossRef]
de Baene, Wouter, Wouter Duyck, Marcel Brass, and Manuel Carreiras. 2015. Brain circuit for cognitive control Is shared by task and language switching. *Journal of Cognitive Neuroscience* 27: 1752–65. [CrossRef] [PubMed]
de Bruin, Angela. 2019. Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences* 9: 33. [CrossRef] [PubMed]
de Bruin, Angela, Barbara Treccani, and Sergio Della Sala. 2015. Cognitive advantage in bilingualism. *Psychological Science* 26: 99–107. [CrossRef]
Diamond, Adele. 2013. Executive Functions. *Annual Review of Psychology* 64: 135–68. [CrossRef] [PubMed]
Dick, Anthony Steven, Nelcida L. Garcia, Shannon M. Pruden, Wesley K. Thompson, Samuel W. Hawes, Matthew T. Sutherland, Michael C. Riedel, Angela R. Laird, and Raul Gonzalez. 2019. No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour* 3: 692–701. [CrossRef] [PubMed]
Dijkstra, Ton, and Judith F. Kroll. 2005. Bilingual visual word recognition and lexical access. In *Handbook of Bilingualism: Psycholinguistic Approaches*. Edited by Judith F. Kroll and Annette. M. B. de Groot. Oxford: Oxford University Press, pp. 179–201.
Fuller-Thomson, Esme, and Diana Kuh. 2014. The healthy migrant effect may confound the link between bilingualism and delayed onset of Alzheimer's disease. *Cortex* 52: 128–30. [CrossRef] [PubMed]
Garcia, Fredi, Diana Mendez, Chris Ellis, and Casey Gautney. 2014. Cross-cultural, values and ethics differences and similarities between the US and Asian countries. *Journal of Technology Management in China* 9: 303–22. [CrossRef]
Gathercole, Virginia C. Mueller, Enlli M. Thomas, Ivan Kennedy, Cynog Prys, Nia Young, Nestor Viñas Guasch, Emily J. Roberts, Emma K. Hughes, and Leah Jones. 2014. Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology* 5: 11. [CrossRef]
Goriot, Claire, Mirjam Broersma, James M. McQueen, Sharon Unsworth, and Roeland Van Hout. 2018. Language balance and switching ability in children acquiring English as a second language. *Journal of Experimental Child Psychology* 173: 168–86. [CrossRef]
Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1: 67–81. [CrossRef]
Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef] [PubMed]
Grosjean, François. 1992. Another view of bilingualism. *Cognitive Processing in Bilinguals* 83: 51–62. [CrossRef]
Grosjean, François. 2015. Bicultural bilinguals. *International Journal of Bilingualism* 19: 572–86. [CrossRef]
Harkness, Sara, Charles M. Super, and Nathalie Van Tijen. 2000. Individualism and the "Western mind" reconsidered: American and Dutch parents' ethnotheories of the child. In *Variability in the Social Construction of the Child*. Edited by Sara Harkness, Catherine Raeff and Charles M. Super. San Francisco: Jossey-Bass, pp. 23–39.
Hartanto, Andree, and Hwajin Yang. 2016. Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition* 150: 10–19. [CrossRef]
Hernández, Mireia, Clara D. Martin, Francisco Barceló, and Albert Costa. 2013. Where is the bilingual advantage in task-switching? *Journal of Memory and Language* 69: 257–76. [CrossRef]

Hofstede, Geert. 1980. *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills and London: Sage Publications.
Hofstede, Geert. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Thousand Oaks: Sage.
Hofstede, Geert. 2011. Dimensionalizing cultures: The hofstede model in context. *Readings in Psychology and Culture* 2. [CrossRef]
Hong, Ying-yi, Michael W. Morris, Chi-yue Chiu, and Veronica Benet-Martinez. 2000. Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist* 55: 709–20. [CrossRef]
Imada, Toshie, Stephanie M. Carlson, and Shoji Itakura. 2012. East-West cultural differences in context-sensitivity are evident in early childhood. *Developmental Science* 16: 198–208. [CrossRef]
Jaramillo, Jorge M., María I. Rendón, Lorena Muñoz, Mirjam Weis, and Gisela Trommsdorff. 2017. Children's self-Regulation in cultural contexts: The role of parental socialization theories, goals, and practices. *Frontiers in Psychology* 8: 923. [CrossRef]
Jia, Fanli, and Tobias Krettenauer. 2017. Recognizing moral identity as a cultural construct. *Frontiers in Psychology* 8: 412. [CrossRef]
Kelkar, Amit S., Monica S. Hough, and Xiangming Fang. 2013. Do we think alike? A cross-cultural study of executive functioning. *Culture and Brain* 1: 118–37. [CrossRef]
Kim, Young-Hoon, Chi-Yue Chiu, Siqing Peng, Huajian Cai, and William Tov. 2009. Explaining east-west differences in the likelihood of making favorable self-evaluations: The role of evaluation apprehension and directness of expression. *Journal of Cross-Cultural Psychology* 41: 62–75. [CrossRef]
Kitayama, Shinobu, Batja Mesquita, and Mayumi Karasawa. 2006. Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology* 91: 890–903. [CrossRef] [PubMed]
Kiyokawa, Sachiko, Zoltán Dienes, Daisuke Tanaka, and Ayumi Yamada. 2010. Cross cultural differences in implicit learning. Paper presented at the Annual Meeting of the Cognitive Science Society, Portland, OR, USA, August 11–14. Available online: https://escholarship.org/uc/item/0c98t8p9 (accessed on 1 September 2022).
Kiyokawa, Sachiko, Zoltán Dienes, Daisuke Tanaka, Ayumi Yamada, and Louise Crowe. 2012. Cross cultural differences in unconscious knowledge. *Cognition* 124: 16–24. [CrossRef] [PubMed]
Kousaie, Shanna, and Natalie A. Phillips. 2012. Ageing and bilingualism: Absence of a "bilingual advantage" in Stroop interference in a nonimmigrant sample. *Quarterly Journal of Experimental Psychology* 65: 356–69. [CrossRef] [PubMed]
Kovács, Ágnes Melinda, and Jacques Mehler. 2009. Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences* 106: 6556–60. [CrossRef] [PubMed]
Krassner, Ariye M., Maria A. Gartstein, Curie Park, Wojciech Ł. Dragan, Felipe Lecannelier, and Samuel P. Putnam. 2016. East–west, collectivist-individualist: A cross-cultural examination of temperament in toddlers from Chile, Poland, South Korea, and the U.S. *European Journal of Developmental Psychology* 14: 449–64. [CrossRef]
Kroll, Judith F., Susan C. Bobb, and Zofia Wodniecka. 2006. Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition* 9: 119–35. [CrossRef]
Kroll, Judith F., Susan C. Bobb, Maya Misra, and Taomei Guo. 2008. Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica* 128: 416–30. [CrossRef]
Kroll, Judith F., Paola E. Dussias, Kinsey Bice, and Lauren Perrotti. 2015. Bilingualism, mind, and brain. *Annual Review of Linguistics* 1: 377–94. [CrossRef]
Kuwabara, Megumi, and Linda B. Smith. 2012. Cross-cultural differences in cognitive development: Attention to relations and objects. *Journal of Experimental Child Psychology* 113: 20–35. [CrossRef]
LaFromboise, Teresa, Hardin LK Coleman, and Jennifer Gerton. 1993. Psychological impact of biculturalism: Evidence and theory. *Psychological Bulletin* 114: 395–412. [CrossRef] [PubMed]
Lan, Xuezhao, Cristine H. Legare, Claire Cameron Ponitz, Su Li, and Frederick J. Morrison. 2011. Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology* 108: 677–92. [CrossRef] [PubMed]
Lee, Hyo-Seon, and Seok-Ki Kim. 2014. A qualitative study on the bicultural experience of second-generation Korean immigrants in Germany. *Pacific Science Review* 16: 97–109. [CrossRef]
Legare, Cristine H., Michael T. Dale, Sarah Y. Kim, and Gedeon O. Deák. 2018. Cultural variation in cognitive flexibility reveals diversity in the development of executive functions. *Scientific Reports* 8: 16326. [CrossRef] [PubMed]
Lehtonen, Minna, Anna Soveri, Aini Laine, Janica Järvenpää, Angela De Bruin, and Jan Antfolk. 2018. Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin* 144: 394–425. [CrossRef]
Ling, Xiaoli, Li Zheng, Xiuyan Guo, Shouxin Li, Shiyu Song, Lining Sun, and Zoltan Dienes. 2018. Cross-cultural differences in implicit learning of chunks versus symmetries. *Royal Society Open Science* 5: 180469. [CrossRef]
Lowe, Cassandra J., Isu Cho, Samantha F. Goldsmith, and J. Bruce Morton. 2021. The bilingual advantage in children's executive functioning is not related to language status: A meta-analytic review. *Psychological Science* 32: 1115–46. [CrossRef]
Luk, Gigi, and Ellen Bialystok. 2013. Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology* 25: 605–21. [CrossRef]
Luna, David, Torsten Ringberg, and Laura A. Peracchio. 2008. One individual, two identities: Frame switching among biculturals. *Journal of Consumer Research* 35: 279–93. [CrossRef]

Lunt, Laura, Jessica Bramham, Robin G. Morris, Peter R. Bullock, Richard P. Selway, Kiriakos Xenitidis, and Anthony S. David. 2012. Prefrontal cortex dysfunction and 'jumping to conclusions': Bias or deficit? *Journal of Neuropsychology* 6: 65–78. [CrossRef]
Luo, Lin, Gigi Luk, and Ellen Bialystok. 2010. Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition* 114: 29–41. [CrossRef] [PubMed]
Marian, Viorica, and Michael Spivey. 2003. Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics* 24: 173–93. [CrossRef]
Markus, Hazel Rose, and Shinobu Kitayama. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* 98: 224–53. [CrossRef]
Markus, Hazel Rose, and Shinobu Kitayama. 2010. Cultures and Selves. *Perspectives on Psychological Science* 5: 420–30. [CrossRef]
Martin, Clara D., Kristof Strijkers, Mikel Santesteban, Carles Escera, Robert J. Hartsuiker, and Albert Costa. 2013. The impact of early bilingualism on controlling a language learned late: An ERP study. *Frontiers in Psychology* 4: 815. [CrossRef]
Martin-Rhee, Michelle M., and Ellen Bialystok. 2008. The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition* 11: 81–93. [CrossRef]
Masuda, Takahiko, and Richard E. Nisbett. 2001. Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology* 81: 922–34. [CrossRef]
Masuda, Takahiko, Phoebe C. Ellsworth, Batja Mesquita, Janxin Leu, Shigehito Tanida, and Ellen Van de Veerdonk. 2008. Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology* 94: 365–81. [CrossRef]
Meca, Alan, Raha F. Sabet, Colleen M. Farrelly, Cynthia G. Benitez, Seth J. Schwartz, Melinda Gonzales-Backen, and Elma I. Lorenzo-Blanco. 2017. Personal and cultural identity development in recently immigrated Hispanic adolescents: Links with psychosocial functioning. *Cultural Diversity and Ethnic Minority Psychology* 23: 348–61. [CrossRef] [PubMed]
Mezzacappa, Enrico. 2004. Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development* 75: 1373–86. [CrossRef] [PubMed]
Miyake, Akira, and Naomi P. Friedman. 2012. The nature and organization of individual differences in executive functions. *Current Directions in Psychological Science* 21: 8–14. [CrossRef] [PubMed]
Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter, and Tor D. Wager. 2000. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology* 41: 49–100. [CrossRef] [PubMed]
Miyamoto, Yuri, Richard E. Nisbett, and Takahiko Masuda. 2006. Culture and the physical environment: Holistic versus analytic perceptual affordances. *Psychological Science* 17: 113–19. [CrossRef]
Morton, J. Bruce, and Sarah N. Harper. 2007. What did Simon say? Revisiting the bilingual advantage. *Developmental Science* 10: 719–26. [CrossRef] [PubMed]
Nayak, Srishti, Hiba Z. Salem, and Amanda R. Tarullo. 2020. Neural mechanisms of response-preparation and inhibition in bilingual and monolingual children: Lateralized Readiness Potentials (LRPs) during a nonverbal Stroop task. *Developmental Cognitive Neuroscience* 41: 100740. [CrossRef]
Nguyen, Angela-MinhTu D., and Verónica Benet-Martínez. 2007. Biculturalism unpacked: components, measurement, individual differences, and outcomes. *Social and Personality Psychology Compass* 1: 101–14. [CrossRef]
Nguyen, Angela-MinhTu D., and Verónica Benet-Martínez. 2012. Biculturalism and adjustment. *Journal of Cross-Cultural Psychology* 44: 122–59. [CrossRef]
Nickels, Lyndsey, Solène Hameau, Vishnu KK Nair, Polly Barr, and Britta Biedermann. 2019. Ageing with bilingualism: Benefits and challenges. *Speech, Language and Hearing* 22: 32–50. [CrossRef]
Nisbett, Richard E., and Takahiko Masuda. 2003. Culture and point of view. *Proceedings of the National Academy Of Sciences* 100: 11163–70. [CrossRef]
Nisbett, Richard E., and Takahiko Masuda. 2006. Culture and Point of View. In *Biological and Cultural Bases of Human Inference*. Edited by Riccardo Viale, Daniel Andler and Lawrence A. Hirschfeld. Hillsdale: Lawrence Erlbaum Associates Publishers, pp. 49–70.
Nisbett, Richard E., and Yuri Miyamoto. 2005. The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences* 9: 467–73. [CrossRef]
Nisbett, Richard E., Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review* 108: 291–310. [CrossRef] [PubMed]
Norenzayan, Ara, Edward E. Smith, Beom Jun Kim, and Richard E. Nisbett. 2002. Cultural preferences for formal versus intuitive reasoning. *Cognitive Science* 26: 653–84. [CrossRef]
Oh, Seungmi, and Charlie Lewis. 2008. Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development* 79: 80–99. [CrossRef] [PubMed]
Ooi, Seok Hui, Winston D. Goh, Antonella Sorace, and Thomas H. Bak. 2018. From Bilingualism to bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition* 21: 867–79. [CrossRef]
Oyserman, Daphna, and Spike WS Lee. 2008. Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin* 134: 311–42. [CrossRef]

Paap, Kenneth R. 2019. Bilingualism in cognitive science: The characteristics and consequences of bilingual language control. In *The Cambridge Handbook of Bilingualism*. Edited by Annick De Houwer and Lourdes Ortega. Cambridge: Cambridge University Press, pp. 435–65.

Paap, Kenneth R., and Zachary I. Greenberg. 2013. There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* 66: 232–58. [CrossRef]

Paap, Kenneth R., and Oliver Sawi. 2014. Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology* 5: 962. [CrossRef]

Paap, Kenneth R., Hunter A. Johnson, and Oliver Sawi. 2015. Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69: 265–78. [CrossRef]

Paap, Kenneth R., Regina Anders-Jefferson, Lauren Mason, Katerinne Alvarado, and Brandon Zimiga. 2018. Bilingual advantages in inhibition or selective attention: more challenges. *Frontiers in Psychology* 9: 1409. [CrossRef]

Paap, Kenneth R., Regina Anders-Jefferson, Roman Mikulinsky, Shigeaki Masuda, and Lauren Mason. 2019. On the encapsulation of bilingual language control. *Journal of Memory and Language* 105: 76–92. [CrossRef]

Padilla, Amado M. 2006. Bicultural social development. *Hispanic Journal of Behavioral Sciences* 28: 467–97. [CrossRef]

Pavlenko, Aneta. 2005. *Emotions and Multilingualism*. Cambridge: Cambridge University Press.

Peal, Elizabeth, and Wallace E. Lambert. 1962. The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied* 76: 1–23. [CrossRef]

Pelham, Sabra D., and Lise Abrams. 2014. Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40: 313–25. [CrossRef]

Poulin-Dubois, Diane, Agnes Blaye, Julie Coutya, and Ellen Bialystok. 2011. The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology* 108: 567–79. [CrossRef]

Prior, Anat, and Brian MacWhinney. 2009. A bilingual advantage in task switching. *Bilingualism: Language and Cognition* 13: 253–62. [CrossRef]

Ramírez-Esparza, Nairán, Samuel D. Gosling, Verónica Benet-Martínez, Jeffrey P. Potter, and James W. Pennebaker. 2006. Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of Research in Personality* 40: 99–120. [CrossRef]

Sabbagh, Mark A., Fen Xu, Stephanie M. Carlson, Louis J. Moses, and Kang Lee. 2006. The development of executive functioning and theory of mind. A comparison of Chinese and U.S. Preschoolers. *Psychological Science* 17: 74–81. [CrossRef]

Samuel, Steven, Karen Roehr-Brackin, Hyensou Pak, and Hyunji Kim. 2018. Cultural effects rather than a bilingual advantage in cognition: A review and an empirical study. *Cognitive Science* 42: 2313–41. [CrossRef]

Sarma, U. Arathi, and Tissy Mariam Thomas. 2020. Breaking the limits of executive functions: Towards a sociocultural perspective. *Culture & Psychology* 26: 358–68. [CrossRef]

Schwartz, Seth J., Dina Birman, Verónica Benet-Martínez, and Jennifer B. Unger. 2017. Biculturalism: Negotiating multiple cultural streams. In *The Oxford Handbook of Acculturation and Health*. Edited by Seth J. Schwartz and Jennifer B. Unger. Oxford: Oxford University Press, pp. 29–47.

Shih, Margaret, and Diana T. Sanchez. 2005. Perspectives and research on the positive and negative implications of having multiple racial identities. *Psychological Bulletin* 131: 569–91. [CrossRef]

Soffietti, James P. 1960. Bilingualism and biculturalism. *The Modern Language Journal* 44: 275. [CrossRef]

Spiegler, Olivia, and Birgit Leyendecker. 2017. Balanced cultural identities promote cognitive flexibility among immigrant children. *Frontiers in Psychology* 8: 1579. [CrossRef] [PubMed]

Spivey, Michael J., and Viorica Marian. 1999. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10: 281–84. [CrossRef]

Stroink, Mirella L., and Richard N. Lalonde. 2009. Bicultural identity conflict in second-generation Asian Canadians. *The Journal of Social Psychology* 149: 44–65. [CrossRef]

Suárez-Orozco, Carola, Marcelo M. Suárez-Orozco, and Irina Todorova. 2008. *Learning a New Land: Immigrant Students in American Society*. Cambridge: Belknap Press/Harvard University Press.

Tadmor, Carmit T., and Philip E. Tetlock. 2006. Biculturalism. *Journal of Cross-Cultural Psychology* 37: 173–90. [CrossRef]

Thierry, Guillaume, and Yan Jing Wu. 2007. Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences* 104: 12530–35. [CrossRef]

Tomoschuk, Brendan, Victor S. Ferreira, and Tamar H. Gollan. 2018. When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition* 22: 516–36. [CrossRef]

Tran, Crystal D., Maria M. Arredondo, and Hanako Yoshida. 2015. Differential effects of bilingualism and culture on early attention: A longitudinal study in the U.S., Argentina, and Vietnam. *Frontiers in Psychology* 6: 795. [CrossRef]

Tran, Crystal D., Maria M. Arredondo, and Hanako Yoshida. 2018. Early executive function: The influence of culture and bilingualism. *Bilingualism: Language and Cognition* 22: 714–32. [CrossRef]

Treffers-Daller, Jeanine, Zehra Ongun, Julia Hofweber, and Michal Korenar. 2020. Explaining individual differences in executive functions performance in multilinguals: The impact of code-switching and alternating between multicultural identity styles. *Frontiers in Psychology* 11: 561088. [CrossRef]

Triandis, Harry C. 1989. The self and social behavior in differing cultural contexts. *Psychological Review* 96: 506–20. [CrossRef]

Triandis, Harry C. 2001. Individualism-collectivism and personality. *Journal of Personality* 69: 907–24. [CrossRef] [PubMed]

Tse, Chi-Shing, and Jeanette Altarriba. 2012. The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition* 15: 663–76. [CrossRef]

Tse, Chi-Shing, and Jeanette Altarriba. 2014. The relationship between language proficiency and attentional control in Cantonese-English bilingual children: Evidence from Simon, Simon switching, and working memory tasks. *Frontiers in Psychology* 5: 954. [CrossRef] [PubMed]

Valian, Virginia. 2014. Bilingualism and cognition. *Bilingualism: Language and Cognition* 18: 3–24. [CrossRef]

Varnum, Michael EW, Igor Grossmann, Shinobu Kitayama, and Richard E. Nisbett. 2010. The origin of cultural differences in cognition: Evidence for the social orientation hypothesis. *Current Directions in Psychological Science* 19: 9–13. [CrossRef]

Von Bastian, Claudia C., Alessandra S. Souza, and Miriam Gade. 2016. No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General* 145: 246–58. [CrossRef] [PubMed]

Ward, Colleen, Caroline Ng Tseung-Wong, Agnes Szabo, Tamara Qumseya, and Uma Bhowon. 2018. Hybrid and alternating identity styles as strategies for managing multicultural identities. *Journal of Cross-Cultural Psychology* 49: 1402–39. [CrossRef]

Weissberger, Gali H., Tamar H. Gollan, Mark W. Bondi, Daniel A. Nation, Lawrence A. Hansen, Douglas Galasko, and David P. Salmon. 2019. Neuropsychological deficit profiles, vascular risk factors, and neuropathological findings in hispanic Older adults with autopsy-confirmed Alzheimer's disease. *Journal of Alzheimer's Disease* 67: 291–302. [CrossRef]

Wiseheart, Melody, Mythili Viswanathan, and Ellen Bialystok. 2014. Flexibility in task switching by monolinguals and bilinguals. *Bilingualism: Language and Cognition* 19: 141–46. [CrossRef]

Yang, Hwajin, Andree Hartanto, and Sujin Yang. 2016. The Complex nature of Bilinguals' language usage modulates task-switching outcomes. *Frontiers in Psychology* 7: 560. [CrossRef]

Yang, Sujin, and Hwajin Yang. 2016. Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *Journal of Experimental Child Psychology* 146: 121–36. [CrossRef] [PubMed]

Yang, Sujin, Hwajin Yang, and Barbara Lust. 2011. Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism: Language and Cognition* 14: 412–22. [CrossRef]

Ye, Yanyan, Lei Mo, and Qihan Wu. 2017. Mixed cultural context brings out bilingual advantage on executive function. *Bilingualism: Language and Cognition* 20: 844–52. [CrossRef]

Yow, W. Quin, and Xiaoqian Li. 2015. Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology* 6: 164. [CrossRef] [PubMed]

languages

Article

Cultural Experience Influences Multisensory Emotion Perception in Bilinguals

Peiyao Chen [1,*], Ashley Chung-Fat-Yim [2] and Viorica Marian [2]

[1] Department of Psychology, Swarthmore College, Swarthmore, PA 19081, USA
[2] Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208, USA; ashley.chungfatyim@northwestern.edu (A.C.-F.-Y.); v-marian@northwestern.edu (V.M.)
* Correspondence: pchen3@swarthmore.edu

Abstract: Emotion perception frequently involves the integration of visual and auditory information. During multisensory emotion perception, the attention devoted to each modality can be measured by calculating the difference between trials in which the facial expression and speech input exhibit the same emotion (congruent) and trials in which the facial expression and speech input exhibit different emotions (incongruent) to determine the modality that has the strongest influence. Previous cross-cultural studies have found that individuals from Western cultures are more distracted by information in the visual modality (i.e., visual interference), whereas individuals from Eastern cultures are more distracted by information in the auditory modality (i.e., auditory interference). These results suggest that culture shapes modality interference in multisensory emotion perception. It is unclear, however, how emotion perception is influenced by cultural immersion and exposure due to migration to a new country with distinct social norms. In the present study, we investigated how the amount of daily exposure to a new culture and the length of immersion impact multisensory emotion perception in Chinese-English bilinguals who moved from China to the United States. In an emotion recognition task, participants viewed facial expressions and heard emotional but meaningless speech either from their previous Eastern culture (i.e., Asian face-Mandarin speech) or from their new Western culture (i.e., Caucasian face-English speech) and were asked to identify the emotion from either the face or voice, while ignoring the other modality. Analyses of daily cultural exposure revealed that bilinguals with low daily exposure to the U.S. culture experienced greater interference from the auditory modality, whereas bilinguals with high daily exposure to the U.S. culture experienced greater interference from the visual modality. These results demonstrate that everyday exposure to new cultural norms increases the likelihood of showing a modality interference pattern that is more common in the new culture. Analyses of immersion duration revealed that bilinguals who spent more time in the United States were equally distracted by faces and voices, whereas bilinguals who spent less time in the United States experienced greater visual interference when evaluating emotional information from the West, possibly due to over-compensation when evaluating emotional information from the less familiar culture. These findings suggest that the amount of daily exposure to a new culture and length of cultural immersion influence multisensory emotion perception in bilingual immigrants. While increased daily exposure to the new culture aids with the adaptation to new cultural norms, increased length of cultural immersion leads to similar patterns in modality interference between the old and new cultures. We conclude that cultural experience shapes the way we perceive and evaluate the emotions of others.

Keywords: emotion; modality interference; cultural immersion; cultural exposure; bilingualism

Citation: Chen, Peiyao, Ashley Chung-Fat-Yim, and Viorica Marian. 2022. Cultural Experience Influences Multisensory Emotion Perception in Bilinguals. *Languages* 7: 12. https://doi.org/10.3390/languages7010012

Academic Editor: John W. Schwieter

Received: 25 October 2021
Accepted: 23 December 2021
Published: 10 January 2022

1. Cultural Experience Influences Multisensory Emotion Perception in Bilinguals

Every year, thousands of people move to a new country for educational purposes, career opportunities, or personal endeavors. Depending on the destination country, these individuals are exposed to new languages, cultures, and social norms, all of which have been shown to shape the way they perceive, process, and organize information in their environment (Marian, forthcoming). Especially for those trying to adapt to a new culture, the ability to read the socio-emotional cues of others may ease the adjustment process by increasing the likelihood of building new support systems. Although past research has examined the role of cultural immersion on cognitive processes, such as color perception (Athanasopoulos et al. 2010), face perception (Derntl et al. 2009, 2012), object perception (Kitayama et al. 2003), and categorization strategy (Cook et al. 2006), few studies have examined the effect of cultural immersion experience on multisensory emotion perception. The current experiment investigated whether daily exposure to a new culture and the length of immersion influence multisensory emotion perception in bilinguals.

Previous studies have reported that multisensory emotion perception is shaped by the perceiver's cultural background (Liu et al. 2015a, 2015b; Tanaka et al. 2010). Specifically, individuals from East Asian cultures are more influenced by the emotional expression in the auditory modality (i.e., tone of voice), while individuals from Western cultures are more influenced by the emotional expression in the visual modality (i.e., facial expressions). These cultural differences in modality interference have been attributed to display rules (Liu et al. 2015a, 2015b; Tanaka et al. 2010), which are a set of rules learned from an early age through socialization that regulate how to appropriately express emotions (Ekman and Friesen 1971). In East Asian collectivist societies, individuals tend to suppress and control their emotions and personal preferences to accommodate the thoughts and feelings of others (Markus and Kitayama 1991; Matsumoto et al. 1998; Morling et al. 2002; Weisz et al. 1984). On the contrary, in Western individualistic societies, emotions are expressed with the intention to influence others, leading to more direct eye contact amongst Westerners than Easterners (Argyle et al. 1986; McCarthy et al. 2006, 2008). An open question is whether the degree of interference from the visual or auditory modality for individuals who migrate to a new country is fixed across cultures (i.e., they maintain their old pre-existing schemas) or changes depending on the cultural context.

Although there is a prescribed set of rules within each culture, culturally-specific schemas and behaviors are not static and can change as a consequence of new cultural experiences. For example, after being immersed in the United Kingdom for 3.5 years, immigrants from Greece began to perceive the colors blue and green in a way that resembled native speakers from the United Kingdom (Athanasopoulos et al. 2010). Of particular relevance to the current study, Liu et al. (2017) found that Chinese immigrants who moved to Canada were more distracted by irrelevant facial expressions, mirroring the behavior of North American participants. However, Chinese immigrants' brain activity revealed that they were equally distracted by the irrelevant facial and vocal expressions, mirroring the brain activity of native Mandarin speakers from China (Liu et al. 2015a). These results suggest that cultural immersion leads to culturally-specific changes in emotion perception, at least at the behavioral level. Note that the participants in the study by Liu et al. (2017) migrated to Canada between the ages of 10 and 18 and were tested exclusively on emotional stimuli from their old Eastern culture (i.e., Asian face with Mandarin speech). Therefore, some of the participants had more experience immersed in the North American culture than in Eastern cultures; it remains unknown how they would perceive multisensory emotions in their new culture (i.e., Caucasian face with English speech). We aim to build on these findings by having participants judge multisensory emotional stimuli from their old and new culture and splitting participants into two groups based on the duration of time they lived in the United States. This allows us to investigate whether the length of immersion influences multisensory emotion perception.

The process by which emotional behaviors change due to repeated interactions with a new cultural context has been described as emotional acculturation (Cosedine et al. 2014; De Leersnyder et al. 2013). De Leersnyder et al. (2011) compared the emotional patterns of first-generation Korean immigrants living in the United States and Turkish first-generation immigrants living in Belgium to the emotional patterns of members of the majority culture in each country. Participants rated the extent to which they experienced feeling angry, ashamed, happy, proud, and respectful on a scale from 1 (Not at all) to 7 (Extremely) in response to a set of scenarios. The emotional responses of first-generation immigrants were lower than the average responses of members of the majority culture. However, with increased exposure to the new culture and repeated daily social interactions with members of the majority culture, the emotional patterns of first-generation minorities became more similar to those of the majority culture. These findings from the emotional acculturation literature suggest that immigrants may maintain the emotional patterns from their heritage culture, but increased exposure to the new culture changes their emotional patterns to align with the new culture.

While increased cultural experience can result in new patterns of emotion processing, it does not necessarily mean that the behavioral patterns that were developed in the old cultural context are completely discarded. In fact, those who identify with two cultures sometimes exhibit different behavioral patterns, depending on the cultural context (Berry 1997). This is known as the Cultural Frame Switching hypothesis (Hong et al. 2000; LaFromboise et al. 1993), and refers to bicultural individuals accessing and shifting their mental schemas based on the culture that was most recently activated. Culturally-specific images have also been found to elicit different cultural tendencies in attribution styles (Benet-Martínez et al. 2002; Hong et al. 2003; Hong et al. 2000), personality traits (Morris and Mok 2011; Ramírez-Esparza et al. 2006), and self-concept or identity (Cheng et al. 2006; Ross et al. 2002; West et al. 2018). Language has also been shown to elicit culture-specific behaviors (De Leersnyder et al. 2020; Marian and Kaushanskaya 2004; Panayiotou 2004; Perunovic et al. 2007). For example, Greek-English bilinguals reacted to the same story differently depending on the language in which the story was read (Panayiotou 2004). When the story was read to them in Greek, the bilinguals reported feeling concerned for the protagonist. In contrast, when the story was read to them in English, the bilinguals reported feeling indifferent towards the protagonist. Cultural cues can influence social cognition and potentially play a unique role in shaping the way bilinguals who migrate to a new country evaluate multisensory emotions.

In previous studies examining the Cultural Frame Switching hypothesis, the cultural context was primed with cultural images or by languages that were independent from the task itself. For multisensory emotion processing, the cultural context is primed with stimuli that are embedded within the task (i.e., language and face). For bilinguals who learn to communicate in a new language and recognize the faces of a new racial group, both the visual input (i.e., facial expressions of different racial groups) and the auditory input (i.e., vocal expressions in different languages) may serve as strong cues to signal different cultural contexts.

The present study examined whether cultural experience affects multisensory emotion perception when a bilingual moves from an East Asian culture to a Western culture. Chinese-English bilinguals who were born and raised in China and moved to the United States were presented with face-language pairs from their old Eastern culture (i.e., Asian face with Mandarin speech) and new Western culture (i.e., Caucasian face with English speech). There are three possible ways that bilinguals migrating from China to the United States would exhibit modality interference. The first possibility is that bilinguals would adopt new Western cultural norms by increasing their reliance on the visual modality when presented with Eastern and Western emotional information, replicating the behavioral findings by Liu et al. (2017). The second possibility is that bilinguals would hold both Eastern and Western norms and exhibit culturally-specific behaviors depending on the cultural context, supporting the Cultural Frame Switching hypothesis (Hong et al. 2000; LaFromboise et al. 1993). Specifically, they would show greater interference from the visual modality when presented with emotional input from the West, and greater interference from the auditory modality or equivalent interference from both modalities when presented with emotional input from the East. The third possibility is that bilinguals would maintain their old Eastern cultural norms by showing greater interference from the auditory modality across both cultures, in line with the electrophysiological findings by Liu et al. (2017).

In addition, we investigated whether the amount of daily exposure to a new culture and the length of immersion influence multisensory emotion perception. Both the amount of daily exposure and the length of immersion have been previously found to influence emotional acculturation (Cosedine et al. 2014; De Leersnyder et al. 2011). Specifically, we compared Chinese-English bilinguals with shorter immersion experience to those with longer immersion experience in the U.S., as well as Chinese-English bilinguals with low daily exposure to those with high daily exposure to the U.S. culture. A closer look at these two factors would enable us to tease apart the role of accumulated experience (i.e., immersion) from that of everyday experience (i.e., exposure).

2. Materials and Methods

2.1. Participants

Forty-nine Chinese-English bilinguals between the ages of 18 and 35 were recruited through posters around campus and university email listservs. Participants were compensated for their participation with an Amazon gift card at the rate of $10/hour. Chinese-English bilinguals were recruited based on four inclusionary criteria: (1) being proficient in English and Mandarin, (2) living in North America at the time of testing, (3) being born and raised in China to Chinese parents, and (4) previously living in China for at least 6 years before moving to a Western country. The study was conducted remotely via the Internet and informed consent was obtained by all participants in their native language.

Three participants did not complete the task and another four participants performed at or below chance level on the emotion recognition task. In addition, one participant had recently moved to the United States, which resulted in a mean length of time in China that was 3 standard deviations greater than the group's mean. The remaining 41 participants (31 females, M_{Age} = 24.07 years, SD_{Age} = 3.53) had spent an average of 19.3 years in China (*SD* = 3.93 years, range: 9 years to 26 years) and, by the time of testing, had spent an average of 4.70 years in Western countries (*SD* = 3.51 years, range: 4 months to 14 years). Participants' language background information was obtained using an adapted version of the Language Experience and Proficiency Questionnaire (Marian et al. 2007). All participants acquired Mandarin from birth and rated their proficiency in Mandarin as 9.53 out of 10 (*SD* = 0.63). On average, participants learned English before the age of 7 (*M* = 6.95, *SD* = 3.37) and rated their proficiency in English as 7.54 out of 10 (*SD* = 1.21). Their daily exposure to Eastern and Western cultures was 46.54% (*SD* = 17.04) and 49.02% (*SD* = 19.04), respectively. There was no relation between immersion length and percentage of daily exposure to Western culture, $r(41) = 0.12$, $p = 0.45$. All participants had normal or corrected-to-normal vision and no hearing impairments. The study was approved by the local Institutional Review Board.

2.2. Materials

Language Experience and Proficiency Questionnaire (LEAP-Q)

An adapted version of the LEAP-Q (Marian et al. 2007) was used to assess each participant's linguistic and cultural backgrounds. The first set of questions contained demographic questions, including age, years of formal education, gender, and any history of hearing or vision impairments. The second set of questions targeted their language history. Participants were asked to report the languages they spoke, including any non-native languages. For each language listed, participants rated their level of proficiency from 1 to 10 (1 = very low and 10 = perfect) and reported the age of acquisition. The third set of questions pertained to the participant's cultural background. Participants were asked to list the cultures they identified with. For each culture listed, participants rated the extent to which they identified with each culture on a scale from 0 to 10 (0 = no identification and 10 = complete identification), and the percentage of time spent exposed to each culture. Lastly, participants listed the countries they had previously lived in, and the duration of time spent in each country.

2.3. Stimuli

2.3.1. Vocal Stimuli

Twenty Mandarin and 20 English pseudo-sentences were selected from two validated vocal emotional stimuli databases (Mandarin: Liu and Pell 2012; English: Pell et al. 2009). Pseudo-sentences followed the segmental properties of each language respectively but included no semantic information. The pseudo-sentences were spoken in five basic emotions (happiness, sadness, disgust, fear, and anger) by 2 female and 2 male native speakers of each language, resulting in a unimodal voice list of 20 pseudo-sentences (4 voices × 5 emotions = 20 pseudo-sentences) in each language. Based on the normed data within each database, the Mandarin and English pseudo-sentences were matched on recognition rate, emotional intensity, and duration, *t*s < 1 (Table 1).

Table 1. Recognition accuracy and emotional intensity ratings for the sentences and faces from each culture.

Stimuli	Culture	Percent Recognition Rate	Emotional Intensity (0 to 5 for Sentences and 0 to 9 for Faces)	Duration (in Seconds)
Sentences	Mandarin	86 (7.3)	3.3 (0.6)	1.78 (0.26)
	English	88 (7.4)	3.4 (0.4)	1.79 (0.19)
Faces	Asian	83 (12.1)	5.6 (0.6)	-
	Caucasian	84 (12.6)	5.7 (1.0)	-

2.3.2. Face Stimuli

Twenty Asian faces and 20 Caucasian faces were selected from two databases (Asian faces: Taiwanese Facial Expression Image Database by Chen and Yen (2007); Caucasian faces: Karolinska Directed Emotional Faces Database by Lundqvist et al. (1998)). Five different facial expressions (happiness, sadness, disgust, fear, and anger) were displayed by 2 female and 2 male actors from each database, resulting in a unimodal face list of 20 faces for each culture (4 faces × 5 emotions = 20 faces). Based on the normed data in each database, the Asian and Caucasian faces were matched on recognition rate and emotional intensity, *t*s < 1 (Table 1). To maintain consistency in size, brightness, and contrast, the images were re-processed in GIMP 2.9.8 (GIMP Development Team 2015) to the same dimension (345 pixels wide × 430 pixels high), resolution (300 dpi), and intensity (grayscale).

2.3.3. Bimodal Face-Voice Stimuli

For each culture, bimodal face-voice stimuli were created by pairing a unique voice with a unique face of the same gender. The same voice was always paired with the same face to maintain consistency in identity (Figure 1). The emotion displayed across modalities could either be the same (congruent condition; e.g., happy face and happy voice) or different (incongruent condition; e.g., happy face and sad voice). Because a total of five emotions were used in the study with four different speakers/actors, each face was paired once with the voice of the same emotion for a total of 20 bimodal congruent trials and once with each of the remaining four emotions for a total of 80 bimodal incongruent trials. As a result, four bimodal lists were created, each containing 20 congruent trials and 20 of the 80 incongruent trials.

Figure 1. An example of bimodal stimuli from each culture. In the left-hand panel, an Asian face is paired with meaningless Mandarin speech (East Asian culture). In the right-hand panel, a Caucasian face is paired with meaningless English speech (Western culture). (Asian face—Image ID fea118: Adapted with permission from Chen and Yen (2007); Mandarin pseudo-sentence: Adapted with permission from Liu and Pell (2012); Caucasian face—Image ID AM08ANS: Adapted with permission from Lundqvist et al. (1998); English pseudo-sentence: Adapted with permission from Pell et al. (2009).)

2.3.4. Fillers

To discourage participants from developing a strategy, 12 bimodal filler trials with new faces and voices were used. Three trials contained a red dot (radius = 20 mm in size) on the cheek of the face and another three trials had a 500 ms beep inserted in the speech stream.

2.4. Design and Procedure

The emotion recognition task consisted of two tasks. In the face task, participants were instructed to identify the emotion of the face and ignore the emotion of the voice. In the voice task, participants were instructed to identify the emotion of the voice and ignore the emotion of the face. Each participant was assigned one of the four bimodal lists, a unimodal face list, and a unimodal voice list from each culture. The same bimodal list was used for the face task and voice task.

For each trial, a prompt instructing participants to identify either the emotion of the voice (i.e., voice task; "Judge the Voice Emotion") or the emotion on the face (i.e., face task; "Judge the Face Emotion") appeared first. Once the participant clicked on the prompt, a bimodal or a unimodal stimulus appeared in the middle of the screen. On unimodal face trials, a face appeared for a random duration between 1500 to 2000 ms with a 100 ms interval, which is consistent with the duration of 95% of the vocal stimuli. On unimodal voice trials, a fixation cross appeared in the middle of the screen while a pseudo-sentence was presented through the participant's speakers. On bimodal face-voice trials, a face appeared for the duration of the speech. For all three trial types, participants were then presented with a display of five emotion words in English (happiness, sadness, disgust, fear, and anger) and instructed to select the emotion they perceived by clicking the box next to the emotion as quickly as possible. After selecting an emotion, they rated the intensity of the perceived emotion on a scale from 0 (not intense at all) to 6 (extremely intense). On filler trials, a beep in the speech stream or a red dot on the cheek of the face appeared for 500 ms within the last 600 ms to 700 ms of a trial. Instead of rating the intensity of the emotion, participants reported whether they saw a red dot on the face or heard a beep by clicking "Yes" or "No".

The face and voice tasks were presented in separate blocks, with the order of presentation counterbalanced across participants. Within each task, unimodal, bimodal congruent, bimodal incongruent, and filler trials from both cultures were intermixed and randomly presented. The face and voice tasks did not differ in the number of switches between cultures, $t(44) = -0.51$, $p = 0.61$, and between trial types, $t(44) = -0.99$, $p = 0.33$. In addition, for the bimodal trials, the number of switches between congruent and incongruent trials was equivalent in the face and voice tasks, $t(44) = -0.41$, $p = 0.68$. Participants were provided 10 practice trials at the start of each task and three breaks that were embedded throughout. The entire testing session was 60 to 90 minutes long.

3. Results

Response times (RTs) were measured from the onset of the display with the five emotions until a response was made. Only correct trials were included in the analyses for RTs and intensity ratings. RTs below 500 ms and above 5000 ms were discarded from the analyses. For accuracy rates and intensity ratings, modality interference was computed as the difference between the bimodal congruent trials and bimodal incongruent trials. For RTs, modality interference was computed as the difference between the bimodal incongruent trials and bimodal congruent trials. The higher the modality interference in the face task, the greater the interference from the irrelevant voice. The higher the modality interference in the voice task, the greater the interference from the irrelevant face.

3.1. Modality Interference across Participants

The mean accuracy rates, response times, and intensity ratings by condition, culture, and task are presented in Table 2. A 2-way ANOVA with task (face vs. voice) and culture (East vs. West) as within-subjects factors on accuracy rates yielded a main effect of task, $F(1,40) = 10.72$, $p = 0.002$, $\eta_p{}^2 = 0.21$, showing that modality interference was larger in the voice task ($M = 0.13$, $SE = 0.019$) than the face task ($M = 0.065$, $SE = 0.012$), but no effect of culture was observed, $F < 1$. The task by culture interaction was significant, $F(1,40) = 9.13$, $p = 0.004$, $\eta_p{}^2 = 0.19$. When the emotional input was from the West, there was a larger modality interference in the voice task than the face task, $F(1,40) = 15.52$, $p < 0.001$, $\eta_p{}^2 = 0.28$, suggesting that participants experienced greater interference from the visual modality than auditory modality. No difference between modalities was found for emotional input from the East, $F < 1$ (Figure 2). To examine the Cultural Frame Switching hypothesis, we also examined the interaction of task and culture by comparing Eastern and Western stimuli for the voice and face tasks separately. There was a larger modality interference for Western stimuli than Eastern stimuli on the voice task, $F(1,40) = 7.73$, $p = 0.008$, $\eta_p{}^2 = 0.16$, suggesting that participants experienced greater interference from the irrelevant face when evaluating emotional speech from the West. The modality interference was larger for Eastern stimuli than for Western stimuli on the face task, but this difference did not reach significance $F(1,40) = 3.59$, $p = 0.065$, $\eta_p{}^2 = 0.082$, suggesting that participants showed a tendency towards greater interference from the irrelevant speech when evaluating facial expressions from the East. The analyses on RTs and intensity ratings yielded no significant effects or interactions, all *ps* > 0.16.

Table 2. Mean accuracy rates (ACC), response times in ms (RT), and intensity ratings (IR) across cultures, tasks, and conditions (standard deviations are in parentheses).

Measure	Culture	Task	Bimodal Congruent	Bimodal Incongruent	Modality Interference
ACC	East	Face	0.86 (0.097)	0.77 (0.14)	0.087 (0.12)
		Voice	0.88 (0.11)	0.78 (0.15)	0.10 (0.11)
	West	Face	0.80 (0.095)	0.75 (0.11)	0.044 (0.090)
		Voice	0.78 (0.12)	0.61 (0.16)	0.16 (0.18)
RT	East	Face	1484 (310)	1601 (317)	117 (268)
		Voice	1583 (370)	1676 (343)	93 (281)
	West	Face	1474 (290)	1591 (334)	117 (269)
		Voice	1757 (416)	1806 (377)	49 (338)
IR	East	Face	4.05 (0.90)	3.74 (0.94)	0.31 (0.51)
		Voice	4.19 (0.77)	3.93 (0.93)	0.25 (0.46)
	West	Face	3.96 (0.91)	3.71 (0.95)	0.25 (0.43)
		Voice	3.96 (0.94)	3.63 (0.92)	0.33 (0.52)

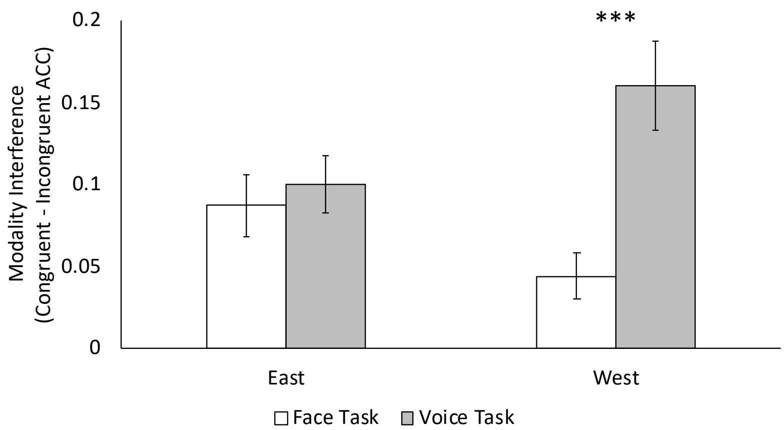

Figure 2. Modality interference effects in accuracy rates (bimodal congruent trials minus bimodal incongruent trials). A larger modality interference in the face task represents greater interference from the auditory modality, whereas a larger modality interference in the voice task represents greater interference from the visual modality. Chinese-English bilinguals experienced a larger modality interference in the voice task than in the face task for the emotional information from the West (i.e., visual interference). Error bars represent standard error. *** $p < 0.001$.

3.2. Long vs. Short Immersion

To examine the effect of immersion length, participants were divided into two groups using a median split based on the duration of time spent in the United States. Bilinguals in the shorter immersion group ($N = 22$) spent less than or equal to 4 years in the United States, whereas those in the longer immersion group ($N = 19$) spent more than 4.5 years in the United States. The groups were matched on Chinese proficiency, age of English acquisition, and years of formal education, *p*s > 0.06. The bilinguals in the short immersion group were younger, $t(39) = -2.19$, $p = 0.034$, and rated their English proficiency lower than the bilinguals in the long immersion group, $t(39) = -2.11$, $p = 0.041$ (Table 3). The difference in English proficiency is not surprising given that the bilinguals in the long immersion group had lived in an English-speaking country for a longer period of time and were likely more confident in their English-speaking abilities. English proficiency correlated with the accuracy rates on the voice task when listening to English speech, $r(41) = -0.32$, $p = 0.044$. Age and years of formal education did not correlate with any of the dependent measures on the emotion recognition task (*p*s > 0.12). Separate three-way ANOVAs with task (face vs. voice) and culture (East vs. West) as the within-subjects factors and length of immersion (short vs. long) as the between-subjects factor were performed on accuracy rates, RTs, and intensity ratings. The mean accuracy rates, RTs, and intensity ratings by condition, task, culture, and length of immersion are presented in Table 4.

Table 3. Language background of short vs. long immersion and low vs. high exposure groups.

	Short Immersion	Long Immersion	*p*-Value	Low Exposure	High Exposure	*p*-Value
N	22	19		18	23	
Gender	3 M, 19 F	6 M, 13 F		4 M, 14 F	5 M, 15 F	
Age in Years	23.00 (2.74)	25.32 (3.99)	0.034 *	23.39 (3.87)	24.61 (3.23)	0.28
Years of Education	16.52 (2.54)	18.23 (3.19)	0.06	16.58 (2.70)	17.91 (3.07)	0.16
Chinese Proficiency	9.68 (0.48)	9.37 (0.76)	0.12	9.56 (0.62)	9.52 (0.67)	0.87
English Proficiency	7.18 (1.10)	7.95 (1.22)	0.041 *	7.17 (1.15)	7.83 (1.19)	0.08
English AoA	6.86 (2.13)	7.06 (4.48)	0.86	7.47 (2.42)	6.55 (3.96)	0.40

Note. * $p < 0.05$. AoA = Age of Acquisition. Chinese proficiency and English proficiency were rated out of 10.

Table 4. Mean accuracy rates (ACC), response times in ms (RT), and intensity ratings (IR) across cultures, tasks, length of immersion, and conditions (standard deviations are in parentheses).

Culture	Measure	Task	Immersion	Bimodal Congruent	Bimodal Incongruent	Modality Interference
East	ACC	Face	Short	0.84 (0.093)	0.74 (0.10)	0.10 (0.13)
			Long	0.88 (0.099)	0.81 (0.16)	0.071 (0.11)
		Voice	Short	0.87 (0.098)	0.77 (0.17)	0.11 (0.12)
			Long	0.89 (0.11)	0.80 (0.12)	0.090 (0.084)
	RT	Face	Short	1521 (318)	1688 (341)	167 (282)
			Long	1441 (304)	1501 (259)	60 (245)
		Voice	Short	1661 (437)	1791 (347)	130 (300)
			Long	1492 (256)	1543 (294)	51 (258)
	IR	Face	Short	4.13 (0.89)	3.83 (0.92)	0.29 (0.48)
			Long	3.97 (0.92)	3.63 (0.97)	0.34 (0.56)
		Voice	Short	4.26 (0.73)	3.92 (0.98)	0.33 (0.44)
			Long	4.11 (0.82)	3.94 (0.88)	0.16 (0.46)
West	ACC	Face	Short	0.77 (0.092)	0.73 (0.084)	0.039 (0.083)
			Long	0.83 (0.090)	0.78 (0.12)	0.050 (0.10)
		Voice	Short	0.78 (0.097)	0.57 (0.15)	0.22 (0.18)
			Long	0.77 (0.14)	0.66 (0.16)	0.11 (0.16)
	RT	Face	Short	1495 (300)	1637 (347)	141 (261)
			Long	1448 (283)	1538 (319)	89 (284)
		Voice	Short	1875 (460)	1940 (405)	65 (378)
			Long	1620 (316)	1650 (277)	30 (294)
	IR	Face	Short	4.08 (0.86)	3.86 (.85)	0.22 (0.32)
			Long	3.82 (0.97)	3.53 (1.04)	0.28 (0.54)
		Voice	Short	4.08 (0.84)	3.72 (0.97)	0.36 (0.60)
			Long	3.82 (1.04)	3.52 (0.86)	0.30 (0.43)

In the accuracy rates analyses, there was a marginally significant three-way interaction between task, culture, and immersion, $F(1,39) = 3.63$, $p = 0.064$, $\eta_p^2 = 0.085$. The main effect of task, $F(1,39) = 10.25$, $p = 0.003$, $\eta_p^2 = 0.21$, and the interaction between task and culture, $F(1,39) = 8.83$, $p = 0.004$, $\eta_p^2 = 0.19$, were significant, but the main effect of immersion was not, $F(1,39) = 2.22$, $p = 0.14$. To breakdown the three-way interaction, separate two-way ANOVAs were conducted for each immersion group. Among the Chinese-English bilinguals who lived in the U.S. for a shorter duration, the task by culture interaction was significant, $F(1,21) = 16.59$, $p < 0.001$, $\eta_p^2 = 0.44$. Modality interference was larger in the voice task than the face task when the emotional input was from the West, $F(1,21) = 16.71$, $p < 0.001$, $\eta_p^2 = 0.44$, but there were no differences in modality interference between the two tasks when the emotional input was from the East, $F < 1$ (Figure 3). Furthermore, the main effect of task was significant, $F(1,21) = 8.46$, $p = 0.008$, $\eta_p^2 = 0.29$, in which the voice task ($M = 0.16$, $SE = 0.029$) produced a larger modality interference than the face task ($M = 0.069$, $SE = 0.016$). The effect of culture was not significant, $F < 1$. Among the Chinese-English bilinguals who were immersed in the U.S. for a longer duration, neither the main effects nor the interaction were significant, *p*s > 0.13. These findings demonstrate that with increased immersion experience, bilinguals experience similar patterns in modality interference across their two cultures. Analyses of RTs and intensity ratings yielded no significant main effects or interactions, *p*s > 0.15.

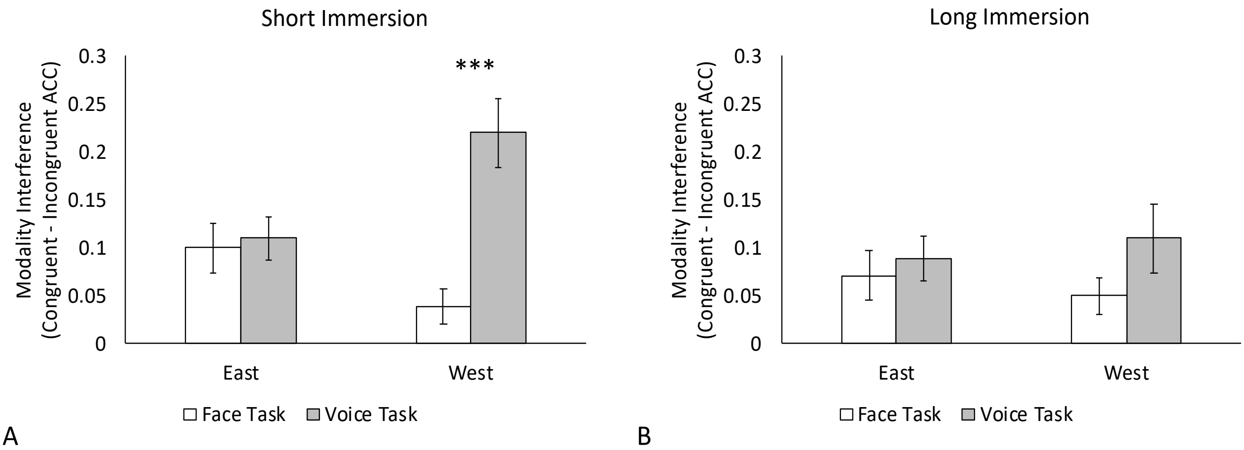

Figure 3. Modality interference effects in accuracy rates (bimodal congruent trials minus bimodal incongruent trials) for Chinese-English bilinguals immersed in North America for a short duration (**A**) and long duration (**B**). The short immersion group showed a larger modality interference in the voice task than the face task for the emotional information from the West (i.e., visual interference). The long immersion group showed no differences between tasks in both cultures. Error bars represent standard error. *** $p < 0.001$.

3.3. Low vs. High Daily Exposure

To examine the effect of cultural exposure on multisensory emotion perception, participants were divided into low (less than 50%; $N = 18$) and high (50% or more; $N = 23$) levels of cultural exposure to Western culture using a median split, based on their self-reported percentage of daily exposure to Western culture. The low and high daily exposure groups were matched on age, years of formal education, English proficiency, Chinese proficiency, and age of English acquisition, all *p*s > 0.082 (Table 3). Separate three-way ANOVAs with task (face, voice) and culture (East, West) as the within-subjects factors and exposure (low vs. high) as the between-subjects factor were performed on accuracy rates, RTs, and intensity ratings. Mean accuracy rates, RTs, and intensity ratings by condition, task, culture, and amount of exposure to the U.S. culture are shown in Table 5.

In the RT analyses, the three-way interaction of task, culture, and exposure was marginally significant, $F(1,39) = 4.03$, $p = 0.052$, $\eta_p{}^2 = 0.094$. Separate two-way ANOVAs were performed for each exposure group. In the bilingual group with low exposure to the U.S. culture, there was a marginally significant culture by task interaction, $F(1,17) = 3.96$, $p = 0.063$, $\eta_p{}^2 = 0.19$. A larger modality interference was found in the face task than the voice task when the emotional input was from the West, $F(1,17) = 6.59$, $p = 0.020$, $\eta_p{}^2 = 0.28$, but there were no differences between tasks when the emotional input was from the East, $F < 1$ (Figure 4a). These findings suggest that Chinese-English bilinguals with low exposure to the U.S. culture are more impacted by vocal expressions than facial expressions when presented with emotional stimuli from the West. This pattern of auditory interference coincides with their East Asian culture. The main effects of task and culture were not significant, *p*s > 0.092. In the bilingual group with high exposure to the U.S. culture, none of the main effects and interactions reached significance, *Fs* < 1.

Table 5. Mean accuracy rates (ACC), response times in ms (RT), and intensity ratings (IR) across cultures, tasks, level of exposure to Western culture, and conditions (standard deviations are in parentheses).

Culture	Measure	Task	Exposure	Bimodal Congruent	Bimodal Incongruent	Modality Interference
East	ACC	Face	Low	0.85 (0.098)	0.77 (0.16)	0.078 (0.11)
			High	0.87 (0.096)	0.78 (0.12)	0.094 (0.13)
		Voice	Low	0.89 (0.097)	0.79 (0.16)	0.097 (0.10)
			High	0.88 (0.11)	0.78 (0.15)	0.10 (0.11)
	RT	Face	Low	1491 (371)	1596 (373)	104 (298)
			High	1478 (262)	1606 (273)	128 (248)
		Voice	Low	1686 (459)	1801 (406)	115 (308)
			High	1502 (266)	1578 (253)	76 (264)
	IR	Face	Low	4.00 (1.03)	3.71 (1.06)	0.30 (0.50)
			High	4.09 (0.80)	3.76 (0.86)	0.32 (0.54)
		Voice	Low	4.07 (0.76)	3.72 (0.93)	0.35 (0.41)
			High	4.28 (0.78)	4.10 (0.91)	0.18 (0.48)
West	ACC	Face	Low	0.76 (0.10)	0.71 (0.13)	0.050 (0.11)
			High	0.83 (0.080)	0.79 (0.077)	0.039 (0.075)
		Voice	Low	0.77 (0.12)	0.60 (0.15)	0.17 (0.19)
			High	0.78 (0.12)	0.62 (0.17)	0.16 (0.17)
	RT	Face	Low	1451 (298)	1628 (337)	177 (246)
			High	1491 (289)	1562 (336)	71 (283)
		Voice	Low	1927 (494)	1877 (421)	-50 (369)
			High	1624 (289)	1750 (338)	126 (298)
	IR	Face	Low	3.99 (0.97)	3.75 (1.02)	0.24 (0.45)
			High	3.94 (0.88)	3.68 (0.91)	0.26 (0.43)
		Voice	Low	3.78 (0.93)	3.60 (0.85)	0.18 (0.42)
			High	4.09 (0.94)	3.65 (0.99)	0.44 (0.57)

In the intensity ratings analyses, the culture by exposure interaction was significant, $F(1,39) = 4.45$, $p = 0.041$, $\eta_p{}^2 = 0.10$, as was the three-way interaction between task, culture, and exposure, $F(1,39) = 5.10$, $p = 0.030$, $\eta_p{}^2 = 0.12$ (Figure 4b). To breakdown the interaction, separate two-way ANOVAs were conducted for each exposure group. In the bilingual group with low exposure to the U.S. culture, the main effects and interactions were not significant, *p*s > 0.24. In the bilingual group with high exposure to the U.S. culture, the culture by task interaction was significant, $F(1,22) = 5.82$, $p = 0.025$, $\eta_p{}^2 = 0.21$. Specifically, there was a larger modality interference in the voice task when the emotional stimuli were from the West ($M = 0.44$, $SE = 0.12$) than the East ($M = 0.18$, $SE = 0.10$), $F(1,22) = 9.41$, $p = 0.006$, $\eta_p{}^2 = 0.30$, but no differences in modality interference between cultures were found in the face task, $F < 1$.

In the accuracy rates analyses, there was a main effect of task, $F(1,39) = 10.38$, $p = 0.003$, $\eta_p{}^2 = 0.21$, with the voice task ($M = 0.13$, $SE = 0.020$) producing a larger modality interference than the face task ($M = 0.065$, $SE = 0.012$). The task by culture interaction was also significant, $F(1,39) = 8.65$, $p = 0.005$, $\eta_p{}^2 = 0.18$. There was a larger modality interference in the voice task ($M = 0.17$, $SE = 0.028$) than the face task ($M = 0.045$, $SE = 0.014$), but only when the stimuli were from the West, $F(1,39) = 14.88$, $p < 0.001$, $\eta_p{}^2 = 0.28$, and not the East, $F < 1$. No other effects were significant, *F*s < 1.

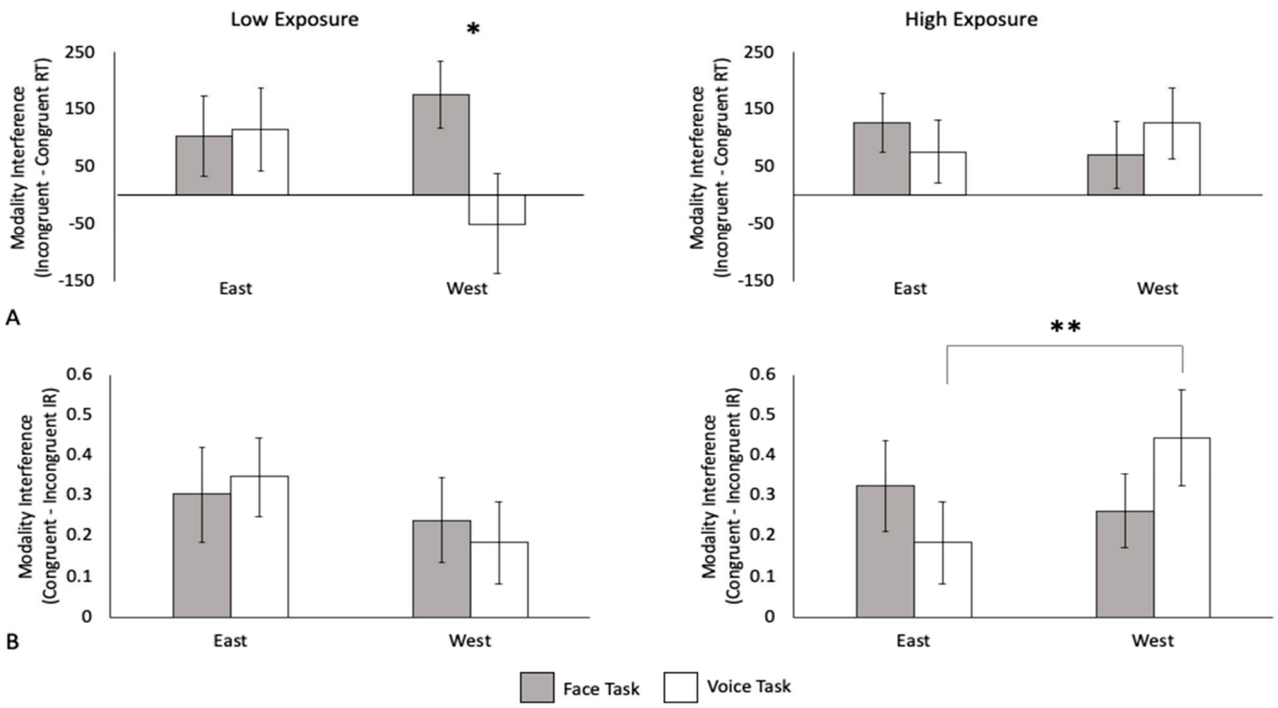

Figure 4. Modality interference effects on (**A**) response times and (**B**) intensity ratings for bilinguals with low (left panels) and high (right panels) levels of exposure to the U.S. culture. The low exposure group showed a larger modality interference in the face task than in the voice task for emotional information from the West (i.e., auditory interference). The high exposure group showed a larger modality interference for Western than Eastern emotional information in the voice task, but no differences in modality interference between cultures in the face task. * $p < 0.05$, ** $p < 0.01$.

4. Discussion

The current study examined whether cultural experience, including length of immersion and amount of daily exposure to a new culture, affects multisensory emotion perception. Overall, bilinguals who migrated to the United States from China experienced greater interference from the visual modality than the auditory modality when asked to judge emotional stimuli from the West. However, when asked to judge emotional stimuli from the East, the bilinguals in the current study experienced similar degrees of interference from both modalities. Interestingly, these patterns were observed in the Chinese-English bilinguals who have been immersed in the American culture for a short duration of time. The differences across modalities and between cultures disappeared in bilinguals with longer immersion experience. Bilinguals with low exposure to the U.S. culture displayed auditory interference, whereas bilinguals with high exposure to the U.S. culture displayed visual interference when presented with Western emotional information. These results reveal that cultural immersion and exposure play important, but distinct, roles in multisensory emotion perception.

The bilinguals in the current study were more impacted by the facial cues than vocal cues when evaluating Western emotional information, replicating previously-observed visual interference patterns exhibited by Westerners (Liu et al. 2015b; Tanaka et al. 2010). When evaluating Eastern emotional information, the bilinguals were equally distracted by the auditory and visual modalities, replicating Liu et al.'s behavioral findings (2015b). At first glance, these findings appear to support the Cultural Frame Switching (CFS) hypothesis. However, this overall modality interference pattern is more similar to that of the bilinguals with shorter immersion experience than that of the bilinguals with longer immersion experience, making the CFS interpretation less plausible. In our findings, a noticeable pattern was that bilinguals with shorter immersion experience relied heavily on the visual modality when presented with emotional information from the West. As immersion experience increased, modality interference across cultures appeared to converge, suggesting that bilinguals with longer immersion experience may merge the two cultures into one. If the CFS hypothesis was supported, bilinguals with longer immersion would have shown visual interference when perceiving emotional information from the West.

Furthermore, a closer look at Figure 3 reveals that the difference between bilinguals with shorter and longer immersion was in the condition where participants were asked to judge English speech, $t(39) = 2.10$, $p = 0.042$. There was a positive correlation between length of immersion in the U.S. and accuracy in judging English speech, showing that accuracy increased as the length of immersion increased, $r(41) = 0.31$, $p = 0.046$. This positive correlation suggests that greater interference from Caucasian faces when judging the English speech could be due to limited English familiarity in bilinguals with shorter immersion experience. This limited familiarity in the auditory modality might cause the bilinguals to over-compensate by excessively relying on the visual modality. In fact, previous studies have shown that during multisensory emotion processing, individuals tend to rely more on the modality they are most familiar with (Chen 2019; Collignon et al. 2008). Therefore, it is likely that the differences in modality interference between Eastern and Western cultures among bilinguals with shorter immersion experience was driven by familiarity rather than switching between different cultural contexts. Future research is needed to disentangle the relative contribution of familiarity from the length of immersion on multisensory emotion perception in individuals who migrate to a new country and are in the process of adopting new cultural norms.

In contrast to the findings by Liu et al. (2017), we did not observe visual interference in the condition where participants were presented with Asian facial expressions and Mandarin speech. The discrepancy between our findings and theirs may be due to the sample under investigation. The participants in Liu et al. (2017) had spent relatively equal time in Eastern and Western countries, whereas the participants in the current study spent more time in an Eastern country than a Western country. The participants in our study may have had less experience with Western cultures than the participants in Liu et al. (2017). Another possibility is that the participants in Liu et al.'s study became less familiar with Mandarin, as they were using English for 68–81 hours per week on average. Previous work has shown that perceivers increase their reliance on the visual modality when multisensory emotional input is less familiar (Chen 2019). Therefore, it is possible that the bilinguals in Liu et al.'s study may have increased their reliance on the visual modality (i.e., Asian face) because the auditory input was less familiar.

Compared with length of immersion, daily exposure to the new culture had a different effect on emotion perception. Response times revealed that bilinguals with low daily exposure to the U.S. culture relied more on the auditory modality when evaluating Western emotional information, but did not rely more on the auditory modality when evaluating Eastern emotional information. A possible explanation of this puzzling finding is that participants consistently relied on the more-familiar modality when evaluating unfamiliar emotional input from the West. However, when evaluating familiar emotional input from the East, modality reliance may have varied depending on ties to the old culture, willingness to assimilate into the new culture, and duration of immersion, with potentially opposite direction of effects across participants. For example, participants who feel strongly about preserving their Eastern culture may be more likely to continue relying on the auditory modality, whereas participants seeking to integrate into the new culture may be less likely to rely on the auditory modality. The opposite patterns may be cancelling each other out, resulting in no overall modality interference for Eastern emotional information in the low exposure participants.

For the individuals with high daily exposure to the U.S. culture, the pattern of intensity ratings showed that they might be adjusting their reliance on the visual modality depending on the context of the interaction, which is consistent with the Cultural Frame Switching hypothesis (Hong et al. 2000; LaFromboise et al. 1993). Specifically, we observed that there was increased interference from the visual modality when participants were presented with emotional information from the West compared to when they were presented with information from the East. Therefore, daily exposure to the new culture seems to change our perceptions to reflect new cultural norms, consistent with the findings from the emotional acculturation literature (Cosedine et al. 2014; De Leersnyder et al. 2011).

The differences between the low and high daily exposure groups may be due to each group's experience with and degree of confidence in recognizing the emotions of Western individuals, as well as to distinctions between what is being assessed by RT versus intensity ratings. On the emotion recognition task, RTs were measured as the amount of time required to select the perceived emotion from five alternatives. The longer RTs in the low exposure group may be indicative of greater uncertainty discriminating between multiple emotions (Arndt et al. 2018; Lischetzke et al. 2011). Intensity ratings, on the other hand, required participants to compare how intense a perceived emotion was relative to input with no emotion and may be capturing a combined cognitive appraisal of the subjective experience (i.e., inferences about the mental state; Matsumoto 1999) and the valence of the emotion (e.g., pleasant or unpleasant; Russell 1980; Sutton et al. 2019). Because response times and intensity measures are potentially tapping into different cognitive phenomena and may have different degrees of sensitivity when measuring modality interference, more research is needed to understand how various measures tap into different aspects of emotional processing.

It remains unclear how much daily exposure is necessary for the Cultural Frame Switching (CFS) to emerge in emotion perception. Does the CFS pattern emerge predominantly amongst those who divide their time more evenly across cultures or do individuals have to immerse themselves completely into the new culture to reduce the effects from the old culture? A possible future direction would be to test participants who grew up with both cultures and have similar levels of exposure to each culture. Another consideration related to testing the CFS hypothesis is to separate the influence of cultural immersion from familiarity. For Chinese-English bilinguals in the current study, the less familiar auditory input (i.e., English) resulted in increased visual interference. However, this increased visual interference is also predicted by having increased Western cultural exposure, making it hard to separate the effect of familiarity from the effect of cultural immersion. An alternative would be to test English speakers moving from a Western culture to China. According to the familiarity account, the less familiar input (i.e., Mandarin speech) would lead to increased visual interference in the new Eastern culture. The CFS hypothesis, on the other hand, would predict an increased auditory interference when processing emotional input from the East.

The current study has several limitations that will need to be addressed in future research. First, the linguistic environment may have impacted the Cultural Frame Switching hypothesis. The task instructions and language of testing were in English, which may have increased Western-culture-specific behaviors. Second, the absence of a monolingual and monocultural control group limits the interpretation of results. In an attempt to address this limitation in future research, we are currently developing a study that will focus on emotion recognition in monolingual native English speakers in the United States and monolingual native Mandarin speakers in China (it is difficult to ascertain the monocultural identity of even monolingual participants, however, because people can identify with multiple cultures despite living in the same country or speaking one language; indeed, the definition of culture is open to debate). Lastly, the face stimuli used in the current study were static images of faces rather than dynamic facial expressions. Moving faces have been shown to improve facial and emotional recognition and lead to more intense responses than static faces (see Alves 2013; Xiao et al. 2014). In addition, contextual information, such as the visual scene, description of the event, and body language, can also serve as important cues for detecting emotions (Aviezer et al. 2017; Barrett et al. 2011). Future research should consider using audio and video recordings to study emotion processing of dynamic facial expressions in variable contexts.

In conclusion, immersion experience and daily exposure to a new culture impacted multisensory emotion perception in different ways. As length of immersion increased, Chinese-English bilinguals demonstrated a similar pattern in modality interference across cultures. This finding suggests that bilinguals who move from an Eastern to a Western culture might merge the norms from their two cultures and demonstrate similar modality interference when perceiving emotional input from both cultures. Furthermore, an individual's exposure to the cultural and social norms of the new culture predicts the likelihood of exhibiting a modality interference pattern that is more common in the new culture. We conclude that length of immersion in a culture and amount of cultural exposure interact to shape perceptual and cognitive processes, such as emotion perception.

Author Contributions: Conceptualization, P.C. and V.M.; Methodology, P.C. and V.M.; Data Collection, P.C.; Formal Analysis, P.C. and A.C.-F.-Y.; Writing—Original Draft Preparation, P.C. and A.C.-F.-Y.; Writing—Review & Editing, P.C., A.C.-F.-Y., and V.M.; Supervision, V.M.; Project Administration, V.M.; Funding Acquisition, V.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health, grant number R01HD059858, to Viorica Marian. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Northwestern University (STU00023477; ethics approval received on 28 April 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available because public sharing of the data was not approved by the Institutional Review Board and because participants did not provide permission to have their data shared publicly at the time of testing.

Acknowledgments: The authors thank Marc Pell for providing the vocal emotion stimuli as well as the members of the Bilingualism and Psycholinguistics Research Group for helpful comments and input on this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

Alves, Nelson Torro. 2013. Recognition of Static and Dynamic Facial Expressions: A Study Review. *Estudos de Psicologia* 18: 125–30. [CrossRef]

Argyle, Michael, Monika Henderson, Michael Bond, Yuichi Iizuka, and Alberta Contarello. 1986. Cross-Cultural Variations in Relationship Rules. *International Journal of Psychology* 21: 287–315. [CrossRef]

Arndt, Charlotte, Tanja Lischetzke, Claudia Crayen, and Michael Eid. 2018. The Assessment of Emotional Clarity via Response Times to Emotion Items: Shedding Light on the Response Process and its Relation to Emotion Regulation Strategies. *Cognition and Emotion* 32: 530–48. [CrossRef]

Athanasopoulos, Panos, Benjamin Dering, Alison Wiggett, Jan Rouke Kuipers, and Guillaume Thierry. 2010. Perceptual Shift in Bilingualism: Brain Potentials Reveal Plasticity in Pre-Attentive Colour Perception. *Cognition* 116: 437–43. [CrossRef]

Aviezer, Hillel, Noga Ensenberg, and Ran R Hassin. 2017. The Inherently Contextualized Nature of Facial Emotion Perception. *Current Opinion in Psychology* 17: 47–54. [CrossRef]

Barrett, Lisa Feldman, Batja Mesquita, and Maria Gendron. 2011. Context in Emotion Perception. *Current Directions in Psychological Science* 20: 286–90. [CrossRef]

Benet-Martínez, Veronica, Janxin Leu, Fiona Lee, and Michael W Morris. 2002. Negotiating Biculturalism. *Journal of Cross-Cultural Psychology* 33: 492–516. [CrossRef]

Berry, John W. 1997. Lead Article Immigration, Acculturation, and Adaptation. *Applied Psychology: An International Review* 46: 5–68. [CrossRef]

Chen, Peiyao. 2019. The Influence of Language and Culture on the Multisensory Perception of Emotion. Ph.D. dissertation, Northwestern University, Evanston, IL, USA.

Chen, Li-Fen, and Yu-Shiuan Yen. 2007. Taiwanese Facial Expression Image Database. Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan. Available online: https://urldefense.com/v3/__https://scholar.google.com/scholar_lookup?title=Taiwanese*Facial*Expression*Image*Database.&author=L*Chen&author=Y*Yen&publication_year=2007&__;KysrKysr!!Dq0X2DkFhyF93HkjWTBQKhk!GJ9yzlfMcnaXqVcHisS2_mo6CI4RcYRHu8rWYKOo_kpSCrXQUw3_b3Gx7Ga1FObmWOIeByKvVHJYDns$ (accessed on 28 April 2017).

Cheng, Chi-Ying, Fiona Lee, and Veronica Benet. 2006. Assimilation and Contrast Effects in Cultural Frame Switching Bicultural Identity Integration and Valence of Cultural Cues. *Journal of Cross-Cultural Psychology* 37: 742–60. [CrossRef]

Collignon, Olivier, Simon Girard, Frederic Gosselin, Sylvain Roy, Dave Saint-Amour, Maryse Lassonde, and Franco Lepore. 2008. Audio-Visual Integration of Emotion Expression. *Brain Research* 1242: 126–35. [CrossRef]

Cook, Vivian, Benedetta Bassetti, Chise Kasai, Miho Sasaki, and Jun Arata Takahashi. 2006. Do Bilinguals Have Different Concepts? *The Case of Shape and Material in Japanese L2 Users of English. International Journal of Bilingualism* 10: 137–52. [CrossRef]

Cosedine, Nathan, Yulia E. Chentsova-Dutton, and Yulia S. Krivoshekova. 2014. Emotional Acculturation Predicts Better Somatic Health: Experiential and Expressive Acculturation among Immigrant Women from Four Ethnic Groups. *Journal of Social and Clinical Psychology* 33: 867–89. [CrossRef]

De Leersnyder, Jozefien, Batja Mesquita, and Heejung S. Kim. 2011. Where Do My Emotions Belong? A Study of Immigrants' Emotional Acculturation. *Personality and Social Psychology Bulletin* 37: 451–63. [CrossRef]

De Leersnyder, Jozefien, Batja Mesquita, and Heejung S. Kim. 2013. Emotional Acculturation. In *Changing Emotions*, 1st ed. Edited by Dick Hermans, Bernard Rimé and Batja Mesquita. London: Psychology Press, pp. 127–33.

De Leersnyder, Jozefien, Heejung S. Kim, and Batja Mesquita. 2020. My Emotions Belong Here and There: Extending the Phenomenon of Emotional Acculturation to Heritage Culture Fit. *Cognition and Emotion* 34: 1573–90. [CrossRef] [PubMed]

Derntl, Birgit, Ute Habel, Simon Robinson, Christian Windischberger, Ilse Kryspin-Exner, Ruben C. Gur, and Ewald Moser. 2009. Amygdala Activation during Recognition of Emotions in a Foreign Ethnic Group Is Associated with Duration of Stay. *Social Neuroscience* 4: 294–307. [CrossRef]

Derntl, Birgit, Ute Habel, Simon Robinson, Christian Windischberger, Ilse Kryspin-Exner, Ruben C. Gur, and Ewald Moser. 2012. Culture but Not Gender Modulates Amygdala Activation during Explicit Emotion Recognition. *BMC Neuroscience* 13: 54. [CrossRef] [PubMed]

Ekman, Paul, and Wallace V. Friesen. 1971. Constants across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology* 17: 124–29. [CrossRef]

GIMP Development Team. 2015. GIMP. Available online: https://www.gimp.org (accessed on 23 May 2017).

Hong, Yingyi, Michael W. Morris, Chi Yue Chiu, and Veronica Benet-Martínez. 2000. Multicultural Minds: A Dynamic Constructivist Approach to Culture and Cognition. *The American Psychologist* 55: 709–20. [CrossRef]

Hong, Ying-Yi, Veronica Benet-Martínez, Chi-Yue Chiu, and Michael W. Morris. 2003. Boundaries of Cultural Influence: Construct Activation as a Mechanism for Cultural Differences in Social Perception. *Journal of Cross-Cultural Psychology* 34: 453–64. [CrossRef]

Kitayama, Shinobu, Sean Duffy, Tadashi Kawamura, and Jeff T. Larsen. 2003. Perceiving an Object and Its Context in Different Cultures: A Cultural Look at New Look. *Psychological Science* 14: 201–6. [CrossRef] [PubMed]

LaFromboise, Terasa, Hardin L. K. Coleman, and Jennifer Gerton. 1993. Psychological Impact of Biculturalism: Evidence and Theory. *Psychological Bulletin* 114: 395–412. [CrossRef] [PubMed]

Lischetzke, Tanja, Rozalina Angelova, and Michael Eid. 2011. Validating an Indirect Measure of Clarity of Feelings: Evidence from Laboratory and Naturalistic Settings. *Psychological Assessment* 23: 447–55. [CrossRef] [PubMed]

Liu, Pan, and Marc D. Pell. 2012. Recognizing Vocal Emotions in Mandarin Chinese: A Validated Database of Chinese Vocal Emotional Stimuli. *Behavior Research Methods* 44: 1042–51. [CrossRef]

Liu, Pan, Simon Rigoulot, and Marc D. Pell. 2015a. Culture Modulates the Brain Response to Human Expressions of Emotion: Electrophysiological Evidence. *Neuropsychologia* 67: 1–13. [CrossRef]

Liu, Pan, Simon Rigoulot, and Marc D. Pell. 2015b. Cultural Differences in On-Line Sensitivity to Emotional Voices: Comparing East and West. *Frontiers in Human Neuroscience* 9: 311. [CrossRef]

Liu, Pan, Simon Rigoulot, and Marc D. Pell. 2017. Cultural Immersion Alters Emotion Perception: Neurophysiological Evidence from Chinese Immigrants to Canada. *Social Neuroscience* 12: 685–700. [CrossRef]

Lundqvist, Daniel, Anders Flykt, and Arne Öhman. 1998. *The Karolinska Directed Emotional Faces—KDEF*. CD ROM from Department of Clinical Neuroscience. Solna: Psychology Section, Karolinska Institute, ISBN 91-630-7164-9.

Marian, Viorica. Forthcoming. *The Power of Language*. New York: Dutton, Penguin Random House, ISBN 9780593187074.

Marian, Viorica, and Margarita Kaushanskaya. 2004. Self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language* 51: 190–201. [CrossRef]

Marian, Viorica, Henrike K. Blumenfeld, and Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech Language and Hearing Research* 50: 940–67. [CrossRef]

Markus, Hazel Rose, and Shinobu Kitayama. 1991. Culture and the Self: Implications for Cognition, Emotion, and Motivation. *Psychological Review* 98: 224–53. [CrossRef]

Matsumoto, David. 1999. American-Japanese Cultural Differences in Judgements of Expression Intensity and Subjective Experience. *Cognition and Emotion* 13: 201–18. [CrossRef]

Matsumoto, David, Sachiko Takeuchi, Sari Andayani, Natalia Kouznetsova, and Deborah Krupp. 1998. The Contribution of Individualism-Collectivism to Cross-National Differences in Display Rules. *Asian Journal of Social Psychology* 1: 147–65. [CrossRef]
McCarthy, Anjanie, Kang Lee, Shoji Itakura, and Darwin W. Muir. 2006. Cultural Display Rules Drive Eye Gaze during Thinking. *Journal of Cross-Cultural Psychology* 37: 717–22. [CrossRef]
McCarthy, Anjanie, Kang Lee, Shoji Itakura, and Darwin W. Muir. 2008. Gaze Display When Thinking Depends on Culture and Context. *Journal of Cross-Cultural Psychology* 39: 716–29. [CrossRef]
Morling, Beth, Shinobu Kitayama, and Yuri Miyamoto. 2002. Cultural Practices Emphasize Influence in the United States and Adjustment in Japan. *Personality and Social Psychology Bulletin* 28: 311–23. [CrossRef]
Morris, Michael W., and Aurelia Mok. 2011. Isolating Effects of Cultural Schemas: Cultural Priming Shifts Asian-Americans' Biases in Social Description and Memory. *Journal of Experimental Social Psychology* 47: 117–26. [CrossRef]
Panayiotou, Alexia. 2004. Switching Codes, Switching Code: Bilinguals' Emotional Responses in English and Greek. *Journal of Multilingual and Multicultural Development* 25: 124–39. [CrossRef]
Pell, Marc D., Silke Paulmann, Chinar Dara, Areej Alasseri, and Sonja A. Kotz. 2009. Factors in the Recognition of Vocally Expressed Emotions: A Comparison of Four Languages. *Journal of Phonetics* 37: 417–35. [CrossRef]
Perunovic, Elaine Wei Qi, Daniel Heller, and Eshkol Rafaeli. 2007. Within-Person Changes in the Structure of Emotion: The Role of Cultural Identification and Language. *Psychological Science* 18: 607–13. [CrossRef]
Ramírez-Esparza, Nairán, Samuel D. Gosling, Verónica Benet-Martínez, Jeffrey P. Potter, and James W. Pennebaker. 2006. Do Bilinguals Have Two Personalities? A Special Case of Cultural Frame Switching. *Journal of Research in Personality* 40: 99–120. [CrossRef]
Ross, Michael, W. Q, Elaine Xun, and Anne E. Wilson. 2002. Language and the Bicultural Self. *Personality and Social Psychology Bulletin* 28: 1040–50. [CrossRef]
Russell, James A. 1980. A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39: 1161–78. [CrossRef]
Sutton, Tina M., Andrew M. Herbert, and Dailyn Q. Clark. 2019. Valence, Arousal, and Dominance Ratings for Facial Stimuli. *Quarterly Journal of Experimental Psychology* 72: 2046–55. [CrossRef]
Tanaka, Akihiro, Ai Koizumi, Hisato Imai, Saori Hiramatsu, Eriko Hiramoto, and Beatrice De Gelder. 2010. I Feel Your Voice: Cultural Differences in the Multisensory Perception of Emotion. *Psychological Science* 21: 1259–62. [CrossRef] [PubMed]
Weisz, John R., Fred M. Rothbaum, and Thomas C. Blackburn. 1984. Standing out and Standing in: The Psychology of Control in America and Japan. *American Psychologist* 39: 955–69. [CrossRef]
West, Alexandria L., Rui Zhang, Maya A. Yampolsky, and Joni Y. Sasaki. 2018. The Potential Cost of Cultural Fit: Frame Switching Undermines Perceptions of Authenticity in Western Contexts. *Frontiers in Psychology* 9: 2622. [CrossRef]
Xiao, Naiqi G., Steve Perotte, Paul C. Quinn, Zhe Wang, Yu-Hao P. Sun, and Kang Lee. 2014. On the Facilitative Effects of Face Motion on Face Recognition and its Development. *Frontiers in Psychology* 5: 633. [CrossRef]

languages MDPI

Article

Being a Student or at Home: Does Topic Influence How Bilinguals Process Words in Each Language?

Veniamin Shiron [1], Huanhuan Liu [2] and Angela de Bruin [1,*]

[1] Department of Psychology, University of York, York YO10 5DD, UK; vs814@york.ac.uk
[2] Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian 116029, China; 104130@lnnu.edu.cn
* Correspondence: angela.debruin@york.ac.uk

Abstract: Research has assessed how language use differences between bilinguals (e.g., whether two languages are used approximately equally often or not) influence language processing. However, first (L1) and second (L2) language use might also differ within bilinguals, depending on the topic of conversation. For example, a Mandarin–English bilingual studying in North America or the UK might talk about exams in English but about their childhood in Mandarin. In this study, we therefore examined how topics associated with either the L1 or L2 can influence language processing. Twenty-nine Mandarin–English students in North America/the UK completed a lexical decision task in single-language contexts (all words/pseudowords in one language) and in dual-language contexts (alternating between Mandarin and English). Half of the words referred to L1-associated topics (childhood and family life) and half were L2-associated (studying and life at university). Topic influenced L2 processing, with L2-associated topics being processed faster than topics associated with the L1 in single- and dual-language contexts. In contrast, topic did not influence L1 processing. This suggests that L2 processing might not only be influenced by differences between bilinguals but also by differences within bilinguals. In contrast, L1 processing might be less susceptible to influences of topic-specific language use.

Keywords: bilingualism; language processing; language context; language use

Citation: Shiron, Veniamin, Huanhuan Liu, and Angela de Bruin. 2021. Being a Student or at Home: Does Topic Influence How Bilinguals Process Words in Each Language? *Languages* 6: 150. https://doi.org/10.3390/languages6030150

Academic Editor: John W. Schwieter

Received: 22 July 2021
Accepted: 2 September 2021
Published: 9 September 2021

1. Introduction

In our international society, many people speak more than one language, including a large number of people who are studying and working in a language they did not grow up with. These many bilinguals differ from each other in various ways, including in their proficiency (i.e., how competent they are in each language) and their language use (i.e., how often or in which contexts they use each of their languages). Various research studies have examined how these differences *between bilinguals* in terms of overall proficiency and use in their second language (L2) can influence first language (L1) and L2 processing (e.g., Blumenfeld and Marian 2007; Chaouch-Orozco et al. 2021; Van Hell and Tanner 2012).

However, language proficiency and use do not just differ between bilinguals but also *within* bilinguals. As Grosjean's Complementarity Principle (Grosjean 2016) outlines, bilinguals can acquire and use their languages for different purposes, domains, and people. For example, an international student in the UK or North America might mainly acquire English as their second language for purposes related to studying. As a consequence, their proficiency and use might be much higher when considering words related to studying than when assessing other topics (cf. Brysbaert et al. 2021). In contrast, their native language might mainly be used when talking about topics associated with their life in their country of origin. Previous research assessing more global differences between bilinguals (asking, e.g., how often a bilingual uses a language in their daily lives) overlooks these more domain- or topic-specific differences *within* bilinguals. Studying the relationship between topic-specific language use and language processing within the same group of bilinguals is crucial to better understand how bilinguals comprehend and process words in their native and non-native language. In this study, we examined if and how topic-specific words associated with either the participant's L1 or L2 influence bilingual processing. We focused on Mandarin–English bilinguals who grew up speaking Mandarin but are studying in their L2 (English) in North America or the UK. As a measure of "topic-specific" language use, we focused on words referring to family life/childhood at home (L1) or to life as a student (L2) (see also Brysbaert et al. 2021).

1.1. Cues Associated with a Language or Culture

Previous research has shown that cues associated with a language or culture can influence both bilingual comprehension and production. For example, Zhang et al. (2013) looked at how L1 culture cues (images and faces) influenced language production in Chinese students living in the US. They found that cues associated with the L1 disrupted L2 fluency, suggesting that bilingual processing is flexible and can differ within bilinguals depending on the surrounding context being associated with a certain language or culture. Similar findings were shown by Roychoudhuri et al. (2016). Bengali–English participants were slower to name pictures in English (L2) when they were presented with an image typically associated with Bengali (L1) culture. Increased activation of the L1, as a consequence of L1-associated cues, potentially led to increased interference with the L2 and, consequently, slower L2 naming. Similar findings have been observed in studies using faces as cues (e.g., Molnar et al. 2015; Woumans et al. 2015).

Beyond contextual cues, Jared et al. (2013) showed that the type of object can influence L1 and L2 naming. Chinese–English bilinguals were asked to name images associated with a specific country (e.g., a variety of cabbage more typically eaten in China or in Canada). Participants' responses were faster in Chinese than English when naming a "Chinese" cabbage, while the opposite was found when naming a "Canadian" cabbage. These results suggest that bilinguals associate newly acquired words with specific concepts that are closely associated with a given culture (cf. Berkes et al. 2018, for similar findings in comprehension). This reinforces the idea underlying Grosjean's Complementarity Principle, stating that bilingual proficiency and use might differ between bilinguals depending on the domain (in this case, objects associated with a specific culture).

These previous studies, however, have focused on non-linguistic cues (e.g., Zhang et al. 2013) or concepts that are uniquely associated with one culture/language (e.g., Jared et al. 2013). An open question remains how bilinguals process words that do exist in both languages/cultures but that are more closely associated with one language due to the bilingual's personal experiences (e.g., using one language at home and one language for studying). Although this is a situation in which many bilinguals find themselves, little is known about how topic-dependent associations influence L1 and L2 word processing.

1.2. Language Interference and Control

In addition, from the previous literature, it is unclear whether any topic-dependent associations should influence the L2 only or both languages. In some studies, only one language is examined (e.g., Zhang et al. 2013), while studies that do directly compare both languages show mixed results. The L2 (when acquired later in life and when having a lower proficiency) might be more easily influenced than the L1 that was acquired from birth, potentially because L1 representations are more strongly entrenched in memory (e.g., Diependaele et al. 2013). For example, previous research has suggested that frequency effects, such as faster processing of high-frequency words, are more pronounced in an L2 than L1 (e.g., Van Wijnendaele and Brysbaert 2002). Roychoudhuri et al.'s (2016) findings are, to some extent, in line with this. Cues had a significantly larger influence on L2 than on L1 naming, with the cue-naming effect only being significant when using the L2. However, this study included only L1 cues, making it difficult to directly compare the two conditions. Furthermore, while the effect was not significant, L1 cues did numerically facilitate L1 naming. Other studies do find effects of language-/culture-specific cues on both L1 and

L2 naming or processing (e.g., Berkes et al. 2018; Jared et al. 2013; Woumans et al. 2015). However, when using faces or objects that are unique to one language/culture (such as cabbages associated with Western or Asian cultures), having to generate a word incompatible with the face or object (such as the Chinese word for a Canadian cabbage) might be so rare that it might cause a mismatch effect regardless of the language. It therefore remains unclear whether words that actually exist in both languages can influence L1 as well as L2 processing, depending on associations between the topic and the language.

Furthermore, it remains unclear (cf. Hartsuiker 2015) whether any effects of language or culture associations are due to matching associations facilitating performance (e.g., faster L2 processing because L2 words are more easily retrieved in combination with an L2 cue or object, cf. Li et al. 2013) or due to associations with the "other" language interfering with performance (e.g., slower L2 processing because L1 cues/objects activate L1 words that consequently interfere more strongly with L2 processing, cf. Zhang et al. 2013). In our study, we therefore included single- and dual-language contexts to manipulate the amount of language interference. In single-language conditions, words were presented in one language, while in dual-language conditions, words alternated between the two languages. Previous research has suggested that language co-activation, and consequently, interference and competition, is larger in dual- than single-language conditions, including in lexical decision tasks (e.g., Vanlangendonck et al. 2020). If topic associations influence language processing by facilitating word retrieval, we would expect these facilitatory effects to occur in both single- and dual-language conditions. However, if topic-associations influence language processing by (also) increasing interference from the "other" language, we would expect these effects to be larger in the dual-language condition (where the other language is present within the same context). Roychoudhuri et al. (2016) observed similar effects of cultural cues in single- and dual-language contexts. However, their English "single-language" contexts included Bengali cues, making it difficult to create truly "English-only" contexts. By using words without any cues, we were able to create pure single-language contexts without any direct task-induced interference from the other language.

1.3. Current Study

We examined whether topics associated with either the L1 or L2 can influence language processing in single- and dual-language contexts. We focused on Mandarin–English bilinguals who grew up in a Mandarin-language environment but are currently studying in English in the UK or North America. Contrary to previous studies (e.g., Jared et al. 2013) that focused on concepts uniquely associated with one culture/language, we used words with similar objective frequency scores in both languages. All words were highly frequent words that occur in both Mandarin and English. However, we selected words that *these bilinguals* more closely associated with L1 (family life) or L2 (studying).

Participants completed a lexical decision task in single-language contexts (all words presented in one language) and in dual-language contexts (words presented in Mandarin and English interchangeably). Our main question examined whether there were reaction time (RT) differences between words representing a topic associated with that language (e.g., words related to studying in English) versus topics that are associated with the other language (e.g., words related to family life in English). We hypothesised that both L1 and L2 should show faster processing of words associated with that language than of words associated with the other language. This would show whether language processing in L1 and L2 is stable and topic-independent or rather flexible and topic-specific. In addition, we also examined differences between single- and dual-language contexts. The interchangeable use of two languages in the dual-language context might increase cross-language interference. If this interference increases the effects of topic, these effects should be larger in the dual- than single-language context.

2. Methods

2.1. Participants

Thirty Mandarin–English bilinguals were recruited through Prolific.co. One participant was excluded from further analysis as their accuracy was below chance. The remaining 29 participants (20 female) were, on average, 23 years old (*SD* = 2.9). All participants had normal or corrected-to-normal vision, and no known neurological, hearing, or reading impairments. All apart from one were right-handed.

Given that it was difficult to determine a realistic effect size (which largely depends on the stimulus materials created for this study) and run a power analysis, we based our sample size calculations on the recommendation to use 1600 observations per cell (e.g., 40 trials × 40 participants) for linear mixed-effects analyses (Brysbaert and Stevens 2018). In combination with 48 trials per condition, we aimed to recruit 32 participants. Due to difficulties recruiting more participants as a consequence of our strict language profiles, the final sample size of 29 is a little below our aim but is, in combination with our 48 trials per condition, still close to the recommended number of 1600 observations.

All participants grew up in a Mandarin-speaking household and spent their primary-school years in China, with limited exposure to other languages and were, at the moment of testing, living and studying in the UK (*N* = 19) or North America (*N* = 10). On average, participants had spent 4.4 years in an English-speaking country (*SD* = 3.3). Participants completed an English picture-naming task to measure vocabulary (De Bruin et al. 2017) and a language-background questionnaire based on (Anderson et al. 2018; see Tables 1 and 2). This confirmed that participants used English more often for education-related topics and Mandarin for topics related to family life (see Table 1). Participants generally used Mandarin more during childhood and to communicate with family members, while they used English more with roommates and classmates (Table 2).

Table 1. Means (and SDs) in the English picture-naming task, proficiency self-ratings, language use per topic, and overall use/exposure.

Item	English Mean	SD	Mandarin Mean	SD
Age of Acquisition	7.9	3.5	0	0
Picture naming (0–65) [a]	59.0	3.1	X	X
Self-rated proficiency (0–10)				
Speaking	6.6	1.4	9.4	1.0
Understanding	7.7	1.3	9.6	0.7
Writing	6.9	1.4	8.8	1.4
Reading	7.7	1.5	9.6	0.6
Language use per topic: (0–100)				
Family life/childhood	31.7	17.5	69.7	20.5
Studying	56.7	27.1	44.7	27.0
Overall exposure [b] (0–100)	68.9	25.5	56.7	23.9
Overall use [b] (0–100)	60.5	23.6	56.9	26.1

[a] Data from one participant could not be included because they responded in Mandarin. [b] For these questions, participants were asked to only consider their time spent in the UK/North America (i.e., their current/recent exposure and use). Participants responded on a scale from 0 (“never”) to 100 (“all the time”), without those percentages necessarily adding up to 100% for the two languages together.

Table 2. Summary of self-rated language use in different contexts and tasks. Higher values (1–5) indicate more Mandarin use.

Item	Mean	SD
Language exposure during childhood: (1–5) [a]		
Infancy	4.9	0.3
Preschool age	4.6	0.7
Primary school	4.0	0.4
High school	3.6	0.7
Language use by activity [c]: (1–5) [a]		
Reading	2.3	0.9
Emailing	1.4	0.6
Texting	2.5	0.9
Social media	2.4	0.9
Watching TV	2.5	1.0
Listening to music	2.4	0.9
Watching movies	2.2	0.8
Language use with different people [c]: (1–5) [a]		
Roommates	2.9	1.3
Classmates	2.4	1.1
Friends	3.0	1.1
Parents	4.9	0.3
Grandparents	5.0	0.2
Other relatives	4.7	0.6
Siblings	4.7	0.6
Partner	3.7	1.3
Language switching per context [c]: (1–5) [b]		
On a daily basis	3.3	0.9
In a conversation	3.0	0.9
In a sentence	2.6	0.9
With family	1.6	0.9
With classmates	2.8	1.2
Talking about family life/childhood	2.1	1.0
Talking about studying/student life	3.1	1.2

[a] Scoring system: 1 (all English)–5 (only Mandarin). [b] Scoring system: 1 (Never)–5 (Always). [c] Participants rated their use/switching while being in the UK/North America (i.e., recent/current use and switching).

2.2. Materials

We created two sets of 24 words each. One set focused on words associated with family life and childhood (L1-association). The other set focused on life as a student (L2-association). At the end of the study, our bilingual participants rated daily life language use for each stimulus (0 "always English" to 100% "always Mandarin"). Participants indicated lower Mandarin (L1) use for words in the L2-association condition (M = 39.3%, SD = 13.8) than for words in the L1-association condition (M = 58.3, SD = 19.3; $t(28) = -6.791$, $p < 0.001$).

Conditions were matched on frequency, valence, and word length (see Appendix A). Within each language, we matched the two conditions (words associated with L1 and L2) on frequency and length (number of letters, phonemes, and syllables for English, and number of characters for Mandarin). Using ZIPF scores[1] to compare frequency across databases (Cai and Brysbaert 2010; Van Heuven et al. 2014), we also matched the English and Mandarin words on objective frequency. While we initially considered creating neutral words too, we decided against this as it is very difficult to find words that are truly neutral in their language association across a group of participants.

We also created a pseudoword for each word using Wuggy (Keuleers and Brysbaert 2010) by changing, on average, half of the letters of the English word and one character per Mandarin word to create a non-existing combination.

While we carefully matched our stimuli on various background characteristics, we wanted to make sure that our stimulus conditions would not differ in lexical processing times when processed by monolinguals. We therefore ran a pilot study with 15 English monolingual native speakers (five additional participants were excluded due to low performance or not meeting background criteria) and 24 Mandarin native speakers (with limited English use) to assess whether there were any baseline differences between the L1- and L2-associated stimuli. Neither accuracy nor RTs showed a significant difference between conditions in either English or Mandarin native speakers (see Table 3).

Table 3. Mean (and SDs) for accuracy and RTs in the English native and Mandarin native control groups (pseudowords not included in the analysis).

Measure	English/L2-Topic	Mandarin/L1-Topic	Paired T-Test
English native			
Accuracy	97.9 (2.8)	96.1 (4.6)	t(14) = 1.233, p = 0.238
RTs	594.5 (73.1)	601.8 (78.3)	t(14) = −0.844, p = 0.413
Mandarin native			
Accuracy	95.9 (5.8)	96.0 (5.1)	t(23) = −0.100, p = 0.921
RTs	694.6 (100.0)	689.6 (107.6)	t(23) = 0.824, p = 0.418

2.3. Tasks and Procedure

The study was run online in one session using Gorilla (www.gorilla.sc, Anwyl-Irvine et al. 2020) and lasted 30–45 min. Participants completed a lexical decision task in which they had to press a button indicating whether a letter/character string was an existing word or not. The task included three within-subject variables: Language (Mandarin/English); Condition (L1-associated Topic/L2-associated Topic); Context (Single-/Dual-language). The task followed the order single-language/dual-language/single-language (a frequently used design in studies using single- and dual-task contexts, e.g., Rubin and Meiran 2005). In the single-language contexts, all stimuli were presented in either English or Mandarin, with the order of languages counterbalanced across participants. Instructions were provided in the language of the context. Prior to the first single-language task, participants completed a practice lexical decision task on each word/pseudoword in each language. This way, we aimed to minimise item-repetition or task-practice effects by making sure that participants had already responded to each word once before the task started. In the dual-language task, participants interchangeably saw English and Mandarin words and pseudowords. Instructions were given in both languages. Participants were asked to indicate whether the stimulus was a word or not regardless of the language. This task was preceded by eight practice trials.

The single-language context included 24 words and pseudowords per condition (L1/L2 association) that were presented twice each (48 words per condition per language). The dual-language context included 192 words (48 words per condition per language). Across the experiment, each word/pseudoword was repeated four times per language. Each condition had a similar number of words and pseudowords. Each word occurred an equal number of times in each condition and was preceded by a word/pseudoword an equal number of times. There were no more than three consecutive stimuli of each type (word/pseudoword) or condition (L1/L2 association) and words and their corresponding pseudowords were not presented consecutively. Each trial started with a fixation cross that was presented for 300 ms. Next, the stimulus was shown on the screen for 2500 ms or until a response was given.

After the lexical decision task, participants completed a task indicating for each target word how likely they were to use Mandarin or English in daily life (0–100%).

This was followed by the picture-naming task and language-background questionnaire described above.

2.4. Data Analysis

The data and the analysis script are available at: https://osf.io/vjgwx/.

Analyses only included word responses. Data were analysed using generalised linear mixed-effect models (for accuracy) and linear mixed-effect models (for RTs) using lme4 package version 1.1-23 and lmerTest 3.1.-2 in R 4.0.2. Analyses started with a maximal structure that included fixed effects and random intercepts and slopes for all within-participant and within-item predictors. We first removed correlations between random slopes and intercepts when models did not converge, followed by removal of by-item slopes explaining the lowest amount of variance until convergence was reached. The script with final models can be found on the OSF page. The two-level categorical variables were coded as −0.5 and 0.5. Continuous fixed effects were centred and scaled.

We first analysed accuracy by using incorrect (0) and correct (1) responses as the dependent variable. We included condition (L2-associated topic = −0.5; L1-associated topic = 0.5), language (Mandarin = −0.5; English = 0.5), and context (single = −0.5; dual = 0.5) as predictors. The same model was run for RTs as the DV, which were log transformed. Next, we replaced condition (L1-/L2-associated topic) with the actual daily life language use for each item, using the average language-use ratings participants provided for each item, with higher scores indicating more Mandarin use. This analysis examined whether there was a direct relationship between how often specific words were used in daily life by our participants in each language and their RTs in the lexical decision task.

3. Results

3.1. Accuracy

Accuracy was very high overall (>90% for all participants). It was higher for the Mandarin words (M = 96.7%, SD = 2.5) than the English words (M = 95.2%, SD = 3.2; β = −0.358, SE = 0.178, z = −2.012, p = 0.044). Accuracy was similar in the single- (M = 96.0%, SD = 2.1) and dual-language contexts (M = 95.9%, SD = 2.7; β = −0.012, SE = 0.101, z = −0.117; p = 0.907) and for L2-associated (M = 96.9%, SD = 2.2) and L1-associated topic words (M = 94.9, SD = 3.1; β = −0.412, SE = 0.230, z = −1.793, p = 0.073). There was a significant interaction between condition and language (β = −0.727, SE = 0.297, z = −2.449, p = 0.014). In English, there was an effect of condition, with higher accuracy for L2 (English)-associated topic words (M = 96.9%, SD = 3.1) than for L1 (Mandarin)-associated topic words (M = 93.4%, SD = 4.8; β = −0.658, SE = 0.279, z = −2.355, p = 0.019). In Mandarin, accuracy did not differ between conditions (L2-topic M = 96.9%, SD = 2.6; L1-topic M = 96.5%, SD = 2.9; β = −0.143, SE = 0.225, z = −0.636, p = 0.525). There were no other significant interactions (all zs < |1.88|).

3.2. Reaction Times

First, we removed incorrect responses and RT outliers below or above 2.5 SD per participant and per condition, language, and context (2.7% of correct responses).

3.2.1. Main Analysis

The main analysis[2] included condition, context, and language. There was a main effect of condition (β = 0.035, SE = 0.014, t = 2.596, p = 0.012) that interacted with language (β = 0.048, SE = 0.019, t = 2.537, p = 0.014). This reflected that L1- and L2-associated topic words were processed differently in English and Mandarin. In English, there was a main effect of condition (β = 0.060, SE = 0.019, t = 3.151, p = 0.003), with L2 (English)-associated topic words (M = 612.6 ms, SD = 50.0) being processed faster than L1 (Mandarin)-associated topic words (M = 655.4 ms, SD = 67.9; see Figure 1). In Mandarin, there was no effect of condition (β = 0.011, SE = 0.014, t = 0.810, p = 0.422; L2-topic M = 604.8 ms, SD = 71.8; L1-topic M = 613.0 ms, SD = 81.1; see Figure 1 and Table 4).

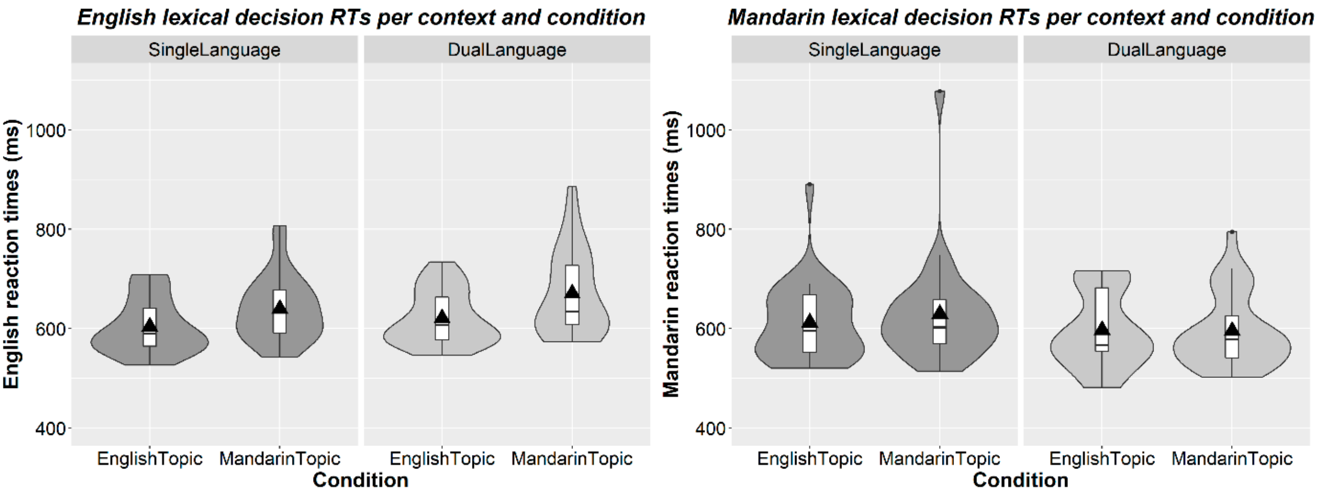

Figure 1. Violin plots showing the distribution of the lexical decision RTs for English (**left**) and Mandarin (**right**). Within each language plot, the single-language context is presented on the left and the dual-language context on the right. Within each context, English/L2 topic RTs are presented on the left and Mandarin/L1 topic RTs on the right. The outline of the violin plot shows the density of data points for the different RTs, with the boxplot showing the interquartile range. The thin horizontal line shows the median and the black triangle the mean.

Table 4. Table showing the mean RTs (and SDs) per context (single-language/dual-language), language (English/Mandarin) and condition (L2/English or L1/Mandarin topic).

Condition	L2/English Topic	L1/Mandarin Topic
Single-language		
English	604.0 (53.9)	639.6 (64.7)
Mandarin	612.3 (78.3)	629.8 (104.1)
Dual-language		
English	621.1 (53.7)	670.9 (85.1)
Mandarin	597.0 (75.0)	595.6 (74.3)

In addition, there was a main effect of language (β = 0.047, *SE* = 0.012, *t* = 4.027, $p < 0.001$), reflecting faster overall responses in Mandarin (*M* = 608.9 ms, *SD* = 75.1) than in English (*M* = 633.6 ms, *SD* = 57.5). There was no main effect of single/dual-language context (β = 0.0001, SE = 0.011, t = 0.010, p = 0.992), nor an interaction between context and condition (β = −0.005, *SE* = 0.011, *t* = −0.451, *p* = 0.655). However, there was an interaction between language and context (β = 0.078, *SE* = 0.015, *t* = 5.042, $p < 0.001$). This reflected that while Mandarin responses were faster in the dual- (*M* = 596.3 ms, *SD* = 72.1) than in the single-language condition (*M* = 621.3 ms, *SD* = 89.5; β = −0.039, *SE* = 0.014, *t* = −2.750, *p* = 0.010), the opposite was true for English responses (single *M* = 621.4 ms, *SD* = 55.9, dual *M* = 645.6 ms, *SD* = 67.3; β = 0.039, *SE* = 0.012, *t* = 3.240, *p* = 0.003). Figure 1 suggests that English responses to L1 (Mandarin)-topic words were especially slow in the dual-language condition. However, the three-way interaction among context, language, and condition did not reach significance (β = 0.037, *SE* = 0.020, *t* = 1.903, *p* = 0.068).

3.2.2. Percentage Language Use

Next, we included percentage English/Mandarin use per item rather than the two-level condition English/Mandarin topic[3]. This analysis showed very similar results, with a significant interaction between %language use and language (β = 0.026, *SE* = 0.009, *t* = 2.914, *p* = 0.005). In English, there was a significant relationship between %language use and RTs (β = 0.023, *SE* = 0.010, *t* = 2.374, *p* = 0.021), with faster RTs for items used more often in English in daily life (see Figure 2). In contrast, in Mandarin, there was no such relationship (β = −0.003, *SE* = 0.007, *t* = −0.387, *p* = 0.701; see Figure 2). There was no three-way interaction between %language use, language, and context (β = 0.008, *SE* = 0.009,

$t = 0.910$, $p = 0.371$), suggesting that the relationship between %language use and English RTs but not Mandarin RTs was similar in the single- and dual-language contexts.

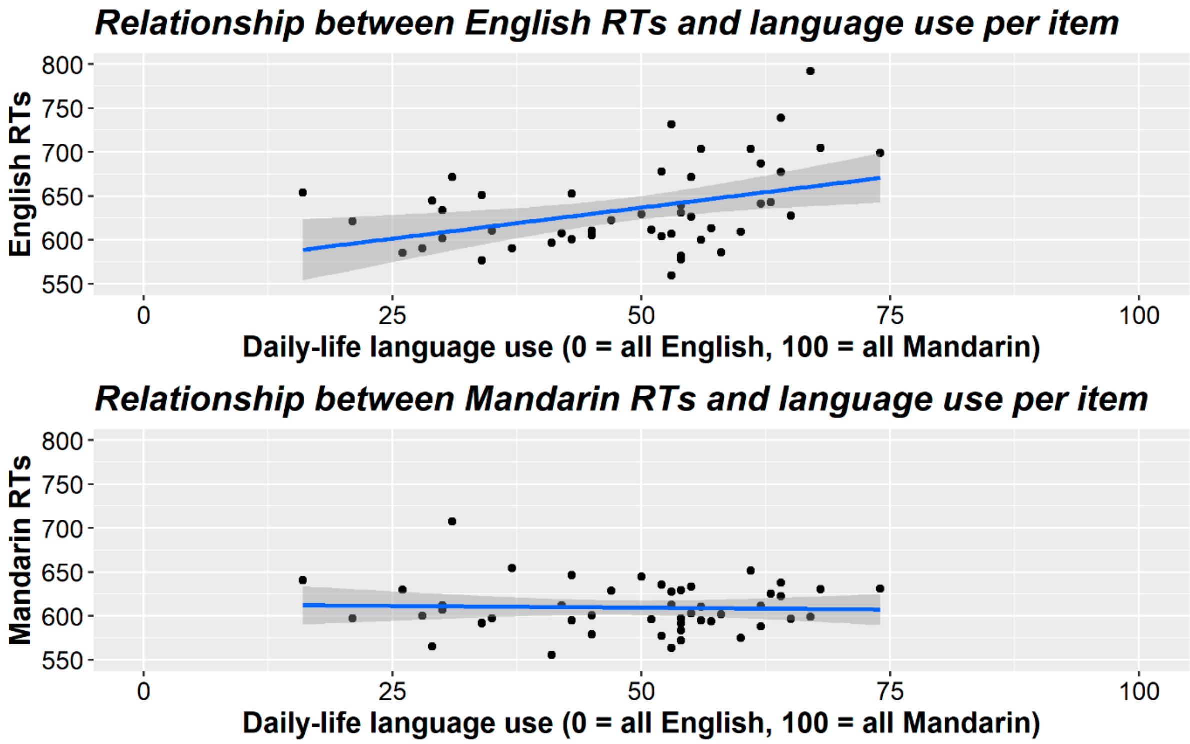

Figure 2. Scatter plots showing the correlation between daily-life language use and lexical decision RTs in English (**top**) and Mandarin (**bottom**). Each dot represents an individual item. Daily-life language use scores are averaged across participants for each item, with lower scores representing more English use and higher scores representing more Mandarin use in daily life.

4. Discussion

We assessed if topic influences processing of words more closely associated with L1 or L2. Both accuracy and RT analyses showed an influence of topic on L2 processing, with faster and more accurate responses to words associated with L2 than L1. This was directly associated with daily life language use: Words that were used more often by our participants in their L2 English were processed faster than words used more often in their L1 Mandarin. Effects of topic on L2 processing were found in both single- and dual-language contexts. However, this effect was not observed in L1 processing.

4.1. Topic Specific Effects on Language Processing

These topic-specific findings support the Complementarity Hypothesis (Grosjean 2016), which highlights that a bilingual's use and proficiency might depend on the domain, topic, and people one is interacting with. Here, we show that topic can influence lexical processing in a bilingual's non-native language. While these findings are specific to international students who associate their L1 with home/family life and their L2 with studying, associating certain languages with certain topics is not unique to this group of participants. Many bilinguals will recognise the feeling of using one language more often with certain people (e.g., a different language when talking with grandparents than with friends) or for certain topics (e.g., researchers using English to write a paper but perhaps not when talking about their favourite food at home). Our study emphasises that L2 proficiency and use are not global and topic-independent (cf. Brysbaert et al. 2021), even though it is often assessed as if it is.

Our findings are also in line with studies showing influences of L1-associated cues on L2 processing (e.g., Jared et al. 2013; Zhang et al. 2013). However, in the current study, words existed and were of similar objective frequency in both languages. Topic effects thus did not occur because stimuli were uniquely associated with or far more frequent in a language/culture but rather because of the bilingual's personal experiences with certain topics. This is an important aspect to consider in future studies on bilingual processing, which often match languages/stimulus conditions on objective frequency but typically do not consider how closely stimuli are associated with each of their bilingual participants' languages.

4.2. Language Context

While topic effects on L2 processing were somewhat larger in the dual-language context, the difference with single-language contexts was not significant. This suggests that these topic-dependent processing effects are present even when the other language is not directly present in the context. The finding that these topic-effects were present in both single- and dual-language contexts suggests that effects of topic might arise because of frequency of use facilitating language processing. That is, because bilinguals use and experience certain words more frequently in their L2 (e.g., words related to studying at an English-speaking University), their lexical processing of these words might be facilitated. This explanation would be in line with the common finding that words with a higher objective frequency level (e.g., words such as "table" that occur more often than words such as "hammer") take less time to process. While these "objective" (e.g., derived from corpora) global frequency differences are often considered in experiments and language models, more subjective, individual (in this case topic-dependent) frequency differences are often not. Our findings, however, can be explained by considering subjective, topic-specific frequency of use in commonly used models of bilingual word comprehension such as the BIA (Dijkstra and Van Heuven 1998) or BIA+ (Dijkstra and Van Heuven 2002). These models incorporate frequency differences between languages, with L2 words having lower resting level activation due to lower frequency of use in bilinguals who use their L1 more often. In a similar manner, L2 words more frequently used in the L2 could have a higher resting level activation level than L2 words more frequently used in the L1.

We hypothesised that if topic effects are due or related to language interference, these effects should be larger in a dual-language environment, in which the "other" language is likely to interfere more than in a single-language environment in which it is not present. Effects of topic were somewhat larger in dual-language contexts with more cross-language interference, suggesting that topic effects in the L2 might increase when L1 interference increases. However, this interaction was not significant and future research is needed to evaluate whether topic effects are indeed modulated by contexts with different levels of language competition.

Overall, RT differences between the single- and dual-language contexts varied between the languages. In English, participants showed longer RTs in dual- than single-language contexts. This dual-language cost is likely to reflect a combination of increased language interference and competition in the dual-language context as well as switches between scripts that were present in the dual- but not in the single-language context. In contrast, however, Mandarin RTs were slower in the single-language context than in the dual-language context. Closer inspection suggested that this was purely due to the first Mandarin single-language condition eliciting relatively long RTs. The second Mandarin single-language condition was completed faster than the dual-language condition, in line with English responses. Responses in this first Mandarin single-language condition were the slowest of all conditions, including the English trials. This is a surprising finding but could potentially be due to participants living in North America/UK and having limited exposure to Chinese characters at the moment of testing, which could potentially lead to participants needing more time to process the Mandarin words at the start. Although participants completed a practice round with all stimuli prior to starting the first single-language block, this might not have been sufficient practice.

4.3. First versus Second Language

Effects of topic were observed only in the L2 but not in the L1. These findings are in line with recent research showing greater topic specificity in the non-dominant language (Tiv et al. 2020). It also suggests that the L1 is less susceptible to topic–language associations, in line with studies showing larger frequency effects in the L2 than L1 (e.g., Van Wijnendaele and Brysbaert 2002). L1 representations, which in this case were acquired earlier in life, are likely to be more strongly entrenched and, therefore, less susceptible to differences in (subjective) frequency (e.g., Diependaele et al. 2013). Furthermore, it is possible that L1 processing is only influenced when objects are uniquely associated with one language (e.g., Jared et al. 2013) but not when words or concepts exist in both languages.

In our study, each word was presented multiple times, which could have diminished effects as the words became associated with both languages throughout the task. It is possible that this especially diminished the effects of topic in the L1, the language that might be influenced less by topic effects in the first place. However, even in the very first practice block (in which participants saw each word for the very first time), topic effects were found in the L2 (L2-associated topic responses were 81ms faster than L1-associated topics responses) but not in the L1 (a difference of 3ms between L1- and L2-associated topics). This suggests that even upon first presentation, effects of topic were only present in the L2 and not in the L1.

5. Conclusions

While much research has focused on differences between bilinguals (for example, some bilinguals using a language more often than other bilinguals), language use might also differ within bilinguals depending on the topic of conversation. Here, we show that the topic can influence bilingual processing times, with L2 words associated with an L1 topic requiring longer processing times than L2 words associated with an L2 topic. These findings show the importance of considering *within-participant* patterns of language use over and beyond global proficiency or language use differences between bilinguals. The topic of conversation, and the materials researchers use, can influence bilingual processing and can consequently lead to, or mask, differences between bilinguals and languages.

Author Contributions: All authors contributed to the development of the study's methodology and stimulus creation. V.S. and A.d.B. led the conceptualisation, project administration, programming and data collection, formal analysis, and data curation. V.S. and A.d.B. wrote the original draft, which was reviewed and edited by H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee in the Department of Psychology at the University of York (approval given on 24 June 2020, identification number 869).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available at: https://osf.io/vjgwx/.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The English (L2) and Mandarin (L1) words and pseudowords are presented in the table below. Words in the "L2-associated Topic" condition were associated with life as a student (e.g., life at university, activities associated with student life). Words in the "L1-associated Topic" condition were associated with life at home (e.g., words related to childhood experiences, family, and the home environment). The languages and conditions were matched on ZIPF frequency scores from SUBTLEX-UK (Van Heuven et al. 2014) and SUBTLEX-CH (Cai and Brysbaert 2010): English words associated with English/L2 topic:

$M = 4.0$, $SD = 0.9$; English words associated with Mandarin/L1 topic: $M = 4.0$, $SD = 0.7$; Mandarin words associated with Mandarin/L1 topic: $M = 4.1$, $SD = 0.7$; Mandarin words associated with English/L2 topic Topic: $M = 4.1$, $SD = 0.6$. English words in the L2-topic and L1-topic conditions were similar in terms of number of letters ($M = 7.3$, $SD = 2.1$ vs. $M = 6.7$, $SD = 2.4$), syllables ($M = 2.4$, $SD = 0.9$ vs. $M = 2.2$, $SD = 0.8$), and phonemes ($M = 6.5$, $SD = 2.0$ vs. $M = 5.3$, SD = 2.1). We also matched the two English conditions on valence ($M = 5.8$, $SD = 0.8$ vs. $M = 6.3$, $SD = 1.1$; Warriner et al. 2013). Only two- and three-character Chinese words were included, and the number of three-character words was similar for the two conditions.

Table A1. Stimuli (words and pseudowords) used in the experiment by condition (English or Mandarin topic).

Condition	Word English	Pseudoword English	Word Mandarin	Pseudoword Mandarin
EnglishTopic	Exam	emap	考试	灭试
EnglishTopic	Assignment	antointint	作业	作仁
EnglishTopic	Plagiarism	scafeirism	剽窃	蚕窃
EnglishTopic	Degree	dechie	学位	额位
EnglishTopic	Lecture	tuffure	讲座	讲兔
EnglishTopic	Seminar	selilir	研讨会	研讨炎
EnglishTopic	Tutorial	bumonial	教程	甘程
EnglishTopic	Library	fidrory	图书馆	图书饼
EnglishTopic	Research	resonfed	研究	研怕
EnglishTopic	Notes	nopto	笔记	笔逃
EnglishTopic	Society	somiely	社团	社楠
EnglishTopic	Event	enant	事件	事门
EnglishTopic	Pub	wuy	酒吧	酒答
EnglishTopic	Community	corcaroty	社区	者区
EnglishTopic	Trip	snix	行程	灰程
EnglishTopic	Classroom	clorckheam	教室	勾室
EnglishTopic	Classmate	drastnent	同学	窝学
EnglishTopic	Roommate	reannent	室友	室填
EnglishTopic	Teacher	tielder	老师	枝师
EnglishTopic	Instructor	inchroctor	辅导员	厚导员
EnglishTopic	Assistant	attontant	助教	助努
EnglishTopic	Football	loodbads	足球	足壳
EnglishTopic	Basketball	batvonpell	篮球	篮晚
EnglishTopic	Sport	snurs	运动	运忽
MandarinTopic	Adoption	asoxtion	收养	收碧
MandarinTopic	Orphan	urswan	孤儿	孤另
MandarinTopic	Father	tascer	父亲	父贯
MandarinTopic	Brother	crither	兄弟	兄彤
MandarinTopic	Sister	douter	姐妹	固妹
MandarinTopic	Kitten	kollen	小猫	晓猫
MandarinTopic	Puppy	vopty	小狗	晓狗
MandarinTopic	Decorations	digocations	装饰品	装饰滔
MandarinTopic	Vase	vost	花瓶	花逆
MandarinTopic	Toy	tra	玩具	徒具
MandarinTopic	Playground	planploust	操场	喂场
MandarinTopic	Doll	vils	洋娃娃	迈娃娃
MandarinTopic	Kindergarten	windergannin	幼儿园	幼儿曼
MandarinTopic	Swings	blongs	秋千	秋克
MandarinTopic	Playhouse	plaghound	剧场	剧斯
MandarinTopic	Piano	reino	钢琴	受琴
MandarinTopic	Violin	vionad	小提琴	小提泰
MandarinTopic	Cello	tenlo	大提琴	大恰琴
MandarinTopic	Sofa	sogo	沙发	沙久
MandarinTopic	Kitchen	jaipsen	厨房	厨着
MandarinTopic	Tableware	fawrepare	餐具	餐各
MandarinTopic	Balcony	bercoty	阳台	促台
MandarinTopic	Photo	twolo	照片	照欠
MandarinTopic	Doorbell	roorbews	门铃	门淡

Notes

1 The ZIPF score is a standardised measure that is independent of the size of the database, thus allowing for closer comparisons between different databases that might differ in the number of words included (cf. Van Heuven et al. 2014, for more details on how to compute ZIPF scores).

2 We also analysed switching (difference between switch and non-switch trials) and mixing (difference between non-switch and single-language) effects. Here, we only included the twelve words per condition that were preceded by another word (and not the trials preceded by pseudoword). There were no interactions between trial type and condition, suggesting that topic did not influence switching/mixing effects.

3 Individual language-use percentages per participant and item (rather than per item across participants) also showed a significant interaction between %language use and language.

References

Anderson, John A. E., Lorinda Mak, Aram Keyvani Chahi, and Ellen Bialystok. 2018. The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods* 50: 250–63. [CrossRef] [PubMed]

Anwyl-Irvine, Alexander L., Jessica Massonnié, Adam Flitton, Natasha Kirkham, and Jo K. Evershed. 2020. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods* 52: 388–407. [CrossRef] [PubMed]

Berkes, Matthias, Deanna C. Friesen, and Ellen Bialystok. 2018. Cultural context as a biasing factor for language activation in bilinguals. *Language, Cognition and Neuroscience* 33: 1032–48. [CrossRef]

Blumenfeld, Henrike K., and Viorica Marian. 2007. Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes* 22: 633–60. [CrossRef]

Brysbaert, Marc, and Michaël Stevens. 2018. Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition* 1: 9. [CrossRef]

Brysbaert, Marc, Emmanuel Keuleers, and Paweł Mandera. 2021. Which words do English non-native speakers know? New supernational levels based on yes/no decision. *Second Language Research* 37: 207–31. [CrossRef]

Cai, Qing, and Marc Brysbaert. 2010. SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE* 5: e10729. [CrossRef] [PubMed]

Chaouch-Orozco, Adel, Jorge González Alonso, and Jason Rothman. 2021. Individual differences in bilingual word recognition: The role of experiential factors and word frequency in cross-language lexical priming. *Applied Psycholinguistics* 42: 447–74. [CrossRef]

De Bruin, Angela, Manuel Carreiras, and Jon Andoni Duñabeitia. 2017. The BEST dataset of language proficiency. *Frontiers in Psychology* 8: 522. [CrossRef]

Diependaele, Kevin, Kristin Lemhöfer, and Marc Brysbaert. 2013. The word frequency effect in first-and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology* 66: 843–63. [CrossRef] [PubMed]

Dijkstra, Ton, and Walter J. B. Van Heuven. 1998. The BIA model and bilingual word recognition. In *Localist Connectionist Approaches to Human Cognition*. East Sussex: Psychology Press, pp. 189–225.

Dijkstra, Ton, and Walter J. B. Van Heuven. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5: 175–97. [CrossRef]

Grosjean, François. 2016. The complementarity principle and its impact on processing, acquisition, and dominance. In *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*. Cambridge: Cambridge University Press, pp. 66–84.

Hartsuiker, Robert J. 2015. Visual cues for language selection in bilinguals. In *Attention and Vision in Language Processing*. New Delhi: Springer, pp. 129–45.

Jared, Debra, Rebecca Pei Yun Poh, and Allan Paivio. 2013. L1 and L2 picture naming in Mandarin–English bilinguals: A test of bilingual dual coding theory. *Bilingualism: Language and Cognition* 16: 383–96. [CrossRef]

Keuleers, Emmanuel, and Marc Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42: 627–33. [CrossRef] [PubMed]

Li, Yunqing, Jing Yang, K. Suzanne Scherf, and Ping Li. 2013. Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and Language* 127: 452–62. [CrossRef] [PubMed]

Molnar, Monika, Antonio Ibáñez-Molina, and Manuel Carreiras. 2015. Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language* 81: 91–104. [CrossRef]

Roychoudhuri, Kesaban S., Seema G. Prasad, and Ramesh K. Mishra. 2016. Iconic native culture cues inhibit second language production in a non-immigrant population: Evidence from Bengali-English bilinguals. *Frontiers in Psychology* 7: 1516. [CrossRef] [PubMed]

Rubin, Orit, and Nachshon Meiran. 2005. On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 1477. [CrossRef]

Tiv, Mehrgol, Jason W. Gullifer, Ruo Ying Feng, and Debra Titone. 2020. Using Network Science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics* 56: 100913. [CrossRef] [PubMed]

Van Hell, Janet G., and Darren Tanner. 2012. Second language proficiency and cross-language lexical activation. *Language Learning* 62: 148–71. [CrossRef]

Van Heuven, Walter J. B., Pawel Mandera, Emmanuel Keuleers, and Marc Brysbaert. 2014. SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology* 67: 1176–90. [CrossRef]

Vanlangendonck, Flora, David Peeters, Shirley-Ann Rueschemeyer, and Ton Dijkstra. 2020. Mixing the stimulus list in bilingual lexical decision turns cognate facilitation effects into mirrored inhibition effects. *Bilingualism: Language and Cognition* 23: 836–44. [CrossRef]

Van Wijnendaele, Ilse, and Marc Brysbaert. 2002. Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance* 28: 616–27. [CrossRef]

Warriner, Amy Beth, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45: 1191–207. [CrossRef] [PubMed]

Woumans, Evy, Clara D. Martin, Charlotte Vanden Bulcke, Eva Van Assche, Albert Costa, Robert J. Hartsuiker, and Wouter Duyck. 2015. Can faces prime a language? *Psychological Science* 26: 1343–52. [CrossRef]

Zhang, Shu, Michael W. Morris, Chi-Ying Cheng, and Andy J. Yap. 2013. Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proceedings of the National Academy of Sciences USA* 110: 11272–77. [CrossRef] [PubMed]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Languages Editorial Office
E-mail: languages@mdpi.com
www.mdpi.com/journal/languages

Academic Open
Access Publishing

www.mdpi.com

ISBN 978-3-0365-8038-8