Title
Catching the cognitive consequences of bilingual language processing on the fly: An approach to reframe the discussion about bilingualism

Author
Navarro-Torres, Christian Alberto

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Catching the cognitive consequences of bilingual language processing on the fly: An approach to reframe the discussion about bilingualism

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Language Science

by

Christian A. Navarro-Torres

Dissertation Committee:
Distinguished Professor Judith F. Kroll, Chair
Assistant Professor Xin Xie
Associate Professor Gregory Scontras
Professor Elizabeth Peña
Associate Professor Susanne Jaeggi
Professor Paola E. Dussias

2021

DEDICATION

To

My family, and anyone who finds a way to overcome suffering through the pursuit of meaning

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to express the deepest appreciation to my committee for being incredibly accommodating as a transfer graduate student, and for providing a context in which I could pursue the research program I had already started before arriving at UCI. I also thank the research assistants that have worked with me since my arrival at UCI, including Eliud, Ania, Abimael, and Nam, all of which helped me build an intellectually engaging context to talk about research, and pave the way for future generations. Dayra, thank you for being incredibly patient and resilient in this whole process! Lupe, Andy, and Audrey, without your support, this dissertation would have not been possible without your tremendous efforts with data collection and coding.

I would also like to thank the following people who have shaped my graduate experience in one form or another:

Rosa Guzzardo-Tamargo and Jorge Valdes Kroff (for opening the doorway to my entire career that night we met at BUCLD); Eleonora Rossi, Jason Gullifer, Grant Berry, Nick Ray, Sarah Fairchild, Yanina Prystauka, Randi Goe, Antje Stoehr, and Manuel Pulido (for helping create an amazing community at the CLS throughout its various stages, and for being genuine friends); Megan Zirnstein and Melinda Fricke (for creating wonderful first impressions of a post-doc life); Gerrit Jan Kootstra (for that thoughtful conversation we had at the hotel in DC); Holger Hopp (for giving me an opportunity to conduct amazing research abroad in the culture I admire the most); Dalia Garcia (for being an exemplary role model for other students through your unshakable determination); Gerry Altman (for providing words of encouragement during my first semester in grad school); Jared Novick (for inspiring the creation of this dissertation through your work); David Rosenbaum (for teaching me the most valuable lesson I learned in graduate school: how to give a presentation); Rena Torres Cacoullos (for shaping how I view and value the study of language); Adele Goldberg (for shaping my research career since the first talk you gave at the CLS); David Green (for being willing to share the canvas and paintbrush through which we create and explore ideas); Victor Castro (for instilling a sense of discipline, responsibility through the application of *Bildung*, and for allowing me to see the world differently through German culture); Letitia Naigles (for picking me up from the ground and giving me an opportunity to have a future); Giuli Dussias (for being my second mother); Annie Beatty-Martínez (for the unconditional friendship and for always pushing me to be a better person); Judy Kroll (for teaching me the most important lesson of all: that no matter what you do, the world will often judge you for what you represent, and not for who you are).

Lastly, I would like to thank the funding agencies that supported the creation, implementation, and writing of this dissertation. The work presented here was supported in part by the National Science Foundation [Award number DGE-1255832 and Grant number BCS-1946051] and the National Institute of Health [Grant number HD098783].

VITA

Christian A. Navarro-Torres

2011	B.A. in Philosophy, University of Puerto Rico, Río Piedras
2011-13	Research Assistant, UConn Child language lab, University of Connecticut
2013-14	Research technologist and lab manager, Center for Language Science, Pennsylvania State University
2016	M.S. in Cognitive Psychology, Pennsylvania State University
2016-19	Ph.D. student in Cognitive Psychology, University of California, Riverside
2019-pres	Ph.D. in Language Science, University of California, Irvine

FIELD OF STUDY

The cognitive and neural bases of language processing in bilinguals

PUBLICATIONS

Navarro-Torres, C. A., Dussias, P. E., & Kroll, J. F. (*under review)*. When Exceptions Matter: Bilinguals regulate their dominant language to exploit structural constraints. *Language, Cognition and Neuroscience.*

Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language*, *222*, 105014.

Hoshino, N., Beatty-Martínez, A. L., Navarro-Torres, C. A., & Kroll, J. F. (2021). Do Cross-Language Script Differences Enable Bilinguals to Function Selectively When Speaking in One Language Alone?. *Frontiers in Communication*, *6*, 119.

Kroll, J. F., Takahesu Tabori, A., & Navarro-Torres, C. A., (2021). Capturing the variation in language experience to understand language processing and learning. In *Language, Interaction and Acquisition* (Vol. 12, Issue 1, pp. 82–109). John Benjamins.

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. In J. Treffers-Daller, E. Ruigendijk, & J. E. Hofweber, (Eds.)., Behavioral and Neurophysiological Approaches to Code-Switching and Language Switching [Special Issue]. *Frontiers in Psychology*, 11:1699.

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, T. Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Kroll, J. F. & Navarro-Torres, C. A. (2020). Bilingualism. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, J. T. Wixted (Ed.).

Navarro-Torres, C. A., Garcia, D. L., Chidambaram, V., & Kroll, J. F. (2019). Cognitive Control Facilitates Attentional Disengagement during Second Language Comprehension. *Brain sciences*, *9*(5), 95.

Fricke, M., Zirnstein, M., Navarro-Torres, C. A., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition*, *22*(1), 200-207.

Zhang, H., Eppes, A., Beatty-Martínez, A. L., Navarro-Torres, C. A., & Diaz, M. T. (2018). Task difficulty modulates brain-behavior correlations in language production and cognitive control: Behavioral and fMRI evidence from a phonological Go-NoGo picture naming paradigm. *Cognitive, Affective, and Behavioral Neuroscience, 18,* 964-981.

Weiss, D. J., Bulgarelli, F., Navarro-Torres, C., & Morales, J. (2017). Translating research to practice in the language sciences. *Translational issues in psychological science*, *3*(1), 1.

ABSTRACT OF THE DISSERTATION

Catching the cognitive consequences of bilingual language processing on the fly: An approach to reframe the discussion about bilingualism

by

Christian A. Navarro-Torres

Doctor of Philosophy in Language Science

University of California, Irvine, 2021

Distinguished Professor Judith F. Kroll, Chair

An aim of research on bilingualism is to understand how the brain adapts to the use of more than one language. Although several important discoveries and insights about the consequences of bilingualism have been generated over the last several decades, concerns about replicability have narrowed the scope of inquiry and discussion to the application of prescriptions about sample size and method. In this dissertation, I critique this approach and reformulate its value by placing it in the broader context of science as a discovery process, in which incremental understanding, methodological and analytical diversification, the framing of our questions, and even underpowered studies, are essential to advance. I propose that this necessitates research practices and tools that: 1) focus on examining variation in the relation between language processing and cognitive functioning; 2) allow us to identify meaningful interactions rather than main effects only; and 3) provide a rich characterization of the participant sample to identify bilingual phenotypes: the adaptive variety induced by the interplay between biology and culture. In seeking to

apply this framework, I then present two empirical studies investigating the cognitive mechanisms that enable fluent language use, and examine the hypothesis that proficient bilingualism is characterized by the active engagement of a regulatory mechanism that adapts to the demands of the language environment. To test this, I examine individual differences in language production and cognitive control across four groups of proficient bilingual phenotypes who differ with respect to the interactional contexts of language use. In Study 1, the results suggest that lexical access is mediated by a systematic interaction between regulation of the dominant language and cognitive control, but that the manifestation of this interaction depends on whether bilinguals are immersed in a second-language environment, a context in which active regulation is required. In Study 2, the results further confirm that dominant-language regulation is not analogous to proficiency per se; rather, it may more adequately reflect the coordination of language-related and domain-general resources that interactively contribute to the accessibility of information in the language network. Together, these studies suggest that language regulation is a fundamental feature of bilingual brains that engages a combination of language-related and domain-general cognitive resources to enable proficient language use.

PREFACE

A central goal of this dissertation is to examine how bilinguals who grow up in different contexts of language use come to exploit cognition to enable fluent speech. In its initial form, this dissertation sought to do this by investigating language processing using a variety of methods that assessed behavior, eye-movements, and brain electrophysiology in two groups of proficient bilinguals. However, the creation, implementation, and writing of this dissertation coincided with three notable events that changed its outcome. One is that, following an initial move from Penn State University to the University of California, Riverside between my second and third year of graduate school, there was a second move to the University of California, Irvine during my fifth year as the dissertation was being created. A second event involves the COVID-19 pandemic, which coincided with the timeline in which I was expected to travel to Granada, Spain to begin data collection. This made in-person research impossible, requiring a shift to online behavioral experimentation and preventing electrophysiological data collection. Lastly, my health was compromised following the pandemic in a manner that rendered me unable to conduct research for a period of time.

As such, the dissertation in its current form is unusual in that it represents an attempt to explore the central aim outlined above by drawing from a combination of already-published work, together with new behavioral data that was collected using an online platform. Specifically, Chapters 3 and 5 contain two published studies that were written collaboratively (one as first author, and the second as a first co-author). The publication included in Chapter 3 (Navarro-Torres, Beatty-Martínez, Kroll, & Green, 2021), which appeared in *Brain and Language*, provides a conceptual framework that reformulates the discussion about bilingualism and its consequences in the form of discovery science. The theoretical questions explored in Chapter 4, as well as the empirical data presented in Chapters 5 and 6, follow from this conceptual framework. The second publication (Beatty-Martínez, Navarro-Torres, et al., 2020), which appeared in *Journal of Experimental Psychology: Language, Memory, and Cognition*, is included in Part A of Chapter 5. This study is included in its original form to provide an adequate context of the results reported in Part B of Chapter 5, which consist of a reanalysis on a subset of the data reported in Beatty-Martínez, Navarro-Torres et al. (2020). Finally, in Summary and Conclusions, I attempt to bring the data presented in Chapters 5 and 6 back to the conceptual and practical issues raised in Chapters 2 and 3.

CHAPTER 1: INTRODUCTION

Bilinguals are individuals who actively use more than one language. However, bilingualism arises in many ways: some individuals grow up with two languages from birth, while others acquire a second language (L2) after early childhood, once the native language (L1) has been established. In some cases, these individuals continue to use both languages throughout their lives. In other cases, the use of a language is restricted to certain contexts (e.g., home, school, work) because of the socio-political circumstances surrounding the individual. In recent years, linguists, psychologists, and neuroscientists have begun to understand that bilingualism is a natural circumstance of human experience, and not a special case that results in disordered speaking or thinking. This is in large part due to a series of important discoveries that researchers have made about bilingualism that have been vital to understand the relation between language, cognition, and the brain. Below I briefly allude to four discoveries that have fundamentally shaped the current enterprise on research on bilingualism.

One important discovery is that when bilinguals use either of their languages, the language not in use also becomes active (for reviews, see Costa, 2005; Hanulová, Davidson, & Indefrey, 2011; Hartsuiker & Bernolet, 2017; Kroll, Bobb, & Wodniecka, 2006; Kroll et al., 2012; Lauro & Schwartz, 2019; Palma & Titone, 2020; Van Assche, Duyck, & Gollan, 2013). Evidence for the presence of cross-language activation comes from a variety of tasks and methods examining particular features of bilinguals' two languages. For instance, when cross-linguistic features converge, cross-language activation results in online facilitation, as in the case of cognates (i.e., translation-equivalent words that share form and meaning; for reviews, see Dijkstra, 2005; Lauro & Schwartz, 2019; Schwartz & Van Hell, 2012) or when there is cross-linguistic overlap in word order (for a review, see Hartsuiker & Bernolet, 2017). When cross-linguistic features conflict with one another, as in the case of interlingual homographs (i.e., words that share form but differ in meaning) or structures that partially share word order, cross-language activation tends to generate interference (e.g., Jared, 2015; Sanoudaki & Thierry, 2015; Vingron et al., 2021). Notably, the active presence of the non-target language seems to persist in languages with different scripts, such as Japanese-English and Mandarin-English (e.g., Hoshino & Kroll, 2008; Thierry & Wu, 2007), or different modalities, such as American Sign Language and English (Morford et al., 2011), suggesting that the phenomenon reflects a fundamental feature of the bilingual brain.

A second discovery is that the experience associated with cross-language activation may generate an openness for the two languages to dynamically interact with one another. Evidently, it seems intuitive that the dominant L1 would come to influence the weaker L2 (Dijkstra & Van Heuven, 2002; Kroll & Stewart, 1994), especially in late L2 learners with limited L2 proficiency (Ambridge & Brandt, 2013; Austin, Pongpairoj, & Trenkic, 2015; Bates & MacWhinney, 1987; Finn & Hudson Kam, 2015; Foucart & Frenck-Mestre, 2012; Hernandez, Li, & MacWhinney, 2005; Sabourin, Stowe, & de Haan, 2006). But remarkably, the research has shown that the opposite is also true. For instance, in proficient bilinguals, lexical retrieval in the dominant L1 is typically disrupted after speaking in the L2 (Guo et al., 2011; Misra et al., 2012; Rossi et al., 2018; Van Assche et al., 2013) or during prolonged immersion in an L2 environment, as in the case of learners who are studying abroad (Baus, Costa, & Carreiras, 2013; Linck, Sunderman, & Kroll, 2009).

Similar effects of L2-to-L1 influences can also be observed regarding phonetic changes (Chang, 2012, 2013) as well as changes in syntactic or collocational preferences in the L1 (Dussias & Sagarra, 2007; Otwinowska et al., 2021). These bidirectional influences are likely not incidental by nature. In fact, neuroimaging studies suggest that the neurophysiology supporting the two languages is largely the same (Perani & Abutalebi, 2005), and where differences in brain activation are observed for the bilingual's two languages, they tend to be associated with activation of control areas responsible for processing the less dominant or proficient language (Abutalebi & Green, 2016). Relatedly, research examining language production also shows that the lexical-semantic category boundaries of each language tend to converge with one another (Ameel et al., 2005, 2008; Zinszer et al., 2014), suggesting that the inter-related architecture of bilinguals' languages is likely a natural outcome of experience-dependent neuroplasticity.

A third discovery is that bilingual experience changes not only the processes and brain networks associated with language, but also those associated with more general control mechanisms (for reviews, see Bialystok, 2017; Bialystok et al., 2009; DeLuca et al., 2020). Evidence for this discovery came from a range of behavioral studies examining cognitive control (also known as executive function or attentional control), a hypothesized set of general-purpose processes that help regulate our thoughts and actions in a goal-oriented manner (Botvinick et al., 2001). In these studies, adult bilinguals were shown to outperform monolinguals on cognitive tasks that required generating quick decisions among competing responses (e.g., Bialystok, Craik, & Luk, 2008; Bialystok et al., 2004; Costa, Hernández, & Sebastián-Galles, 2008), monitoring for conflict in advance (Costa et al., 2009), shifting attention away from irrelevant information (e.g., Bialystok et al., 2006; Prior & MacWhinney, 2010).

At the same time, converging evidence from neuroscience showed that bilinguals engage brain regions associated with control mechanisms to a greater extent (Bialystok et al., 2005; Luk et al., 2010; for a review, see Pliatsikas & Luk, 2016) and in a more efficient manner (Abutalebi et al., 2012). These effects were not exclusive to brain function, as several other studies showed that L2 experience induced structural changes in brain regions associated with domain-general control (for reviews, see Li, Legault, & Litcofsky, 2014; Pliatsikas, DeLuca, & Voits, 2020). Notably, the neurological consequences of bilingualism have become vital in our understanding of brain pathology, as they have led to the discovery that bilingualism offsets the symptoms of dementia (for a review, see Bialystok, 2021a).

What accounts for these observed cognitive and neural changes in bilinguals? The initial interpretation was relatively straightforward: By virtue of the interrelated and non-selective nature of the bilingual language system, bilingual brains develop a unique ability to engage control processes to temporarily suppress the language not in use, which in turn generates neural changes that result in cognitive benefits over the lifespan. However, the fundamental challenge with this interpretation is that not all bilinguals reveal the same benefits, as evidenced by behavioral studies unable to replicate the bilingual-monolingual differences in cognitive tasks, particularly among young adult populations (for a review, see Bialystok, 2021b), and by the fact that the anatomical and functional changes observed in bilingual brains varied across studies (for reviews, see DeLuca et al., 2020; García-Pentón et al., 2015; Pliatsikas et al., 2019). But as noted at the beginning, bilingualism arises in many forms across communities and can also change throughout an individual's lifespan.

This leads us to the final and perhaps most important discovery in the last decade, which is that the bilingualism does not yield a singular outcome regarding adaptive change in cognitive and brain functioning (Luk & Bialystok, 2013). Rather, the consequences of bilingualism depend on the contexts of language use that bilinguals come to exploit throughout their lives. This is an important discovery especially in the context of the controversies surrounding the putative consequences of bilingualism for cognition (e.g., Nichols et al., 2020; Paap, Mason, & Anders-Jefferson, 2021), as the evidence in support of cognitive benefits for bilinguals relative to monolinguals has yielded mixed findings (an issue that I will reference throughout the dissertation and revisit in Summary and Conclusions).

To reconcile this idea with the initial reported effects of bilingualism on cognition, Green and Abutalebi (2013) proposed the adaptive control hypothesis. According to the hypothesis, distinct interactional contexts lead to specific adaptive changes to control processes. To illustrate, in a single-language context, only one language is typically used. Codeswitching contexts, in which bilinguals may freely alternate between stretches of the two languages within a conversation at will, offer opportunities for language integration. Finally, in dual-language contexts, both languages are used in the same environment but typically between speakers. Critically, dual-language contexts are hypothesized to increase the demands on cognitive control processes over and above the other contexts.

Support for this hypothesis comes from recent studies showing that bilinguals who operate in dual-language contexts exhibit reduced task-switching costs (Hartanto & Yang, 2016) and more efficient conflict resolution (Ooi, Goh, Sorace, & Bak, 2018). Similarly, increased diversity in language use across social contexts (another proxy for dual-language use) has been related to better behavioral cognitive performance in older adult bilinguals (Pot et al., 2018), and greater connectivity between brain regions associated with control in the form of goal maintenance and monitoring (Gullifer et al., 2018, Li et al., 2021). These findings are important because they illustrate how speakers can vary in how language and cognitive resources are engaged, and that such variation may partly stem from how bilinguals come to use their languages across communicative contexts throughout their lives.

Another illustration of context dependency comes from studies examining how cognitive control engagement dynamically changes in real time within the same speaker. For instance, greater efficiency in cognitive control engagement is observed when bilinguals are conducting a conflict-resolution task embedded in a dual-language context (e.g., pseudorandomly presenting words in both languages) relative to a single-language context (e.g., when only words from one of the languages are presented; Bosma & Pablos, 2020; Jiao et al., 2020a, 2020b; Timmer, Wodniecka, & Costa, 2021; Wu & Thierry, 2013). Although we still know little about how these dynamic changes relate to the long-term consequences that have previously been established in the literature, they begin to reveal the ways in which the bilingual brain adapts to context.

What is critical to highlight from these four discoveries is that the study of bilingualism allows us to explore fundamental questions about language, cognition, and the brain that may otherwise be difficult to examine (or even ask) with monolingual speakers only. We have learned that there is a great deal of interconnectedness between brain regions and networks, that such interconnectivity is open to change throughout an individual's lifespan, that speakers can vary in how they engage language and cognitive resources given their language experience, and that an understanding of how the brain adapts to context will be essential to generate effective theories about language and the brain more generally. It is in this sense that research on bilingualism is an illustrative case of a scientific enterprise characterized by discovery.

CHAPTER 2: ROADMAP

1. Discovery and replicability

Around the same time of these remarkable discoveries and insights, there was another research boom developing in parallel centered around the issue of replicability, the ability to obtain a consistent result across studies. In 2015, for instance, a large-scale replication study of psychological research reported replicating just under half of the studies that were included (Open Science Collaboration, 2015). Such findings were declared by many as indicative of a replication crisis within psychological sciences and rippled throughout media and several scientific outlets. The emergence of the replication crisis also coincided with concerns around the issue of misconduct, as exemplified in the case of Dutch social psychologist Diedrik Stapel, who was found to perpetrate academic fraud by deliberately falsifying data reported across several notable published studies (Markowitz & Hancock, 2014). The changing attitude centered around these issues shifted research practice from a scientific enterprise focused on discovery to one dominated by skepticism and whose priority was the fastidious evaluation of existing evidence (Ioannidis, 2005, 2012).

Since the replication crisis, there have been increased efforts to implement Big Data practices that seek to optimize replicability, including the use of renewed quantitative meta-analytic procedures, multisite collaborative replication efforts, and statistical power considerations (e.g., Ioannidis, 2012; Klein et al., 2018; Maxwell, Lau, & Howard, 2015; Simons, Holcombe, & Spellman, 2014; Szucs & Ioannidis, 2020; Wagenmakers et al., 2012). Several of these practices have been advocated for in research on bilingualism as well (Brysbaert, 2020; Lehtonen et al., 2018; Marsden, Morgan-Short, Trofimovich, & Ellis, 2018; Mertzen, Lago, & Vasishth, 2020) given that replication failures have also been widely publicized in the field, some of which remain hotly debated to this day (Nichols et al., 2020; Paap et al., 2021). Curiously, the emphasis on such efforts, though noble and important, has been exclusively on the application of method via prescriptive guidelines rather than on the generation and development of new ideas, or in the refinement of already established ideas. This raises the difficult, but crucial, question of how to reconcile these two lineages, scientific discovery on the one hand, and replication efforts on the other, in a manner that leads to progress.

Much like in the context of medicine, a notable feature of scientific prescriptions is that they often attempt to diagnose and treat causal problems using relatively simple heuristics; for instance, that properly powered studies increase the ability to detect true effects (Brysbaert, 2019), that in the context of mixed models, we should keep our random effects structure maximal (Barr et al., 2013), or that pre-registered studies mitigate questionable research practices involving data interpretation (Mertzen et al., 2020).

Unequivocally, prescriptions such as these have value. The challenge with them, however, is that it is not always clear how they can or should be implemented across different circumstances. For example, although the use of a maximal random effects structure decreases the probability of committing Type 1 errors, it is also the case that it increases Type 2 errors (Matuschek et al., 2017). Similarly, though large samples can generate robust effects for the validity of a hypothesis or construct, as in the case of recent research providing support for a critical period in L2 acquisition (Hartshorne, Tenenbaum, & Pinker, 2018), the same data can yield support for different conclusions (i.e., that L2 age-of-acquisition effects are not due to biological constraints) using alternative analytical procedures (van der Slik et al., 2021).

In short, though scientific prescriptions allow us to examine complex phenomena in relatively simple terms, they often do so at the expense of our ability to fully capture such complexity. This is because scientific prescriptions may be prone to ignore inherent variation in the data. Just like individuals can react very differently to a medical drug, individuals, as well as biological and cognitive systems more generally, can also generate different responses to particular environments or situations.

Remarkably, there is a great deal of systematicity in the ways in which complex systems adapt. Such adaptive responses can be best understood in the context of the variation inherent to evolutionary processes. Biological systems are functionally degenerate: they develop different structural configurations to perform an equivalent function (Deacon, 2010; Edelman & Gally, 2001; Green et al., 2006). We recognize such degeneracy in our everyday lives: We can wave a greeting with one arm or the other. Likewise, just as we can use different expressions to communicate a particular meaning, proficient bilinguals who habitually codeswitch explore degeneracy cross-linguistically by seeking alternative means to convey their intentions (Beatty-Martínez, Navarro-Torres, & Dussias, 2020). Language regions in the brain are asymmetrically organized with a left hemisphere dominance for production, but lateralization can dynamically shift for comprehension in adult L2 learners (Gurunandan et al., 2020).

Degeneracy also enables inter-individual variation in cognitive and brain functioning more generally. For instance, in studies examining proactive vs. reactive control tendencies (i.e., whether goal-relevant information is monitored and maintained before the onset of cognitively challenging tasks or whether it is engaged as needed to changing task demands; Braver, 2012), group comparisons between bilinguals and monolinguals can yield similar behavioral outcomes, but electrophysiological and individual differences analyses reveal different strategies as to how each group coordinates both styles of control (Morales et al., 2013, 2015).

The key point is that degeneracy enables recognition of how biological and cognitive systems adapt to external demands. More generally, however, this raises a fundamental issue regarding research practice: critical for advancing our understanding of language and the brain is the development of ideas and theories that can help us understand the nature of such variation.

2. Dissertation aims

Although variation in how individuals process language and engage cognitive processes has long been a topic of interest in language and psychological sciences, it is not often the case that researchers examine variation to generate principles about a particular cognitive phenomenon (see Mashburn, Tsukahara, & Engle, 2020 for an example involving working memory). Instead, individual differences are often treated as an interesting, though peripheral, feature in the process of generating and developing ideas (Fricke et al., 2019). As such, a major aim of this dissertation is to generate a conceptual framework grounded on the study of variation in language processing and cognitive functioning as it relates to research on bilingualism. Central to this aim is the need to reconceptualize the value of replication efforts and scientific prescriptions as being in the service of discovery, rather than the other way around.

The rest of the dissertation is organized as follows: Chapter 3, which consists of a recently published article in *Brain and Language*, attempts to provide an overarching conceptualization of the fundamentals of science concerning discovery in research on bilingualism. A critique of the rationale behind scientific prescriptions in the context of research on bilingualism is offered using illustrative examples from the history of language, cognitive, and neuroscience. I then attempt to articulate some of the conceptual and methodological prerequisites for establishing a viable framework to examine variation in bilingual language processing and its relation to cognition.

In Chapter 4, I present an emerging hypothesis focused on characterizing the cognitive mechanisms that bilinguals exploit when using language, here referred to as *language regulation*, and review some of the relevant research in support of a language regulatory mechanism with the goal of formalizing the hypothesis on empirical grounds.

I then present two empirical studies in Chapter 5 (Part A, consisting of a published article in *Journal of Experimental Psychology: Language, Memory, & Cognition*, and Part B, which includes a novel reanalysis of the data reported in Part A) and Chapter 6 examining the language regulation hypothesis. Following the framework proposed in Chapter 3, I examine individual differences in four groups of Spanish-English bilingual phenotypes, with a particular focus on the relation between language production and cognitive control ability that emerges in each phenotype, to identify the contexts that are more likely to reveal evidence for language regulation. Finally, I conclude by discussing how the conceptual framework, together with the two studies, presented here can contribute to the question of how replication efforts can become fruitful moving forward.

CHAPTER 3: RESEARCH ON BILINGUALISM AS DISCOVERY SCIENCE[1]

"We have to remember that what we observe is not nature in itself, but nature exposed to our method of questioning."

Werner Heisenberg (1958)

1. Introduction

Research on bilingualism generates debate on the neural bases of language that address fundamental questions about language learning (e.g., the role of critical periods), the specificity of language networks (e.g., the nature of any modularity) and their control (e.g., the domain-generality of such control). More recently, specific aspects of the field, namely the putative cognitive and neural consequences (often framed in the form of advantages) of bilingualism, have become a hotspot for controversy tied to the replication crisis in psychology. The critique of this research appears to be broad, addressing issues of power and sample size (e.g., Brysbaert, 2020; Nichols et al., 2020), failures to replicate (e.g., Paap & Greenberg, 2013), noise in samples and methods (e.g., García-Pentón et al., 2016a, 2016b; Valian, 2015), and publication bias (e.g., de Bruin et al., 2015a; but see Bialystok et al., 2015), suggesting that the effects of bilingualism on cognitive and brain functioning are the result of questionable research practices. Consequently, several prescribed remedies,

[1] This chapter, written collaboratively, recently appeared as a published article in *Brain and Language:*

Reference:

Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language*, 222, 105014.

such as large samples (Brysbaert, 2020) and uniform[2] experimental procedures (García-Pentón et al., 2016a, 2016b), have been marketed as solutions (see also Szűcs & Ioannidis, 2020 for an example involving neuroscience more generally). However, such critiques and remedies, though well intended, often fail to place discussions in the broader context of science and its function throughout history. This raises the question of how the implementation of compulsory prescriptions would come to affect research on bilingualism more generally.

Here we argue that the remedies and prescriptions put forward are deceptively simple and place us on a misleading path as they are based on a mischaracterization of the fundamentals of the scientific endeavor. While this paper is geared toward discussing current issues in research on bilingualism, we necessarily draw from the history of science to make the argument self-evident. Our position is that both large samples and conventionalized methods are important, but their role needs to be understood in the context of science as a discovery process, in which research findings are generated through interrelated iterations of exploration and falsification, which in turn lead to new insights and allow for the formulation of new questions. Fundamental to this process is the generation of *variety*[3] that permits incremental advance. The generation of variety serves

[2] We use "uniform" and "uniformity" to describe the hypothetical state in which scientific practice would require the application of a single idealized methodology and/or method to assess replicability of an effect.

[3] We use "variety" to reflect what allows science to act as a discovery process (i.e., diversification in the application of ideas, methods, and scientific practices), as opposed to "variation", which refers to a number of different referents in the world such as the interactional contexts of language use, within-language variation, typological similarities and differences between languages, individual differences, and an individual's response to encountered variation. We therefore use a more specific term for clarity's sake when occasion demands.

two purposes: to identify reliable signals in the noise of our observations and to allow the formulation of effective theories and constructs about our world. Hypotheses, for instance, that the shape of the head is correlated with psychological traits (Simpson, 2005) or that bilingualism negatively impacts intelligence (Peal & Lambert, 1962), are discarded along the way. Constraints on the exploration of variety, such as those imposed by prescriptive remedies (e.g., keeping experimental designs as simple as possible), hinder the discovery process and so it is imperative in our view to ensure that methodological injunctions and publication practices are understood within the context of science as a discovery process.

The remainder of the paper is organized as follows: in the next section (*Why a prescriptive science is problematic*), we provide a critique on both practical and conceptual grounds of the rationale for power and uniformity prescriptions. We then consider the implications for research practice in bilingualism (*Articulating the research enterprise of bilingualism*), where we emphasize the value of rich characterization of the sample, practices that enable the assessment of interactions rather than main effects on their own, and the application of sensitive tools. The implication is that without appropriate characterization, and without research practices and tools that lead to effective signal extraction, replication and large samples may be void of scientific interest. In both sections, we illustrate the manifestation of science as a discovery process with a range of past and contemporary examples drawn from research on bilingualism as well as from other fields. I necessarily draw on a range of examples, including those outside bilingualism, because these points are not unique to research on bilingualism; rather, they reflect a healthy and productive scientific enterprise. We do not argue against the importance of replication, the analytic value of Big Data, nor the application of sensitive and conventionalized research tools. Rather, we suggest that the application of method should be grounded in science as a discovery process.

2. Why a prescriptive science is problematic

We proposed above that prescriptions to remedy poor research practice fail to adequately acknowledge science as a discovery process. Curiously, in applying these prescriptions to research on bilingualism, the analogy invoked is bringing an image into focus: just as glasses improve blurry vision, larger samples have been claimed to increase the resolution of data (Brysbaert, 2020). Similarly, methodological uniformity transforms haziness into a well-defined picture (García-Pentón et al., 2016a, 2016b). Such analogies, rhetorically persuasive perhaps, are misleading. There is a sensible motivation to establish the stability of effect sizes for a given class of data (see Lorca-Puls et al., 2018 for actual rather than simulated data in the context of the relation between brain damage and speech articulation), but it is *signal quality*, our ability to detect a relevant signal from noise, that is key, not sample size or uniformity *per se*. Below we comment on four points to illustrate why power and uniformity prescriptions are insufficient for effective signal extraction.

2.1. Ambiguity is independent of power

Studies of individuals who speak two or more languages have demonstrated a range of consequences for cognition (see Bialystok, 2017 for a review) but controversy surrounds some of these effects. Many large sample studies have yielded null results (Antón et al., 2014; Dick et al., 2019; Duñabeitia et al., 2014; Kałamała et al., 2020; Nichols et al., 2020;

Paap et al., 2018) and other meta-analyses report inconsistency (Anderson et al., 2020; Donnelly et al., 2019; Grundy & Timmer, 2017; Lehtonen et al., 2018; Mukadam et al., 2017; Schroeder, 2018; Sulpizio et al., 2020). On the face of it, such reports have called putative bilingualism effects into question. However, if statistical power were indeed the solution to ambiguity, then we would expect greater consistency across studies with large samples. Problematically, from a naive prescriptive approach, other large sample studies do report effects of bilingualism (Bak et al., 2014; Hartanto et al., 2018; Santillán & Khurana, 2017). Are these latter studies like "black swans" reducing our belief in the generalization that "all swans are white"?

We need not take the current impasse at face value. Consider a contemporary example. The COVID-19 pandemic made it urgent for scientific communities to address a critical question: does the human body develop long-term immunity to the virus? While some large-scale studies suggest that it does (e.g., Iyer et al., 2020; Wu et al., 2020), other large-scale studies show that the effects are limited (e.g., Liu et al., 2020; Pollán et al., 2020). Curiously, it is the collection of single-case patients with reinfection (e.g., Tillett et al., 2020) that initially became more decisive in addressing this question. The point here is that, without an understanding of the boundary conditions of an effect, power, in the form of large sample studies, does not, on its own, improve our ability to extract a signal. In fact, as reinfection cases suggest, and as it is further illustrated below, small-n studies that exploit the features of the sample can be more informative than studies with poorly characterized large samples, which are bound to increase noise in our signals[4].

2.2. Discovery and the power of the small

The history of science suggests that we should recognize the value of the small sample to increase signal, as illustrated in the case of the discovery of penicillin.[5] Research on bilingualism, as well as cognitive and neuroscience research more generally, also attests to this point. Consider a fundamental question: How does the brain adapt to input deprivation? One claim is that brain specialization is determined by input senses. Vetter and colleagues (2020) examined brain activity in the primary visual cortex in healthy blindfolded (n = 10) and congenitally blind (n = 5) individuals while listening to natural sounds. Using multivariate pattern analysis, they found that the blindfolded participants activated the primary visual cortex in response to the sounds despite not having access to visual input in the moment. Remarkably, the same pattern of activation was observed in the congenitally blind group, suggesting that it is not sensorial input *per se*, but rather, the

---

4 As a related point, it is important to note that limitation on inference from poorly-characterized data is not overcome by meta-analyses of studies using such data.

5 Alexander Fleming had searched for antimicrobial agents for years before recognizing the chance finding in a petri dish that led him to examine its anti-microbial properties on mice. But development requires a community of practitioners: it was a decade later before the drug was purified by Florey, Chain, and Heatley in Oxford – their work made urgent by war. Furthermore, the significance of a finding is a community-agreement. The first patient treated (see Barrett, 2018) was Constable Albert Alexander, who had developed sepsis. His immediate recovery was remarkable, but the original penicillin formulation was not optimal, and he died as it was excreted too rapidly. Despite the shortcomings, the constable's remarkable temporary reprieve was sufficient to convince the team (a community of researchers) that a cure would have been possible if only sufficient drug could have been made.

tasks performed by a brain region, that shape brain specialization. It is not sample size, but signal quality that is key here.

Neuropsychological data, typically based on a small number of cases, have been instrumental for our understanding of memory systems[6], but have also been critical to the emergence of research on language control in bilinguals, as they establish the face validity of the distinction between language networks and their control (Green & Kroll, 2019). For example, S.J., a Friulian-Italian speaker (Fabbro et al., 2000), had intact clausal processing for speech comprehension and speech production in both languages, combined with an inability to avoid switching inappropriately in a conversation (e.g., into Friulian when speaking to an Italian-only speaker). More complex control problems reveal dissociations between speech production in one language and translation into it, as exemplified in the alternate antagonism and paradoxical translation of two bilingual patients (Paradis et al., 1982), modelled narratively in Green (1986) and neurocomputationally in Noor et al. (2020). Such cases pave the way for neuroimaging research on the nature of recovery in bilingual aphasia in which we can ask, for example, whether recovery depends on perilesional activation or the use of a previously inhibited alternative network.

A final example exploits the presence of bilingualism in two different modalities – speech and sign language. Hearing bimodal bilinguals are a small population of speakers.

---

[6] For instance, once there was a theory that entry to long-term memory required an intact short-term memory. The theory was rendered less tenable by an n = 1 –a patient with a severely damaged short-term memory but an intact long-term memory (Shallice & Warrington, 1970). Conversely, the discovery of patients (n = 6) with damage to long-term memory but relatively intact short-term memory (e.g., Baddeley & Warrington, 1970) undermined proposals that short-term memory is the activation of representations in long-term memory.

They are typically either children of deaf adults or sign interpreters. Bimodal bilinguals are able to do something that is impossible to do with two spoken languages, namely speaking one language while simultaneously signing another (i.e., code-blending). Because of this feature, bimodal bilingualism provides a unique opportunity to test claims about how the bilingual's languages are controlled. Initial naturalistic production data from two children (Petitto et al., 2001) and 11 adults (Emmorey et al., 2008) showed that bimodal bilinguals strongly prefer code-blending over switching between sign and speech, suggesting that combining the two languages is relatively free of control demands. More recent work using magnetoencephalography (Blanco-Elorrieta et al., 2018; see also Emmorey et al., 2020 for converging behavioral evidence) confirmed this finding, but also showed that increased cognitive effort is required when bimodal bilinguals switch out of a code-blend to either language alone, suggesting that it is the disengagement of one language to switch into the other that requires active control.

The point is that, so long the data are adequately characterized (see 'A rich characterization of the sample and an identification of boundary conditions' subsection) and measures are sensitive (see 'Realizing signal extraction' subsection), a small sample, even a single case, that exploits the special properties of a particular population allow for effective signal extraction that can generate new observations and move the field forward. This is not to say that large-sample studies cannot be equally informative, or that these findings should not be replicated, but that the force of the evidence is not based on statistical power alone.

2.3. Discovery acts on variety

Science acts as a discovery process with Darwinian-like properties, except it also possesses a time-binding property in which earlier ideas and methods can be recruited at a later point in time.[7] Just as natural selection depends on biological diversification to ensure the continuation of evolution, science relies on variety to ensure incremental improvements in our signal-extraction abilities. Progress, in the form of new discoveries and insights, is made by the gradual accumulation of patterns that emerge over distinct data and methods, a process that William Whewell referred to as *consilience* (Laudan, 1971).

In some cases, consilience is relatively straightforward: Converging evidence is obtained from variations of a method (see Green & Abutalebi, 2015 for an example of left caudate involvement in language control). For instance, in examining the question of whether bilingualism changes the engagement of control processes, Wu and Thierry (2013) found in a group of Welsh-English bilinguals a modulation of Flanker performance by experimentally inducing a shift in the language context. Using a novel paradigm in which Flanker was interleaved with words from Welsh or English (single-language context) or both languages (dual-language context), they showed that exposure to words in a dual-

---

[7] Consider the case of Hockett (1985), who hypothesized that hunter-gatherer societies showed a marked lack of labiodentals (e.g., /f/ and /v/) because these incurred greater articulatory effort with their diet-induced edge-to-edge bites. The hypothesis, deemed a *just-so story*, was widely refuted at the time (see Brace, 1986 for a commentary on the matter) based on apparent inconsistencies between the decline of the edge-to-edge bite and the development of agricultural and food processing technologies. Using converging methods from paleoanthropology, linguistics, and evolutionary biology, Blasi et al. (2019) revisited Hockett's conjecture almost three decades later and provided evidence for how changes in fundamental aspects of the ecology (dietary and behavioral practices concerning what food we eat and how we process it) enriched human sound systems by enabling the innovation of a new class of speech sounds.

language context led to greater electrophysiological efficiency in Flanker performance. Since then, several studies have also reported electrophysiological Flanker modulations using variations of the paradigm in different bilingual populations (Bosma & Pablos, 2020; Jiao, Grundy et al., 2020; Jiao, Liu et al., 2020), suggesting that the effect reflects a more general feature of bilingualism. This is an important discovery not only because it shows how control processes adapt to the language context, but also because it makes a more general point that the relative involvement of control processes on a particular task will depend on the control state of an individual at a particular time (see Hsu et al., 2020; Salig et al., 2021 for an elaboration of this argument).

In other cases, consilience requires us to bring together evidence from different methods and populations. Consider the claim that language processing is determined by factors unique to the language system. Converging evidence from neuroscience suggests that domain-general processes also play a vital role. For instance, in studies examining monolingual brain activity across a variety of linguistic and non-linguistic conflict-related tasks (Hsu et al., 2017), co-localization and functional connectivity analyses reveal that, although activation of the Multiple Demands system varies across tasks, engagement of the left inferior frontal gyrus is constant across tasks while also co-activating with other task-specific networks. Research on bilingualism, too, attests to this idea (Nair et al., 2021). In proficient bilinguals, brain potentials reveal that the ability to recover from prediction errors during L2 sentence reading is mediated by individual differences in control ability, but this effect depends on L1 verbal fluency (Zirnstein et al., 2018). The interaction between control and fluency suggests that successful L2 prediction may depend on language-related processes that are partially overlapping with more domain-general control processes. It is the coordination, not the presence or absence, of particular processes or brain regions that is relevant (see also Bialystok, 2011; Morales et al., 2013).

What makes these ideas (i.e., that control processes are state dependent, or that language draws from both domain-general and language-specific resources) compelling is not the ability to replicate a finding using the same method ad infinitum. Rather, it is the fact that we can identify converging patterns *despite* the use of different tools, procedures, and populations with varying sample sizes, all of which might seem to work against us but may in fact improve identification (i.e., signal quality). Such insights allow us to ask new and more useful questions. That is why deep insights about language and the brain emerge through the application of variety.

2.4. Focus is meaningless without context

Analogies invoking focus of an image via large samples and uniformity minimize the conceptual basis on which an observation is made and are fundamentally misleading for a simple reason. We only know the significance of increased focus because we already know the picture (i.e., the conceptual ground). Experience generates the conceptual ground for our everyday lives: we learn to recognize different objects and entities through our ability to interact with them over time. But the conceptual ground for the processes and causal mechanisms underlying the brain and behavior are typically unknown. Science as a discovery process fundamentally concerns the identification of effective theories and constructs of those unknowns. Such theories are based on our justified true beliefs given the evidence and are an intersubjective agreement about that evidence. Theories contest for that agreement. They are necessarily an intersubjective agreement because our senses and scientific tools do not provide immediate access to the physical world[8], as the introductory quote by German physicist Werner Heisenberg suggests. Thus, we come to know the significance of increased focus not by power or uniformity, but by generating and exploring the conceptual ground, as Figure 1 gently illustrates. The problem lies not in the pursuit of statistical power or conventionalized tools *per se*, but in assuming that increasing power or achieving uniformity will generate a picture. As the Blind Men and an Elephant parable suggests, a thousand blind people inspecting separate parts of an elephant will yield an enormous effect size for their one bit but yield zilch for totality!

[8] How does the brain reconstruct the physical world? Salzman et al. (1992) showed that microstimulation of neurons in the middle temporal area selectively distorts motion perception in monkeys. For instance, when applying microstimulation to neurons that selectively respond to objects with upward motion, the direction of motion reported by the monkeys in a direction discrimination task is upwards even when the physical stimuli are projecting downward motion, suggesting that perception is fundamentally abstracting, rather than merely reproducing, the physical world. It is in this sense that our observations are fundamentally an interpretative act (Barret, 2017).

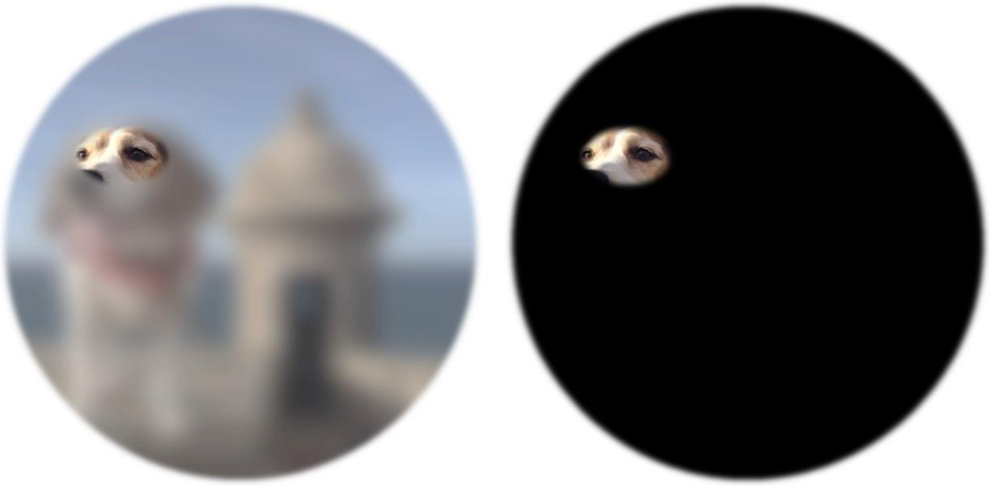

Figure 1. Figure 1a (left) and 1b (right). We formulate our hypotheses in the context of discovery – a possible picture of the world (1a). Bringing an event into focus without context (i.e., testing a hypothesis without a conceptual ground) is meaningless (1b).

3. Articulating the research enterprise of bilingualism

The key question is how we generate the conceptual ground for effective theories and constructs on research on bilingualism and its consequences. We do so by recognizing that language and the brain are byproducts of evolutionary and ecological processes. Such recognition is a generator of the expertise and intuitions for researchers and can play an important role in recognizing the significance of a chance observation or novel finding just as experience furnishes the hunches of everyday life (Bowers et al., 1990). We propose two key factors. First, pre-existing evolutionary older systems are coordinated and the use of these systems exapts[9] mechanisms for the control of language and action more generally (Stout & Chaminade, 2012). Bilingual speakers necessarily must select and control the language of use (e.g., Green, 1986, 1998; see Pliatsikas & Luk, 2016 for a review of data on the overlap). Second, and critical for research on bilingualism, there is a need to characterize the socio-cultural niche in which speakers act. To this end, we emphasize the need for research practices and tools that provide a rich characterization of the participant sample, and ultimately envision a research enterprise focused on the identification of bilingual phenotypes. Under this notion, interactions become of high relevance, and questions eliciting binary outcomes (e.g., is there a bilingual advantage?) become inadequate as they mask the richness of the science. Below we consider each of these points more carefully.

3.1. A rich characterization of the sample and an identification of boundary conditions

The interactional contexts of language use constrain which languages can be used and how they can be used. Characterizing speakers in terms of their habitual community practices, as well as their trajectory through particular contexts, is essential if we are to understand adaptive change in the context of degeneracy (see Chapter 2). We refer to such characterization as bilingual phenotyping (Adamou, 2010; Beatty-Martínez & Dussias, 2017; Beatty-Martínez, Navarro-Torres, Dussias et al., 2020; Poplack, 1998). Such phenotyping helps determine the boundary conditions for any adaptive effect because we

[9] Critical to evolutionary biology is the distinction between "adaptation", features that are the byproduct of natural selection, and "exaptation", features that attain a new function for their present role regardless of their evolutionary history (Gould & Vrba, 1982). See footnote 6 for an illustration.

might predict an effect for one phenotype but not for another (Bak, 2016). On the grounds of degeneracy (see footnote 9), we can ask what kinds of cognitive and neural changes might be expected given the demands of particular contexts on language and the control processes supporting it (DeLuca et al., 2020; Green & Abutalebi, 2013). We propose that a plausible answer to this question will require the application of ethnographic practices in brain and cognitive sciences (see Billig, 2020; Torres Cacoullos & Travis, 2018 for illustrations in psychology and variationist linguistics, respectively) tied to multi-lab collaborations (Leivada et al., 2020). We illustrate with a comparison of two Spanish-speaking locations, San Juan, Puerto Rico, and Granada, Spain, to highlight three key aspects of rich characterization using ethnography.

First, we need to be aware of the diachronic processes that have shaped the culture and history of a community. Granada is located in the community of Andalusia, officially considered one of several monolingual autonomous communities of Spain. Despite having a long-standing influence of Arabic culture, Andalusia has historically perpetuated ideologies tied to monolingualism, especially throughout the 20th century under the Francoist regime (see Lorenzo & Moore, 2009). As a result, foreign-language prevalence has remained lower in Andalusia relative to the rest of Spain and Europe (Ministerio de Educación, Cultura, y Deporte, 2012), creating fewer opportunities for other languages such as English to influence everyday language use. By comparison, Puerto Rico is the byproduct of a rich colonial history spanning across five centuries until the present (see Guzzardo Tamargo et al., 2018). Although Spanish had been the established language following four centuries of Spanish colonial rule, the island became a US territory after the Spanish-American war at the end of the 19th century and continues to this day to be a non-incorporated territory of the US. Unlike Granada, the history of Puerto Rico created conditions in which American culture would become highly influential for the already established Hispanic culture, especially in the metropolitan area of San Juan.

The observation and description of current community practices in a well-defined speech community[10] (Labov, 2001) leads to the second feature of rich characterization. Determining how bilinguals' languages are habitually used (e.g., whether a speaker has extensive experience codeswitching or not) is important, but doing so requires an understanding of how the global environment of a community supports those practices. For instance, Spanish-English codeswitching is often a prominent form of communication among proficient bilinguals in San Juan whereas speakers from Granada tend to use their languages separately (see Beatty-Martínez & Dussias, 2017; Beatty-Martínez, Navarro-Torres, & Dussias, 2020; Beatty-Martínez, Navarro-Torres, Dussias, et al., 2020 for a more comprehensive characterization). But in terms of interactional demands, the key difference between these two communities is not the frequency of switching behavior *per se*[11]; rather,

[10] Speakers can form part of stable speech communities, such as San Juan and Granada (see Torres Cacoullos & Travis, 2018 for another example in Albuquerque, New Mexico), where most individuals are members of the community, but they can also live in more dynamic and/or cosmopolitan communities, such as many major cities (e.g., London, Montreal) and some countries (e.g., Singapore). The distinction is important because it can help us infer the range of possible phenotypes and interactional demands that are likely to emerge in a given location.

[11] Notably, codeswitching is a relatively infrequent behavior even among habitual codeswitchers (Fricke & Koostra, 2016), it can also be observed even among non-habitual codeswitchers such as bilinguals from Granada (see Table 9 in Beatty-Martínez & Dussias, 2017), and great discrepancy can exist even among bilingual communities that display habitual codeswitching (see Poplack, 1987 for a contrast between French-English bilinguals in Ottawa vs. Spanish-English Puerto Rican bilinguals in New York). The cognitive consequences of codeswitching in spontaneous discourse remain to be determined, but for now I make the

it is the fact that, given the history and culture of each location, one context (i.e., San Juan) enables speakers to use both languages more openly and opportunistically with little-to-no interactional costs, whereas in the other context (i.e., Granada), there is a strong tendency to expect the use of Spanish (the L1) most of the time, creating constraints as to when speakers expect the use of the L2. This is not to say that codeswitching is not critical to understand how bilinguals control their languages (see Adler et al., 2019; Green, 2018; Hofweber et al., 2019 for how control processes may be engaged during codeswitching), but that the relative involvement of control processes during different kinds of speech acts may also depend on the demands imposed by the global environment. Characterizing speakers in terms of their habitual community practices is vital to understand such dynamics.

The final feature of rich characterization relates to changes in an individual's trajectory of experiences. While some bilinguals live in homogeneous communities where the language dynamics are relatively stable over the lifespan, other bilinguals undergo radical shifts in language use at particular time points (see Kubota et al., 2020; Pallier et al., 2003 for examples involving international returnees and adoptees). To illustrate, speakers may initially grow up in a home environment where a minority language (e.g., Spanish in the US) is used, but can then become educated and socialized in the majority language of the community (e.g., English in the US) during childhood. As such, some bilinguals (a.k.a., heritage speakers and indigenous-speaking bilinguals) may grow up and become educated

---

point that an aggregate lump of codeswitchers vs. non-codeswitchers is misleading if rich characterization is not provided.

in a context where the L1 is the majority language but then shift to an environment where the L2 becomes the dominant language (e.g., Garraffa et al., 2015; Garraffa et al., 2017; Bonfieni et al., 2019; Polinsky & Scontras, 2020). A similar case is observed with young adults seeking higher education who relocate to a new environment (e.g., a foreign country) with a different predominant language (Beatty-Martínez, Navarro-Torres, Dussias et al., 2020). These shifts in language immersion status are likely to generate unique adaptive brain responses. Indeed, emerging evidence suggests that heritage speakers' initial minority-language experience has long-term consequences for language processing in the majority language (Bice & Kroll, 2021) and that contexts with high linguistic diversity (Gullifer et al., 2018) or L1-to-L2-immersion shifts (Beatty-Martínez, Navarro-Torres, Dussias et al., 2020) may trigger a novel adaptation of control processes in the form of proactive control engagement (see also Blanco-Elorrieta & Pylkkännen, 2018 for a similar observation regarding dual-language contexts).

Although we still know little about the boundary conditions of these effects, the point is that in a main-effect group analysis, different groups of speakers would be assumed to represent the same underlying population of bilinguals (see Weyman et al., 2020, for an illustration), despite having remarkably different community practices and/or individual trajectories that become evident through rich characterization. In making such an assumption, we may miss critical information that can change our conclusions. Hence, research on bilingualism[12] is likely to benefit more from small sample studies with rich characterization.

3.2. Beyond main effects and binary oppositions

Under a traditional lens in research on bilingualism, idiosyncratic patterns are typically discarded as random noise and complexity is equated with complication. Despite several notable critiques to this approach (Baum & Titone, 2014; Fricke et al., 2019; Luk & Białystok, 2013; Tanner et al., 2013), binary classifications and group comparisons, together with recommendations to keep experimental designs as simple as possible (Brysbaert, 2020), continue to dominate much of research on bilingualism, forcing discussions into a binary opposition not unlike those that have recently characterized the consequences of bilingualism (Nichols et al., 2020; c.f. Leivada et al., 2020), as well as psychological research more generally (Newell, 1973). But given the degenerate nature of biological and cognitive systems, solely main effect 'yes' or 'no' questions are unhelpful. As Bronfenbrenner (1977, p. 518) noted, "in ecological research, the principal main effects are likely to be interactions." To illustrate this point, we return to the study by Zirnstein et al. (2018; see 'Discovery acts on variety' subsection).

Following a main-effect analysis, Zirnstein and colleagues found different electrophysiological responses for L1 and L2 speakers to predictions errors during

[12] Although the focus here is on bilingualism, this proposal can serve a role in establishing the value of rich characterization and phenotyping procedures more generally. Evidently, the study of variation in language and cognition is central to any population of speakers, as has been established by research examining learning in monolinguals from different linguistic environments (e.g., Bice & Kroll, 2019), individual differences in language processing (e.g., Beatty-Martínez, Bruni et al., 2020; Pakulak & Neville, 2010; Tanner & van Hell, 2014), as well as the consequences of dialectal experience for lexical and grammatical processing (e.g., Clopper, 2014; Squires, 2014).

sentence reading: Only L1 speakers showed reliable electrophysiological costs when encountering semantically unexpected words. At first glance, it would be tempting to conclude that L2 speakers were unable to generate predictions, consistent with previous claims (e.g., Martin et al., 2013; Grüter et al., 2017). But instead of asking whether L2 speakers can generate predictions (a 'yes' or 'no' question), one can ask about the cognitive processes that enable prediction in the first place. Upon examining individual differences in control and verbal fluency in both groups, Zirnstein and colleagues identified a more complex, but also more insightful, picture. First, both L1 and L2 speakers recruited control processes to recover from prediction errors (i.e., increased control ability related to reduced prediction costs). But for L2 speakers, as mentioned previously, there was an interaction between control and L1 fluency, such that increased L1 fluency related to larger prediction costs in the L2. This suggests that L2 speakers had to overcome the challenge of regulating the L1 in order to engage prediction mechanisms in ways comparable to L1 speakers. More critically, the interaction reveals that the absence of an electrophysiological response in the L2 group stemmed from an aggregate of bilingual phenotypes with different configurations of control and regulatory engagement. If we had only asked whether L2 speakers can generate predictions, we might have come to a different conclusion (see also Pulido, 2021; Tanner & van Hell, 2014 for illustrations with adult L2 learners and monolinguals, respectively).

The point is that a simple main-effects approach focused on attaining large samples or replication-via-uniformity would disregard the fact that the form of language and cognitive engagement varies across individuals. Arguably, there may be some important main effects, but the way to identify them is by first seeking out meaningful interactions that are informed by a rich characterization (see Rohrer & Arslan, 2021 for a discussion on the application of interactions). As a discovery process, science benefits from relatively open-ended questions such as "how do these regions in the temporal lobe dissociate during different tasks?" or "what are the possible range of phenotypes that can emerge in this community?". As Calhoun and Bandettini (2020) point out, such questions cast an effective net in making sense of large amounts of data[13].

3.3. Realizing signal extraction

Rich characterization, as well as the framing of our questions, is vital for effective signal extraction, but just as important is determining task and test sensitivity (e.g., for a given sample, to what extent do we expect a non-verbal task to tap executive processes used in language control so that any putative adaptive response of language experience could be realized?). We comment on four aspects. First, task sensitivity might require a revised conception of the task construct that it is designed to tap, such as using within-subject paradigms that allow us to induce different control states and track how they are engaged during language processing (e.g., Adler et al., 2020; Hsu et al., 2020; Navarro-Torres et al., 2020; Salig et al., 2021), as opposed to exclusively relying on aggregate executive function

[13] Historically, Leibniz (1690/1951, cited in Gigerenzer, 1991, p. 254) likened scientific enquiry to "an ocean, continuous everywhere and without a break or division". Divided later by Reichenbach (1938) into two seas (the contexts of justification –hypothesis-testing— and the contexts of discovery –the generation of novel ideas): some have argued in favor of a sharp distinction between hypothesis-testing and exploration (Mertzen et al., 2020), while others have argued that the only legitimate scientific practice is hypothesis-driven (Kullmann, 2020). But as we have argued, hypotheses arise in the context of an evolving understanding, and can vary in specificity, which is why there is no sharp division between the two contexts. Science as a discovery process entails the mingling of both creativity and empirical verification.

measures (e.g., Stroop effects) that likely mask degenerate patterns. As such, the insensitivity or the appropriateness of the test to tap into control processes engaged in language use limits their relevance for exploring any putative wider effects on non-verbal control tasks, as acknowledged in recent papers reporting data based on more richly characterized large samples (Gullifer & Titone, 2020a; Kałamała et al., 2020; Kheder & Kaan, 2021).

Second, deepening theoretical understanding also requires that we understand the totality of performance for which we need to consider data from a number of modalities – some of which may be more sensitive to the effects of interest than others. For instance, brain measures may better capture some aspects of early L2 learning than overt behavioral responses (e.g., Bice & Kroll, 2015, 2019; Kurkela et al., 2019; McLaughlin et al., 2004), although in other cases, both brain and behavior converge (see Li et al., 2014 for a review on imaging studies). Multi-lab collaborations can be an effective way to explore the boundary conditions of such issues for a given task and sample. Notably, however, the goal should not be replication *per se* –using paradigms that are simple enough to easily reproduce— but to see collaborations in ways that are designed to exploit variation across different labs and different locations (Leivada et al., 2020).

Third, it is important that we use measures that are reasonably commensurate with the questions being asked. For instance, with respect to determining how bilinguals' languages are habitually used, self-reported data can be informative (e.g., Gullifer & Titone, 2020b) but likely insufficient in the absence of conversational data that correspond to the vernacular of the speech community (Labov, 1984) or that reflect engagement of different attentional/control states when bilinguals shift between different modes of communication (Green, 2019).[14] The application of Network Science (Tiv et al., 2020) and Information Theory (Gullifer & Titone, 2020b; Feldman et al., 2021) practices can also be of high value regarding effective phenotyping as they can help us establish correspondence between individual differences in language experience and the extent to which those trajectories reflect (or deviate from) more general community practices.

Finally, more sensitive data analysis practices are likely to be more revealing of individual differences and degenerate patterns. For instance, using ex-Gaussian distributions (Sundh et al., 2021; von Bastian et al., 2020), delta plots (Morales et al., 2013), or Bayesian mixture models (Ferrigno et al., 2020) to infer the possible range of strategies in a given task, rather than simply averaging effects for a condition. Individual differences also allow us to construct generative models of behavior and neuroplasticity (see Parr et al., 2018 for an example in neuropsychology) which can be used to computationally model, say, neuroplastic effects of different interactional contexts given a set of behavioral profiles. Further, within a large sample, there may be different phenotypes and we need to be able to explore and characterize these using data-driven techniques such as multivariate statistics (e.g., cluster and/or factor analyses; Hartanto & Yang, 2020; Rodriguez-Fornells et al., 2012) if the data are sufficiently rich to detect different profiles. However, although individual differences offer an opportunity for effective phenotyping, they potentially

[14] For example, the conversational topics centered on individuals' personal experiences and that involve in-group members from the same speech community have been shown to increase the likelihood of codeswitching in informal contexts fourfold (Poplack, 1983).

involve the same risks as those observed with main-effect practices in the absence of a rich characterization tied to well-defined speech communities.

All four aspects are pertinent to advance. Further, in some cases, and to reinforce our earlier point (see 'Discovery and power of the small' subsection), only small samples may be feasible and yield decisive evidence. For example, localization of phonemic restoration effects in the auditory cortices is best achieved through the high signal-to-noise ratio afforded by electrocorticography arrays implants for clinical purposes (see Leonard et al., 2016), and so establishing convergence in bilingual speakers in two languages may sometimes require small samples with rich characterization (see "A rich characterization of the sample and an identification of boundary conditions" subsection). In short, replication, conventionalized tools, and large samples have value, but their role in the discovery process in research on bilingualism hinges on conceptual, experimental, and analytic advance.

3.4. Concluding remarks

In this chapter, we have emphasized the community-value of incremental contributions via science as a discovery process against the enforcement of prescribed remedies, such as pre-determined sample sizes and/or methods, because we trust in the basic integrity of participants in the enterprise of research on bilingualism and ultimately in the self-correcting dynamic of science itself.

From the point of view of ensuring variety on which the quasi-Darwinian process of science can act, we require the publication of possibilities (e.g., sensitive tasks geared to testing specific processes, rich characterization to identify phenotypic variation, and non-binary questions that enable the exploration of interactions) that may or may not lead to deeper understanding. Replication and reproducibility efforts, important as they are to the scientific enterprise, need to be in service of such aims to advance. As the National Academies of Sciences, Engineering, and Medicine recently acknowledged: "The goal of science is not to compare or replicate [studies], but to understand the overall effect of a group of studies and the body of knowledge that emerges from them" (Fineberg, 2019, as cited in Miceli, 2019). Like the brain, bilingualism is complex, and we are far from having a complete understanding of the boundary conditions of previously reported findings for replicability to be fruitful on its own. And while some have proposed that such understanding lies in the data itself (de Bruin et al., 2015b), we make the point that the answers ultimately lie in the characterization (i.e., the intersubjective agreement) of the data.

Finally, in establishing the need to view science as a discovery process, we wish to return to the question raised in the introduction of how compulsory prescriptions would come to affect research on bilingualism, as well as psychological and neuroscience research more generally. If we choose to allow prescriptions to dominate the scientific enterprise, then we must ask how they will come to shape not only the environment in which research is currently being conducted, but also how they will shape the minds of young and early-career researchers, and ultimately, whether we are willing to live with the consequences of those choices. Thus, in articulating the research enterprise of bilingualism, we hope to contribute to the establishment of a viable future research enterprise more generally.

CHAPTER 4: THE LANGUAGE REGULATION HYPOTHESIS

1. Introduction

As alluded to in Chapter 1, the bilingual language network is highly interactive and dynamic (Kroll et al., 2012). But despite the parallel activation that bilinguals experience when using their languages, they are typically required to function in one language at a time, even in the case of proficient bilinguals who habitually codeswitch between their languages (Fricke & Kootstra, 2016). Paradoxically, speech errors involving intrusions from the unintended language are relatively uncommon (Gollan, Sandoval, & Salmon, 2011; Poulisse, 1997; Poulisse & Bongaerts, 1994; Sandoval et al., 2010). This suggests that bilinguals must actively rely on a cognitive mechanism that allows them to manage the relative activation of each language in accordance with their proficiency, task at hand, and environment (Zirnstein, Bice, & Kroll, 2019).

Although alternative proposals have been generated to address the issue of how bilinguals achieve this (e.g., Costa, Miozzo, & Caramazza, 1999; Finkbeiner, Gollan, & Caramazza, 2006), the predominant view is that appropriate language selection involves the active engagement of an inhibitory mechanism to control the activation of the languages (Kroll et al., 2008). But as we will see in this chapter, converging evidence suggests that, although inhibition may be an important component, the mechanism involved in the language selection process is likely more complex and dynamic. Here I refer to such mechanism as *language regulation*.

In previous research, the core ideas behind the language regulation hypothesis have typically been framed in terms of language control, where the focus is on examining the extent to which language-switching ability (as indexed by language-switching paradigms discussed in Section 2 below) reflects the same mechanism involved in more general forms of switching behaviors (see Declerck et al., 2021; Timmermeister et al., 2021 for recent illustrations). A fundamental challenge with this approach, however, is that, by relying on a narrow set of tasks (e.g., language-switching tasks) and analytic procedures (e.g., whether language-switching performance correlates with general task-switching performance), it fails to adequately characterize the dynamic and interactive nature of such mechanisms and their function. As such, this approach is unable to provide an account of how regulation is engaged to enable fluent speech or efficient language processing more generally. As others have recently argued (e.g., Zirnstein, Bice, & Kroll, 2019), such a question requires an approach that focuses on assessing language processing across different conditions that impose different demands on the language system, and how those demands come to engage cognitive resources.

Here, language regulation is meant to denote the engagement of a host of cognitive processes in a coordinative manner, some of which may be domain-general by nature, but others which may be more specialized for the language network. From this perspective, the function of language regulation is not solely to actively inhibit activation when switching from one language to another, but to up-regulate and down-regulate information flow from the language networks accordingly. This ensures that the differentiation of competing information, be it cross-linguistically or within a language, is maximized or minimized as needed. Below I briefly review some of the evidence supporting the presence of an inhibitory control mechanism in production, followed by additional sources of evidence from neuroscience and language processing that highlight the dynamic nature and function of language regulation.

2. Language regulation and inhibitory control

In an initial attempt to address the selection problem, Green (1998) proposed the inhibitory control (IC) model. At its core, the IC model stipulates that for bilinguals to speak in one language, alternatives in both languages remain active until domain-general control processes suppress information from the unintended language (see Kroll & Navarro-Torres, 2018; Kroll & Tocowicz, 2005; Meuter, 2005 for more comprehensive characterizations of the model). Since suppression should be more difficult for highly activated information relative to weakly activated information, the model predicts that the more dominant L1 will require greater inhibition to enable the weaker L2 to be selected.

Support for this prediction mainly comes from studies reporting asymmetric switch costs when bilinguals are required to switch between the two languages following externally-driven cues (e.g., naming digits or pictures in one or the other language following a color cue; e.g., Meuter & Allport, 1999). The general finding is that, when performance involves unpredictively mixing the two languages within a block, switching from the L2 into the L1 typically results in slower speech relative to switching from the L1 to the L2, particularly in bilinguals whose L1 is more dominant (see Declerck, 2020; Bobb & Wodniecka, 2013 for reviews; see also Bultena, Dijkstra, & Van Hell, 2015 for an example in comprehension). These findings may seem counterintuitive because the dominant language should be more readily accessible, but they are consistent with the IC model in that active suppression of the L1 should be taking place while using the L2, which in turn spills over to subsequent instances of L1 use (see also Misra et al., 2012 for converging electrophysiological evidence in a blocked naming paradigm).

Although not all studies report such asymmetric switch costs (e.g., Christoffels, Firk, & Schiller, 2007; Declerck, Koch, & Phillip, 2012; Gollan & Ferreira, 2009; Hernandez & Kohnert, 1999), they do consistently report slower speech in the L1 than the L2 when switching takes places in a mixed language setting, even in highly and relatively balanced bilinguals who do not show the asymmetry (Costa & Santesteban, 2004). Further, other studies examining individual differences find that switch costs are predicted by measures of cognitive control typically associated with inhibition and interference suppression, though the patterns of association depend on a number of factors, including the bilingual sample, the direction of the switch, the nature of the task, and the cognitive control measures used (e.g., Li et al., 2021; Linck, Schwieter, & Sunderman, 2012, 2020). Notably, there is some debate regarding the interpretation of these findings, including the extent to which these effects reflect the use of domain-general mechanisms (e.g., Declerck, Eben, & Grainger, 2019), and whether they reflect the involvement of different components of inhibition (Bobb & Wodniecka, 2013). And even though these studies frame discussions in terms of inhibitory control, they begin to suggest that the language selection may involve the engagement of multiple cognitive mechanisms rather than a single inhibitory control mechanism.

Some of the more compelling evidence regarding the complexity of language regulation comes from studies examining brain activity while bilinguals produce words in each language either in a mixed fashion or using a blocked design (Guo et al., 2012; Rossi et al., 2018; Yuan et al., 2021). These studies reveal that when L1-dominant bilinguals name pictures in the L1 following L2 naming, or when naming involves switching between the languages, they activate a widespread number of brain regions. Some of these regions are typically associated with the planning, updating, and selection of information from the language network, together with more domain-general regions involving the control of attention, detection of conflict, monitoring, and interference suppression (see Abutalebi & Green, 2007 for a more comprehensive characterization of the networks). Critically, such distributed patterns of activation are not seen in monolinguals (Rossi et al., 2018). These findings, though typically interpreted in the context of an inhibitory mechanism, are important because they not only confirm the involvement of domain-general resources during language use, but they also suggest that, for bilinguals, the simple act of producing a word in a language likely requires the *coordination* of language-related and domain-general brain networks.

3. The dynamic nature of language regulation

Given the neuroimaging evidence, we can ask what the function of language regulation is with respect to language processing and linguistic skill development more generally. In adult L2 learners, learning to engage regulation may be a necessary step to accommodate the emerging and less proficient L2. For instance, when L2 learners are immersed in an L2 environment via study-abroad programs, fluency and lexical accessibility in the L1 is reduced relative classroom learners (Baus et al., 2013; Linck et al., 2009). Such costs may reflect learners' ability to regulate the influence of the L2 on the L1. Bice and Kroll (2015) showed in beginning classroom learners that sensitivity to cross-language overlap with cognates via the L1 was mediated by individual differences in cognitive control ability (measured using the AX-CPT) and L2 proficiency: those with greater cognate sensitivity (measured electrophysiologically) and more efficient cognitive control generated slower responses in the L1 (measured in a lexical decision task) relative to those with reduced cognate sensitivity and less efficient cognitive control.

Notably, the effects of regulation in learners are not limited to single word phenomena, but may also extend to different sources of linguistic information. For example, better learning outcomes are observed for idiosyncratic verb-noun collocations in the L2 (e.g., *launder money* literally translates to *whiten money* in Spanish) when L2 learners undergo training with collocations that overlap and compete with the L1 relative to those who become exposed to non-competing collocations only (Pulido & Dussias, 2019). Such findings suggest that language regulation initially acts as a desirable difficulty that subsequently enables cross-language interactions (Bogulski, Bice, & Kroll, 2020; Bjork & Kroll, 2015), which, as stated in Chapter 1, are a defining feature of proficient bilingualism.

In more skilled and proficient bilinguals, language regulation may maximize the efficiency through which linguistic information is processed and accessed more generally. Currently, this is best evidenced by research examining variation in bilingual language processing and cognitive functioning (see Zirnstein et al., 2019 for a review). Zirnstein et al.

(2018; see also Chapter 3) showed that sensitivity to prediction errors in the L2, and the ability to recover from them, during online sentence reading is dependent on a systematic interaction between fluency in the L1 and cognitive control ability (measured using the AX-CPT): those with more efficient cognitive control engagement were better able to mitigate the costs associated with prediction errors in the L2, but this was true only for bilinguals with high L1 fluency. In turn, those with high L1 fluency but low cognitive control had increased prediction costs. The interpretation is that high L1 fluency reflects the ability to up-regulate activation of the more dominant L1, a process that was hypothesized to partially depend on domain-general resources.

A second source of evidence similar to Zirnstein et al. (2018) comes from a study by Mercier, Pivneva, and Titone (2014), which focused on examining cross-language activation (i.e. the processing of cognates and interlingual homographs) during sentence reading. Upon examining individual differences, the authors found that higher cross-linguistic proficiency and more efficient cognitive control (measured using a Stroop task) led to greater facilitatory effects with cognates and reduced interference effects with homographs, respectively. Although the findings were interpreted in the context of proficiency effects, unlike Zirnstein et al. (2018), who interpreted the fluency effects in terms of regulation, together they suggest that language regulation may impact the language system in different ways, either through the coordination of up-regulatory and down-regulatory effects for language processing in the L2, or through reliance on down-regulation to manage cross-language influences in real time.

A final source of evidence for language regulation comes from a sentence production study by Navarro-Torres, Dussias, and Kroll (under review) examining variation in the use of a-adjectives (e.g., asleep, afraid, alive), which, unlike typical adjectives (sleepy, frightened, lively), disprefer attributive usage (*the asleep dog* is generally dispreferred over *the dog that's asleep*). The results showed that bilinguals' ability to produce canonical structures with a-adjectives was mediated by a systematic interaction between fluency in the dominant language and working memory: those with higher working memory were more likely to use a-adjectives non-attributively, but this effect was only true for bilinguals with high dominant-language fluency.

4. Predictions from the language regulation hypothesis

Following the evidence described above, there are three key features of the language regulation hypothesis to consider which generate three predictions. First, regulation is assumed to emerge from the coordination of language resources that partially overlap with, but are also distinct from, domain-general resources. This implies that regulation is an emergent process that cannot be accounted for by examining its constituents on their own. It also implies that, by definition, regulation is interactive rather than additive. It is in this sense that we should not expect measures tapping into regulatory processes to reliably correlate with one another. Instead, they should yield interactions that account for significant variance in particular measures of language processing and/or learning. The distinction is important given the debate on whether language control (as measured by language switching paradigms; see Section 4.2) and cognitive control are related, which has yielded mixed results on the correlational front (see Timmermeister et al., 2021 for a review).

Second, even though regulation is assumed to enable higher linguistic skill, it should be distinct, but not entirely independent, from proficiency. This implies that the dissociation between the two should be empirically testable. Further, in proficient bilinguals, more efficient regulation should result in more efficient language processing and greater linguistic ability regardless of the language that is being examined, given that regulation is assumed to impact the extent to which the language network remains accessible.

Finally, the degree to which bilinguals engage regulation depends on the language profile of the bilingual phenotype, as well as on the kinds of demands imposed by the language environment. For instance, for bilinguals who grow up in an L1-dominant environment and who retain L1 dominance throughout their lives, effectively regulating the L1 should be key to attain high linguistic skill in the L2. In contrast, bilinguals who undergo a shift in language dominance (becoming L2 dominant) may more readily need to actively regulate the L2 in order to maintain such skills across the two languages. Further, some contexts of language use should more readily require active regulation, such as L2-immersion contexts (see Linck et al., 2009; Navarro-Torres, Garcia, Chidamabaram, & Kroll, 2019; Zirnstein et al., 2018; Zhang et al., in press) or dual-language contexts where there is a high degree of diversity and/or uncertainty regarding which language to use at a given time (see Gullifer et al., 2019). In the following two chapters, I explore the application of these hypotheses by examining individual differences in language production and cognitive control ability across four proficient bilingual phenotypes.

CHAPTER 5: STUDY 1: LANGUAGE REGULATION AND L2 IMMERSION

PART A[15]

1. Introduction

A major issue of contention in the field of bilingualism research centers on whether bilingual experience confers lifelong changes in cognitive functioning. While there has been considerable evidence supporting the positive consequences of bilingualism across the lifespan (Baum & Titone, 2014; Bialystok, 2017; Hervais-Adelman, Moser-Mercer, & Golestani, 2011; Kroll & Bialystok, 2013), other research has raised concerns about the validity of previous claims, arguing that bilingual effects are observed inconsistently (see Antoniou, 2018, for a review). An issue that may be at the source of this controversy is the complexity of characterizing the bilingual experience. Individuals who learn and use more than one language come to be bilingual in many ways (de Bruin, 2019; Luk & Bialystok, 2017; Pot, Keijzer, & de Bot, 2018), which may in turn have unique consequences for both language processing and cognitive functioning. However, it remains less clear which aspects of the bilingual experience are critical for understanding the observed consequences.

---

[15] Part A of this chapter, written collaboratively as a first co-author, appeared as a published article in *Journal of Experimental Psychology: Language, Memory, and* Cognition. As mentioned in the Preface, its inclusion is intended to provide background for contextualizing the data reported in Part B, which consists of a reanalysis of the published data.

Reference:

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1022.

The present study investigates how different contexts of bilingualism affect the ability to produce words in each language, and whether such contexts modulate the relation between language abilities and cognitive control. We explore the idea that some aspects of lexical access are shaped by habitual patterns of language use (i.e., whether the languages are used separately or interchangeably), while other aspects are shaped by environmental demands (i.e., whether a speaker is immersed in a context that allows the use of the two languages, or whether the context restricts the use of one of the two languages). We argue that these two processes are dissociable to some extent, and that their particular configuration will affect how (and under what circumstances) bilingual language production recruits cognitive control. We note that our main focus is not to ask whether different bilingual groups differ in their cognitive ability as a function of the context of language use, but rather how cognitive resources are engaged differentially to enable proficient spoken production in each language.

1.1. Bilingual Language Production

Actively learning and using a second language has consequences for the language system. There is abundant evidence indicating that bilinguals' two languages are momentarily activated in parallel when the intent is to speak in only one (for reviews, see Costa, 2005; Hanulová, Davidson, & Indefrey, 2001; Kroll, Bobb, & Wodniecka, 2006; Kroll, Dussias, Bogulski, & Valdés Kroff, 2012). As a result, the presence of one language can affect performance in the other language (Kroll & Dussias, 2013; Zirnstein, Van Hell, & Kroll, 2018). In some cases, cross-language activation can result in direct facilitation and/or interference from the non-target language, creating conditions in which speech planning is open to cross-language influences (Bobb & Wodniecka, 2013; Chang, 2013; Kroll et al., 2006).

Similarly, the presence of a second language (L2) seems to introduce subtle but noticeable costs during language production more generally. For example, bilinguals are typically slower to name pictures than monolinguals, even when naming in their native or dominant language (L1), and show larger frequency effects in the slower second language (i.e., the difference in naming performance between high and low frequency words is greater in the L2 than in the L1; Kroll & Gollan, 2014). These observations have been taken to indicate difficulties in language fluency due to reduced functional use of the languages (e.g., Gollan, Montoya, Cera, & Sandoval, 2008) or limited proficiency (Bialystok, Craik, & Luk, 2008). However, more recent work suggests that the slower lexical retrieval abilities and frequency asymmetries in bilinguals might be at least in part a consequence of cross-language interference (Sullivan, Poarch, & Bialystok, 2017). This leaves open the question of how bilinguals successfully regulate the relative activation of both languages to allow fluent speech in each language.

For monolinguals, language production requires cognitive control, particularly when related semantic, lexical, and/or phonological information interferes with the selection of a target representation (Freund, Gordon, & Nozari, 2016; Nozari & Novick, 2017; Shitova, Roelofs, Schriefers, Bastiaansen, & Schoffelen, 2017). However, unlike monolinguals, the choice that bilinguals make in selecting one language is also hypothesized to recruit domain general cognitive processes given the potential for unwanted interference from the non-target language (Abutalebi & Green, 2007, 2016;

Linck, Schwieter, & Sunderman, 2012), although the conditions in which such interference affects speech planning depend on the contextual demands of the task. For example, the ability to produce words in the dominant language has been shown to be sensitive to the order in which the languages are spoken (Misra, Guo, Bobb, & Kroll, 2012; Van Assche, Duyck, & Gollan, 2013), and whether the two languages are mixed or blocked (Bobb & Wodniecka, 2013; Christoffels, Firk, & Schiller, 2007; Meuter & Allport, 1999). At the same time, other studies have shown that the effort devoted to producing words is similar in both languages when bilinguals are given the option to choose between the languages (Gollan & Ferreira, 2009; Gollan, Kleinman, & Wierenga, 2014; Kleinman & Gollan, 2016), suggesting that, in addition to contextual demands, there are aspects of language control that are under the control of the speaker (i.e., deciding which language to speak and/or whether codeswitching is appropriate) that can affect the selection process.

1.2. Bilingual Interactional Context

If immediate contextual/situational demands modulate the availability of each language, then real-world interactional contexts should have notable consequences for language performance and cognitive control, even in highly proficient bilinguals. A recent framework to characterize how distinct social environments may impose different demands on cognitive control for bilinguals has been proposed as the adaptive control hypothesis (Green & Abutalebi, 2013). The hypothesis posits that distinct interactional contexts lead to specific adaptive changes to cognitive control processes. In a single-language context, only one language is used. Codeswitching contexts, in which bilinguals may alternate between stretches of the two languages within a conversation at will, offer opportunities for language integration. Finally, in dual-language contexts, both languages are used in the same environment but typically between speakers. Critically, dual-language contexts are hypothesized to increase the demands on cognitive control processes over and above single-language and codeswitching contexts.

Support for the adaptive control hypothesis comes from studies showing that bilinguals who operate in dual-language contexts exhibit reduced task-switching costs (Hartanto & Yang, 2016) and more efficient conflict resolution (Ooi, Goh, Sorace, & Bak, 2018) than bilinguals in a single-language context (see also Wu & Thierry, 2013). Similarly, two recent studies have shown that increased diversity in language usage across social contexts is related to better behavioral cognitive performance in older adult bilinguals (Pot et al., 2018) and greater neural connectivity between brain regions associated with cognitive control engagement (Gullifer et al., 2018). Critically, these effects did not depend on language proficiency or age of acquisition. More generally, these studies suggest that the expertise bilingual speakers gain in their everyday conversational practices will differentially affect cognitive and neural functioning, and that monolingual comparisons may not necessarily provide a comprehensive understanding of such dynamics.

One limitation that is often associated with studies examining the cognitive consequences of bilingualism is that most do not examine language ability (see Bialystok, 2017, for a review). That is, even though bilingualism is about language experience, few studies examining bilingual cognitive functioning have used anything more than measures of self-reported language proficiency to identify who is bilingual, and do not provide a comprehensive characterization of the context of language use and how it may impact language ability (Surrain & Luk, 2017).

An example that illustrates how language processing is influenced by bilinguals' particular linguistic experiences comes from a recent study by Beatty-Martínez and Dussias (2017). This study examined the processing of codeswitched sentences using event-related potentials (ERPs) in two groups of highly proficient Spanish-English bilinguals who differed in their context of language use. One group lived in Spain and used English as the L2 predominantly in specific environments (e.g., at school or work) and therefore rarely switched between languages within a conversation. Another group was immersed in the United States, a predominantly English language environment, but was born and raised in a Spanish speaking country. Unlike the bilinguals in Spain, they had extensive codeswitching experience. The ERP experiment compared the processing of commonly- and rarely-observed codeswitches across the two groups, and participants' codeswitching behavior was objectively measured based on their performance on a semi-spontaneous speech elicitation task. For codeswitchers, the ERP results revealed that although rarely-observed codeswitches were more difficult to process, codeswitches that adhered to codeswitchers' usage patterns did not result in electrophysiological costs. In contrast, non-codeswitchers processed both common and rare codeswitches with similar difficulty, suggesting that they had not developed sensitivity to codeswitching patterns in their linguistic experience.

The Beatty-Martínez and Dussias (2017) results illustrate how experience with codeswitching, independent of proficiency, is crucial in shaping the processing of codeswitched sentences, and are compatible with the adaptive control hypothesis framework (Green, 2018; Green & Abutalebi, 2013; Green & Wei, 2014). This, in turn, opens the question of whether the choice to habitually codeswitch affects language and cognitive abilities more generally. Critically, the two bilingual groups differed not only as a function of codeswitching experience, but also as a function of language immersion status. The non-codeswitching bilinguals were born, raised, and tested in a predominantly Spanish-speaking, L1 environment. The codeswitching bilinguals were tested while living immersed in an English-speaking, L2 environment.

The dissociation between codeswitching experience and immersion status may be critical given that previous research has shown that the ability to process both the L1 and L2 is modulated by immersion status (Baus, Costa, & Carreiras, 2013; Dussias & Sagarra, 2007; Linck, Kroll, Sunderman, 2009; Zirnstein et al., 2018). Therefore, it is possible that different bilingual experiences can (re)shape the demands imposed by linguistic features, which in turn can affect the relation between language and cognitive control processes. In other words, for bilinguals, the pressures of the environment (e.g., having restricted access to the native language after living for many years in an environment with unrestricted access to the native language) may alter how easy or difficult it is to retrieve words when speaking, which will in turn determine when (and how) cognitive control is engaged.

1.3. The Present Study

The present study examines three contexts of bilingualism for bilinguals who speak the same languages (Spanish and English) and who are all highly proficient in both (see Table 1). In the *separated context*, individuals are more likely to use one language at the expense of the other. Bilinguals in this group live in Spain and use English as the L2 predominantly in specific environments (e.g., at school or work) and therefore rarely switch between their languages within a conversation (i.e., codeswitching). In the *integrated context*, virtually most speakers use the same languages across many life contexts. Bilinguals in this second group live in Puerto Rico where many speakers are also Spanish-English bilinguals and where the two languages are used frequently but also codeswitched in some contexts of everyday life.

Finally, in the *varied context*, the environment is more variable with respect to the types of conversational exchanges that are experienced. Bilinguals in this third group are immersed in the United States, a predominantly English language environment, but initially came from a Spanish-speaking environment similar to that of bilinguals in the integrated context. Although other Spanish-English bilinguals are present, this group lives in a context where most speakers sometimes must use their languages separately (i.e., speaking English with monolingual Anglophones) but can also codeswitch with other Spanish-English bilinguals in certain contexts. At the same time, speakers in the varied context have experienced a shift in their language environment following immigration to the United States, which may require readjusting the relative activation of each language, with some members potentially becoming dominant in English, the predominant language of the environment. Two of the contexts, referred here as separated and varied, were identical to those in the Beatty-Martínez and Dussias (2017) study.

Table 1. Characterization of bilinguals' interactional contexts

Language context	Testing location	Predominant language of environment	Behavioral ecology
Separated	Granada, Spain	Spanish	• Languages must be kept separate • Little-to-no codeswitching experience[a]
Integrated	San Juan, Puerto Rico	Spanish	• Either language can be used opportunistically • Codeswitching experience
Varied	State College, PA, United States	English	• Born and raised in a Spanish-speaking environment • Moved to mainland U.S. during childhood or adolescence • Current restricted use of Spanish • Codeswitching experience[a]

[a] Participants' current codeswitching behavior was objectively assessed via a semi-spontaneous speech elicitation task as part of a larger study investigating the role of codeswitching experience in language processing (Beatty-Martínez & Dussias, 2017).

To compare the performance of these three interactional contexts, we examine two research questions. First, to what extent does variation in bilingual experience affect performance on language production measures over and above proficiency? If the proficient use of two languages is sufficient to determine the speed and accuracy of language processing, then individuals from the three contexts compared here should pattern similarly. Alternatively, it is possible that both proficiency (e.g., Luo, Luk, & Bialystok, 2010) and the age at which English was acquired (e.g., Hernandez & Li, 2006; Hirsh, Morrison, Gaset, & Carnicer, 2003) determine language performance. However, if the modulation of these processes depends on the context of language use, then we might expect differences as a function of whether the two languages are used together or separately, whether bilinguals codeswitch between the two languages, and whether they are immersed in a Spanish- or English- predominant environment (e.g., Gullifer et al., 2018; Hartanto & Yang, 2016; Hofweber, Marinis, & Treffers-Daller, 2016; Pot et al., 2018). If codeswitching is the critical factor that determines how cognitive resources are engaged by bilingual speakers, then the bilinguals in the U.S. (varied context) and in Puerto Rico (integrated context) would be expected to pattern similarly. If immersion in the L2 places unique demands on cognitive resources, then the bilinguals in the U.S. (varied context) with little support for their L1, would be expected to differ from the two groups living in Spain (separated context) or Puerto Rico (integrated context), where the environment supports the use of each language, although in different ways.

We compare performance on two measures of lexical production, category verbal fluency and picture naming. A key feature of verbal fluency is that it leaves the generation of words up to the speaker, in theory reflecting everyday language use (Shao, Janse, Visser, & Meyer, 2014). Unlike picture naming, performance on category verbal fluency is contextually supported by the structure of an individual's semantic network (Kavé & Goral, 2017). For these reasons, verbal fluency has been shown to capture how bilinguals control cross-language competition (Sandoval, Gollan, Ferreria, & Salmon, 2010) and regulate cross-language activation (Zirnstein et al., 2018), and has also been shown to be sensitive to whether learners are immersed in an L1 or L2 environment (Linck et al., 2009). On the other hand, picture naming constrains the event that initiates speech planning, forcing the individual to carry out lexical retrieval without global contextual support. The picture naming induces retrieval difficulties based on item-specific frequency, such that low frequency words are typically harder to retrieve than high frequency words.

Although the two production tasks tap into similar planning processes (Van Assche et al., 2013), we hypothesized that they might differently reflect how bilinguals manage lexical access in each language. Specifically, we expected verbal fluency to primarily reflect language accessibility as a function of environmental demands. Therefore, the separated context should create clear language dominance effects (Spanish > English). While Spanish is the predominant language spoken in Puerto Rico, we hypothesized that the choice to use both languages freely and interchangeably in the integrated context might mitigate the effects imposed by the predominant language of the environment (Spanish = English). Finally, we predicted that the varied context would effectively reverse language dominance (English > Spanish) given individuals' extensive experience in an English L2 immersion environment. We predicted that this pattern of results would hold above and beyond differences in English proficiency and English age of acquisition (AoA).

For picture naming, we hypothesized that performance would reflect the relative accessibility of words in each language as a function of how bilinguals use their languages (i.e., whether they only use their languages separately or whether they have codeswitching experience). Following the frequency-lag account (e.g., Gollan et al., 2008), frequency effects should be larger in L2 English than in L1 Spanish for both separated and integrated contexts because these bilinguals live in a context where Spanish is the predominant language (therefore, the difference between high and low frequency words should be smaller in Spanish). For bilinguals in the varied context, frequency effects should be either comparable across the two languages or smaller in English, reflecting increased functional use of English. However, we predicted that such frequency asymmetries would emerge in the separated context, where one language is typically used at the expense of the other, and that bilinguals in the integrated and varied contexts would pattern similarly due to their extensive practice with codeswitching (which would result in similar frequency effects in both languages). We address this issue by examining frequency effects using picture naming response times (RTs).

The second research question asks whether the demands on language use, and particularly on the pressures associated with deciding how each language might be used in different contexts, modulate the relation between cognitive control and language production. To this end, we examined whether, and if so how, cognitive control ability mediated picture naming performance using the AX variant of the Continuous Performance Task (AX-CPT; Braver et al., 2001). The AX-CPT is a cognitive measure of proactive (e.g., goal maintenance, conflict monitoring, and interference suppression) and reactive (e.g., response inhibition) control processes that has been shown to be related to bilingual experience (Morales, Gómez-Ariza, & Bajo, 2013; Morales, Yudes, Gómez-Ariza, & Bajo, 2015; Zhang, Kang, Wu, Ma, & Guo, 2015, Zirnstein et al., 2018).

Based on the adaptive control hypothesis (Green & Abutalebi, 2013), we hypothesized that differential patterns of association between language and cognitive control for the three interactional contexts might emerge. In the separated context, which has characteristics from both single- and dual-language contexts (i.e., languages are generally used in different domains and are only switched when changing conversational partners), switching from one language to the other may require reactive suppression of the non-target language to change the task goal. The integrated context differs from the separated context in that the environment permits the flexible use of both languages. In consequence, the adaptive response to situational demands for bilinguals in this context may be most strongly associated with a dense-codeswitching environment, where speakers rely on opportunistic planning (i.e., making use of whichever language is most accessible at any given moment). The varied context represents a combination of the three interactional contexts from the adaptive control hypothesis, but most closely approximates the dual-language context since bilinguals in this context are more likely to experience variable circumstances which require constant monitoring of the situation in which communication is going to occur (e.g., Who am I speaking to?, What language(s) does the interlocutor speak?, Is it appropriate to codeswitch?, Am I at home or at work?, etc.). At the same time, because English is the predominant language of the environment, it is likely that the opportunities to use Spanish are likely constrained to limited domains, potentially requiring a dynamic reconfiguration of the language system. Control processes associated with proactive control are expected to trigger the strongest adaptive response to environmental demands of the varied context.

2. Method

2.1. Participants

Three groups of Spanish-English bilinguals participated in this study. All participants gave informed consent and the procedures had the approval of the Institutional Review Board of the Pennsylvania State University (IRB 34810). Participants were paid $10 per hour (or an equivalent of $10 per hour in euros for those recruited in Spain) for their participation. Participants' characteristics are shown in Tables 2, 3, and Figure 1. Bilinguals in the *separated context* (*n* = 31, 20 females) were recruited at the University of Granada, Spain, a predominantly Spanish-speaking environment where codeswitching is not a recurrent form of conversational exchange (Beatty-Martínez & Dussias, 2017). Bilinguals in the *integrated context* (*n* = 34, 31 females) were recruited at the University of Puerto Rico, a predominantly Spanish-speaking context but where English is widely used in education, media, and other societal domains (see Figure 1), and codeswitching among bilinguals is very common (Beatty-Martínez, 2019; Casas, 2016; Guzzardo Tamargo, Loureiro-Rodríguez, Acar, & Vélez Avilés, 2018; Pousada, 2017). Bilinguals in the *varied context* (*n* = 31, 25 females) were from Hispanic countries who had moved to the United States during childhood or adolescence and were raised in established Spanish-English codeswitching communities in the U.S. (Fricke, Kroll, & Dussias, 2016; Guzzardo Tamargo, Valdés Kroff, Dussias, 2016; Poplack, 1980; Valdés Kroff, Dussias, Gerfen, Perrotti, & Bajo, 2016). At the time of testing, participants in this group were students at Pennsylvania State University in State College, Pennsylvania, a predominantly English-speaking environment where the Hispanic population is only 4.4% (US Census Bureau, 2015).

Table 2. Participant self-reported characteristics

Measure	Context					
	Separated		Integrated		Varied	
	M	SD	M	SD	M	SD
Age, years	23.6	3.4	19.9	2.5	21.3	3.2
English AoA, years	5.9	2.3	4.2	2.3	5.4	3.2
English immersion, years	1.3	2.0	0.8	2.2	9.5	8.1
Spanish exposure, %	71.6	13.6	63.3	16.3	32.2	14.3
English exposure, %	25.5	18.4	33.3	14.0	64.8	15.5
Spanish proficiency	9.6	0.7	9.1	0.8	9.4	0.8
English proficiency	8.2	0.9	8.9	0.7	9.1	0.9

Note: Means and standard deviations for participants' language history characteristics. AoA = age of acquisition. Proficiency ratings were made on a 10-point scale ranging from 1 (not proficient) to 10 (highly proficient). Not all participants filled in all questions. See Table S1 in the Supplemental Materials for confidence intervals and the valid N for each measure per group.

To assess language experience, participants completed the Spanish version of the LEAP-Q language questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007). All participants were native Spanish speakers who acquired Spanish at birth and English either simultaneously or in early childhood, and reported high levels of proficiency in both languages. Furthermore, while bilinguals in separated and integrated contexts reported higher overall exposure to Spanish relative to English, bilinguals in the varied context, not surprisingly, reported higher overall exposure to English relative to Spanish. Bilinguals in this context also reported having prolonged immersion experience in English. However, Figure 2 shows that the relative language exposure varies across social domains for each context. In all three contexts, Spanish was reported as the predominant language in the family domain, although discrepancies emerged with friends. While bilinguals in the separated context reported Spanish as the predominant language when interacting with friends, bilinguals in integrated and variable contexts reported being exposed to both languages to a similar degree. This discrepancy highlights how bilinguals in the integrated and varied contexts have more opportunities to use English with other speakers compared to bilinguals in the separated context.

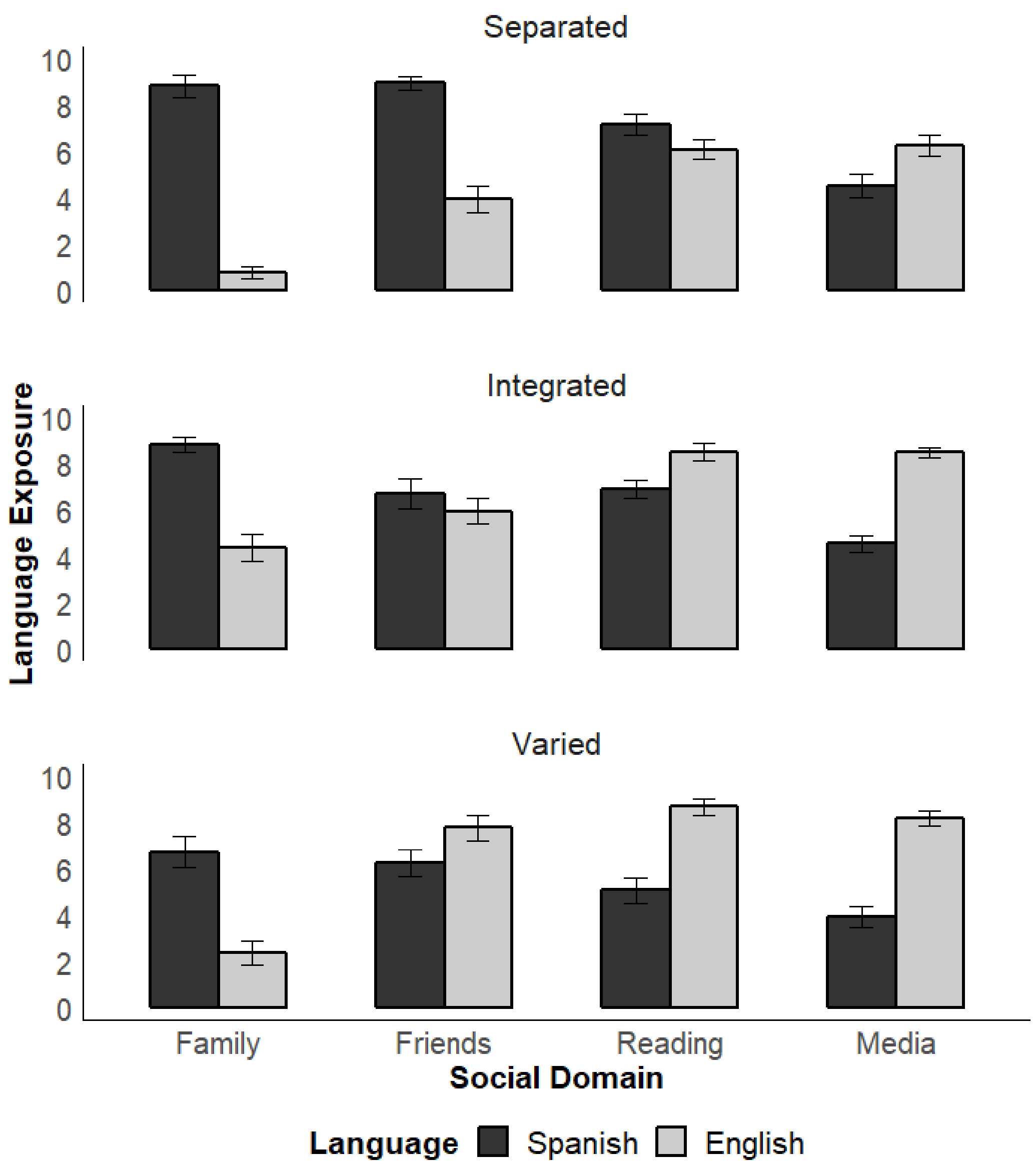

Figure 2. Participants' self-reported exposure to Spanish and English across different social domains. Ratings were made on a 10-point scale ranging from 0 (no exposure) to 10 (high exposure). Error bars indicate standard error of the mean. See Table S2 in the online Supplementary Materials for mean values, standard deviations, confidence intervals, and valid N for each measure per group.

Finally, to measure participants' everyday language switching tendencies, we administered the Bilingual Switching Questionnaire (BSWQ; Rodriguez-Fornells, Kramer, Lorenzo-Seva, Festman, & Münte, 2012). This measure decomposes language switching tendencies into distinct constructs: (a) switching directionality (i.e., switching from the L1 into the L2 or vice versa in order to fill lexical gaps or better convey a message), (b) contextual switching (i.e., whether participants alternate between languages in response to particular sociolinguistic situations or environments), and (c) unintended switching (i.e., awareness of switching languages). Participants answered 12 questions representing these constructs on a 5-point scale varying from never (1) to always (5; see Appendix A for a list of all the questions). Participants' scores on these constructs are shown in Table 3. Bilinguals in integrated and varied contexts reported a greater tendency to switch from Spanish into English, and a higher frequency of contextual switching than bilinguals in the separated context. This is consistent with distributional usage patterns extracted from bilingual corpora of habitual codeswitching communities similar to those examined here (Beatty-Martínez & Dussias, 2017; Beatty-Martínez, Valdés Kroff, & Dussias, 2018; Guzzardo Tamargo et al., 2016; Królikowska et al., 2019; Pfaff, 1979; Poplack, 1980).

Table 3. Mean and standard deviations for scores on the BSWQ subscales

Measure	Context					
	Separated		Integrated		Varied	
	M	SD	M	SD	M	SD
Switching directionality: English into Spanish	2.9	.67	3.2	.61	3.4	.61
Switching directionality: Spanish into English	2.7	.85	3.5	.67	3.5	.49
Contextual switching	3.1	.94	4.0	.73	3.9	.71
Unintended switching	2.7	.50	3.0	.56	2.8	.66

Note: BSWQ = Bilingual Switching Questionnaire (Rodriguez-Fornells et al. (2012). Codeswitching frequency ratings were made on a 5-point scale ranging from 1 (never) to 5 (always). Not all participants filled in all questions. See Table S3 in the Supplemental Materials for confidence intervals and the valid N for each measure per group.

2.2. Materials and Design

Category verbal fluency task. In this task, participants were asked to generate as many exemplars as possible that belong to a semantic category within a 30-second time limit. The task included eight categories (the same categories as in Baus et al., 2013 and Linck et al., 2009) that were counterbalanced and evenly distributed between language blocks. The categories were *animals*, *clothing*, *musical instruments*, and *vegetables* or *body parts*, *colors*, *fruits*, and *furniture*. Participants were asked to avoid producing repetitions and names of people or places. Responses were recorded on a digital recorder. Verbal fluency performance was analyzed by calculating the average number of exemplars produced across categories in Spanish and in English.

Picture naming task. We adapted a version of the picture naming task used by Gollan et al. (2008). Participants named a total of 132 black and white line-drawn pictures over a range of lexical frequencies. The picture names are listed in Appendix B with their corresponding lexical frequency values. Half of the pictures were presented in the Spanish block and the other half were presented in the English block. As depicted in Figure 2, the picture naming trial sequence started with a 500 ms fixation cross ('+') in the middle of screen. Participants initiated each trial by pressing the spacebar which triggered the presentation of a picture. The picture disappeared from the display when the voice-key was triggered or an interval of 3000 ms had passed. Instructions were to name pictures "as quickly and as accurately as possible" in the appropriate language, and to avoid coughs, false starts, and hesitations.

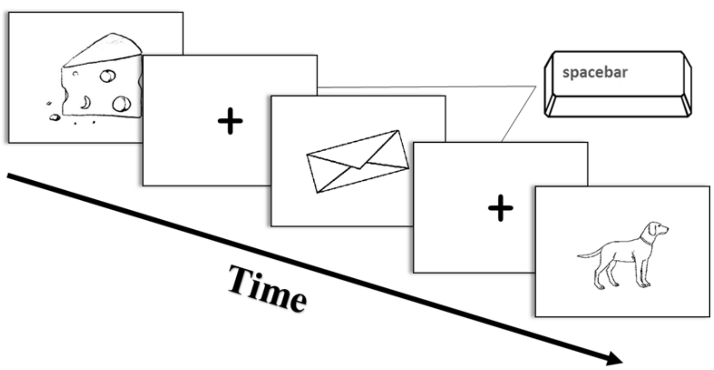

Figure 3. Procedure for the picture naming task.

We collected accuracy and RT data. A response was considered accurate if it matched the intended target name. Where appropriate, alternative dialectal variations were also considered accurate. Three items from the English block (i.e., apron, eggs, and glass) were excluded due to misidentification errors. We excluded any RTs that were associated with inaccurate responses and registration errors (e.g., hesitations and repetitions), or that were either below 300 ms or above 2000 ms. Any remaining RTs that deviated more than 2.5 SDs from the mean of each participant were also excluded.

AX-CPT. The AX-CPT is a non-linguistic task developed to study variability in the use of proactive and reactive control processes (Figure 4). In this version of the task (Ophir, Nass, Wagner, & Posner, 2009), participants were presented with cue-probe pairs in red and were required to respond “yes” only when they detect an AX sequence (i.e., an X-probe preceded by an A-cue), and “no” to any other cue-probe combinations[16] (i.e., AY, BX, BY). Three distractor letters, presented in white, were introduced between cue and probe letters. Participants were instructed to respond “no” to each distractor. While AX trials occurred throughout the experiment with high frequency (70% of the time), each of the other trial types (AY, BX, BY) occurred on 10% of the time. This specific version of the task was chosen because it has been successfully used in previous studies to characterize bilinguals’ reliance on proactive vs. reactive control (Bice et al., 2015; Morales et al., 2013; Zhang et al., 2015; Zirnstein et al., 2018; see also Morales et al., 2015; Gullifer et al., 2018, for similar versions).

Importantly, this design induces two types of context-driven biases in participants. The first bias is an expectancy to make a target “yes” response following A-cues. Context information serves as a predictive function allowing participants to act proactively to prime the selection of a target “yes” response. However, this bias creates the tendency to

16 Letters “B” (for BX, BY) and “Y” (for AY, BY) are used as place-holders for any non-A-cue and non-X-probe letter stimuli (e.g., “J” or “L”) respectively. Letters ‘‘K’’ and ‘‘Y’’ were excluded due to their perceptual similarity with ‘‘X’’.

false alarm on AY trials. In other words, context information should impair performance by creating an inappropriate expectancy bias for AY trials. As such, participants who greatly rely on context are likely to demonstrate increased error rates and slower RTs in AY trials relative to control BY trials where both the cue and the probe always map to a nontarget response. The second bias is to make a target “yes” response on X-probes. On BX trials, context information must be used in an inhibitory fashion to override the tendency to false alarm. Thus, reliance on context information might aid performance on BX trials by inhibiting or overriding the prepotent response tendency, but failures in context monitoring and goal maintenance would produce elevated error rates in BX relative to control BY trials where the probe does not trigger the target "yes" response. Faster RTs in BX relative to BY trials indicate that participants used the cue to correctly predict the probe and override the prepotent response tendency. On the other hand, slower RTs in BX relative to BY trials signal difficulty reactivating context information, which may trigger a need for reactive inhibitory control processes to suppress the incorrect “yes” response. Letters were presented each for 300 ms with a 1000 ms interval between letters.

Participants completed 10 practice trials including all four experimental conditions, and they were provided with feedback on accuracy and RT after each practice trial. Completion of the practice block was followed by the experimental block composed of 100 trials. Error rates and RTs were recorded for each condition. RTs were computed from correct responses. In a first pass, responses that were either below 100 ms or above 1200 ms were removed. For the remaining RTs, extreme outliers were excluded through visual inspection using histograms and boxplots (1% of trials).

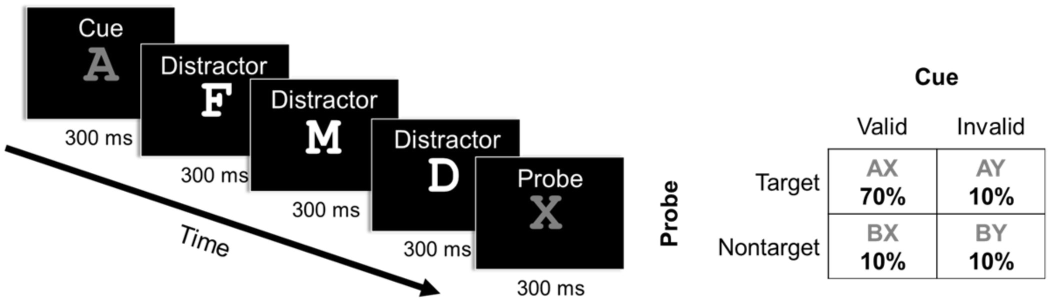

Figure 4. Schematic representation of the procedure for the AX-CPT. AX are target trials that require a "yes" probe response (70% of trials). All other trial sequences (each occurring 10% of the time) require a "no" probe response. AY trials share the cue with target trials, which biases participants to anticipate the target probe. High error rates in these trials indicate failure to suppress an incorrect "yes" response due to high reliance on context. In BX trials, the cue signals a "no" response, but the probe prompts a target response. High error rates in these trials indicate failure to suppress a "yes" response due to minimal or no reliance on context. BY are control trials where the influence of context is reduced, since both the cue and probe differ from target trials.

2.3. Procedure

All tasks were completed on a computer that was connected to a button box and a digital recorder in a sound-attenuated room. At the beginning of each task, participants were carefully briefed on the experimental procedure, and they completed a practice run for each task to ensure that they understood the instructions. Participants performed the verbal fluency and picture naming tasks first. Written instructions indicating the language to be used appeared on the screen, and the order of language of production was blocked such that participants completed all tasks in the L1 (Spanish) first and in the L2 (English) second. After completing the language tasks, participants performed the AX-CPT, followed by the language history questionnaire.

2.4. Analysis

For verbal fluency, we used repeated measures ANCOVAs, with language block (Spanish vs. English) as the within-subjects factor, and context (separated, integrated, and varied) as the between-subjects factor, to analyze the average number of exemplars produced by each group in each language. We used English picture naming accuracy (i.e., the proportion of correct responses for the English naming block) and self-reported English age of acquisition (AoA) as covariates to control for differences that could be attributed to L2 proficiency or to the amount of time spent with the L2 across the lifespan. All other statistical analyses were performed using linear and generalized mixed-effects models in the lme4 software package (version 1.1-18-1; Bates, Mächler, Bolker, & Walker, 2015) in the R programming environment (version 3.5.1; R Development Core Team, 2014). Unlike ANOVAs, mixed models can estimate trial and participant-level data under one analytic framework, therefore increasing the generalizability of results to other individuals and items (Baayen, Davidson, & Bates, 2008).

For picture naming accuracy, the analysis included a contrast coded fixed effect of language block (Spanish = -0.5, English = 0.5), a dummy coded fixed effect of context (separated, integrated, varied), log-transformed word frequency values (used as a continuous factor), and an interaction between language block and context. For picture naming RTs, the analysis additionally included a three-way interaction between language block, context, and frequency. To guard against Type 1 errors and increase generalizability, random effects were fit using a maximal procedure (Barr, Levy, Scheeper, & Tily, 2013), with crossed random effects for participants and items. For accuracy, the final model contained random intercepts for subjects and items, by-participant random slopes for language block and frequency, and by-item random slopes for context. For RTs, the final model additionally included a by-participant random slope for the interaction between language block and frequency.

For the AX-CPT, linear and generalized mixed model analyses included dummy coded fixed effects of condition (AY, BX, BY), context, and a condition-by-context interaction. Of primary interest in this analysis were several comparisons, including AY vs. BY to measure the degree to which the context bias negatively impacted probe responses, and BX vs. BY to measure the degree to which the context facilitated probe responses. In both cases, we used BY as the reference level. We also compared BX vs. AY using BX as the reference level to assess general reliance on proactive vs. reactive control. The final generalized mixed model contained a by-participant random slope for condition and a by-item random slope for group. Due to convergence failures, the final linear mixed model only included a by-participant random slope for condition.

To identify individual differences, follow-up mixed effects models were computed to examine the effect of cognitive control on picture naming performance (see Gullifer & Titone, 2019; Kliegl, Wei, Dambacher, Yan, & Zhou, 2010; Linck, 2016; Mirman, 2011, for different applications of mixed modeling to study individual differences). Based on previous bilingual studies (Bice & Kroll, 2015; Morales et al., 2013; Zirnstein et al., 2018), we extracted four measures from the AX-CPT that included AY and BX error rates, as well as two difference score efficiency measures from the RT data (log AY – log BY, and log BX – log BY). Each individual difference measure was included in a separate model as a fixed effect and allowed to interact with context and frequency in the logistic (accuracy) analyses, and with language, context, and frequency in the linear (RT) analyses. A maximal procedure for the random effects structure was not possible in these models due to convergence failures. Following the recommendation of Bates et al. (2015), we conducted a principal component analysis (PCA) to simplify the random effects structure. The PCA indicated overspecification of the by-participant random slope for frequency in the accuracy models, and overspecification of the by-participant random slope for the language-by-frequency interaction. Therefore, these parameters were removed from the individual difference analyses.

Within the mixed models, significant interactions were examined by refitting a model with a dummy coded categorical factor to examine simple effects at each level of the categorical factor, or by rescaling continuous factors one SD above/below the mean to examine simple effects of categorical factors at high and low values of the continuous factor (see Aiken & West, 1991; Hardy, 1993; Jaccard & Turrisi, 2003). For example, a significant interaction between language block and frequency might indicate that the effect of word frequency is significantly smaller (i.e., less steep) in one language relative to the other language. However, this would not indicate whether the frequency slope significantly differs from zero in each language, and whether each slope is significantly different from one another. To do this, we can refit the model by releveling a given variable (e.g., dummy-coding Language and setting the Spanish naming block as the reference level to determine the significance of the frequency slope for Spanish). Note that refitting or releveling does not affect the goodness of fit of the model or the type-1 error rate. Instead, the model simply re-estimates the parameters with a different reference point, providing a different interpretation of the coefficients while keeping the variance constant (Gelman & Hill, 2007).

Continuous fixed-effects were z-scored to make the intercept in the models reflect average performance. To obtain *p*-values for the fixed effects in the mixed model RT analyses, we used the Satterthwaite approximation with the lmerTest package (version 3.0-1; Kuznetsova, Brockhoff, & Bojesen, 2016). For mixed models, we also report confidence intervals of the estimates to assist in the interpretation of significant and/or meaningful results. A summary of the results for each mixed model analysis, including fixed effects, random effects, and confidence interval estimates, is reported in separate tables. However, estimates involving releveling or follow-up comparisons are reported in the main text.

3. Results

3.1. How Is Language Production Affected by the Interactional Context?

Category verbal fluency. As shown in Table 4, verbal fluency scores revealed high verbal abilities in both languages, although important differences between the three contexts emerged. Bilinguals in the separated context produced more exemplars in Spanish than in English. Bilinguals in the integrated context, on the other hand, produced a similar number of exemplars in both languages, but bilinguals in the varied context produced fewer exemplars in Spanish than in English. After controlling for differences in English proficiency ($F(1, 85) = 3.93$, $p = .050$, $\eta p^2 = .04$) and in English AoA ($F(1, 85) = 3.09$, $p = .082$, $\eta p^2 = .04$), these results were confirmed in the ANCOVA analysis by a language-by-group interaction ($F(2, 93) = 19.96$, $p < .001$, $\eta p^2 = .30$), and follow-up ANCOVAs examining the main effect of language for each context (Separated: $F(1, 30) = 49.56$, $p < .001$, $\eta p^2 = .62$; Integrated: $F(1, 34) = 0.00$, $p = .999$, $\eta p^2 = .00$; Varied: $F(1, 30) = 9.23$, $p = .005$, $\eta p^2 = .24$). Together, they suggest that these context-driven differences likely reflect language accessibility as a function of the *current* dynamics of the language environment (i.e., the degree to which the environment supports the use of one or both languages), and confirm our characterization of the three contexts with respect to how the languages are habitually used (i.e., in an independent or interdependent fashion).

Table 4. Descriptives of language production measures by task and interactional context

Variable	Context					
	Separated		Integrated		Varied	
	M	SD	M	SD	M	SD
Verbal Fluency						
Spanish	54.4	8.1	43.9	7.3	42.9	9.6
English	43.6	5.4	43.9	6.2	48.0	9.4
Picture Naming						
Spanish accuracy	.98	.02	.95	.03	.91	.05
English accuracy	.90	.05	.94	.04	.94	.05
Spanish latency (*ms*)	1122	218	1130	174	1173	187
English latency (*ms*)	1319	184	1080	177	1085	181

See Table S4 in the Supplementary Materials for confidence intervals for each measure per group.

Picture naming. How does language use affect lexical access? Overall picture naming accuracy (Table 4) confirmed that individuals across the three contexts were highly proficient in both languages (i.e., with mean accuracy above 90% in each language), but the analyses revealed important within-context differences. Consistent with verbal fluency performance, individuals in the separated context were more accurate (Table 5) and faster (Table 6) in Spanish than in English, reflecting enhanced lexical accessibility in Spanish, but also reflecting the independent use of both languages. For individuals in the integrated context, however, there was a dissociation between accuracy and RT performance: although picture naming accuracy was higher in Spanish than in English[17] (β = -1.01, SE = 0.45, z = -2.26, p = .024, 95% CI = [-1.89, -0.14]), the two languages had similar latencies (β = 0.00, SE = 0.02, t = 0.11, p = .915, 95% CI = [-0.03, 0.03]). Accuracy performance suggests enhanced lexical accessibility in Spanish, but the latencies suggest interdependent use of both languages. Finally, the varied context yielded similar accuracy (β = -0.32, SE = 0.40, z = -0.80, p = .422, 95% CI = [-1.11, 0.47]) and similar latencies (β = -0.02, SE, = 0.02, t = -1.01, p = .314, 95% CI = [-0.05, 0.01]) in the two languages, reflecting similar lexical accessibility and interdependent use of the two languages.

[17] Note that Table 4 contains the raw means, which suggest that picture naming accuracy was similar in both languages. Mixed models, on the other hand, calculate predicted means that are conditional on the predictor values and random effects, which can differ from raw means (such as in this case).

Table 5. Estimated coefficients from the mixed model on picture naming accuracy

Fixed effects	Estimate	*SE*	*z*-values	*p*	Lower	Upper
(Intercept)	**4.64**	**0.24**	**18.75**	**< 0.001**	**4.15**	**5.12**
Frequency	**1.36**	**0.18**	**7.70**	**< 0.001**	**1.01**	**1.71**
Integrated	0.19	0.30	0.64	0.522	-0.40	0.78
Varied	**-0.55**	**0.26**	**-2.15**	**0.016**	**-1.05**	**-0.05**
Language	**-2.99**	**0.45**	**-6.66**	**< 0.001**	**-3.87**	**-2.11**
Integrated*Language	**1.90**	**0.50**	**3.83**	**< 0.001**	**0.93**	**2.87**
Varied*Language	**2.69**	**0.46**	**5.86**	**< 0.001**	**1.79**	**3.59**

Random effects	Variance	*SD*	Correlation	
Intercept | item	1.9352	1.39		
Integrated | item	1.3005	1.14	-0.06	
Varied | item	0.7231	0.85	-0.15	0.83
Intercept | participant	0.2643	0.51		
Frequency | participant	0.0046	0.07	-0.15	
Language | participant	1.4533	1.21	-0.09	-0.97

Notes. Lower/Upper = lower and upper bounds for 95% confidence intervals of coefficient estimate; SE = standard error of coefficient estimate. The separated-context bilingual group was set as the reference level. Language was contrast coded (-0.5 = Spanish; 0.5 = English), making the coefficient interpretations as follows: Intercept = mean naming accuracy (in log odds) for separated-context bilinguals with an average frequency effect. Frequency = effect of frequency (centered at the sample mean) on accuracy for separated-context bilinguals. Integrated = mean accuracy difference between separated- and integrated-context bilinguals. Varied = mean accuracy difference between separated- and varied-context bilinguals. Bold indicates coefficients that are significantly different from zero.

To what extent does language use modulate lexical frequency effects? Recall that, under the frequency-lag account (e.g., Gollan et al., 2008), bilinguals in separated and integrated contexts should yield larger frequency effects in English relative to Spanish, and bilinguals in the varied context should either show comparable frequency effects in both languages, or smaller frequency effects in English. As Figure 5 shows, bilinguals in the separated context exhibited the predicted asymmetric frequency effects across the languages, supporting the frequency-lag account. However, bilinguals in integrated and varied contexts exhibited similar performance in their two languages and no asymmetric frequency effects. The analysis confirmed this pattern of results via a significant language-by-frequency interaction for separated-context bilinguals (Table 6), but not for integrated (β = 0.00, *SE* = 0.01, *t* = 0.26, *p* = .801, 95% CI = [-0.02, 0.03]) or varied-context bilinguals (β = 0.01, *SE* = 0.01, *t* = 0.46, *p* = .648, 95% CI = [-0.02, 0.03]). Follow-up simple effects analyses revealed that, for bilinguals in the separated context, the frequency effect was reliably smaller in the Spanish block (β = -0.02, *SE* = 0.02, *t* = -3.06, *p* = .003, 95% CI = [-0.04, -0.01]) than in the English block (β = -0.07, *SE* = 0.01, *t* = -8.38, *p* < .001, 95% CI = [-0.08, -0.05]).

Table 6. Estimated coefficients from the mixed model on picture naming RTs

Fixed effects	Estimate	*SE*	*t*-values	*p*	Lower	Upper
(Intercept)	**3.08**	**0.013**	**242.96**	**< 0.001**	**3.05**	**3.10**
Integrated	**-0.05**	**0.017**	**-3.18**	**0.002**	**-0.09**	**-0.02**
Varied	**-0.05**	**0.017**	**-2.63**	**0.010**	**-0.08**	**-0.01**
Language	**0.10**	**0.014**	**6.79**	**< 0.001**	**0.07**	**0.13**
Frequency	**-0.04**	**0.006**	**-8.04**	**< 0.001**	**-0.06**	**-0.03**
Integrated*Language	**-0.10**	**0.016**	**-5.85**	**< 0.001**	**-0.13**	**-0.06**
Varied*Language	**-0.11**	**0.017**	**-6.80**	**< 0.001**	**-0.15**	**-0.08**
Integrated*Frequency	-0.01	0.005	-1.05	0.298	-0.02	0.01
Varied*Frequency	-0.01	0.005	-1.15	0.251	-0.02	0.01
Language*Frequency	**-0.04**	**0.011**	**-4.15**	**< 0.001**	**-0.07**	**-0.02**
Integrated*Language*Frequency	**0.05**	**0.010**	**4.91**	**< 0.001**	**0.03**	**0.07**
Varied*Language*Frequency	**0.05**	**0.009**	**5.31**	**< 0.001**	**0.03**	**0.07**

Random effects	Variance	*SD*	Correlation		
Intercept | item	0.0027	0.05			
Integrated | item	0.0015	0.04	-0.05		
Varied | item	0.0012	0.03	-0.19	0.65	
Intercept | participant	0.0042	0.06			
Language | participant	0.0031	0.06	-0.07		
Frequency | participant	0.0001	0.01	-0.09	-0.13	
Language*Frequency | participant	0.0001	0.01	-0.28	-0.92	-0.04
Residual	0.0142	0.12			

Notes. The separated-context bilingual group was set as the reference level. Intercept represents mean log RTs for separated-context bilinguals with an average frequency effect.

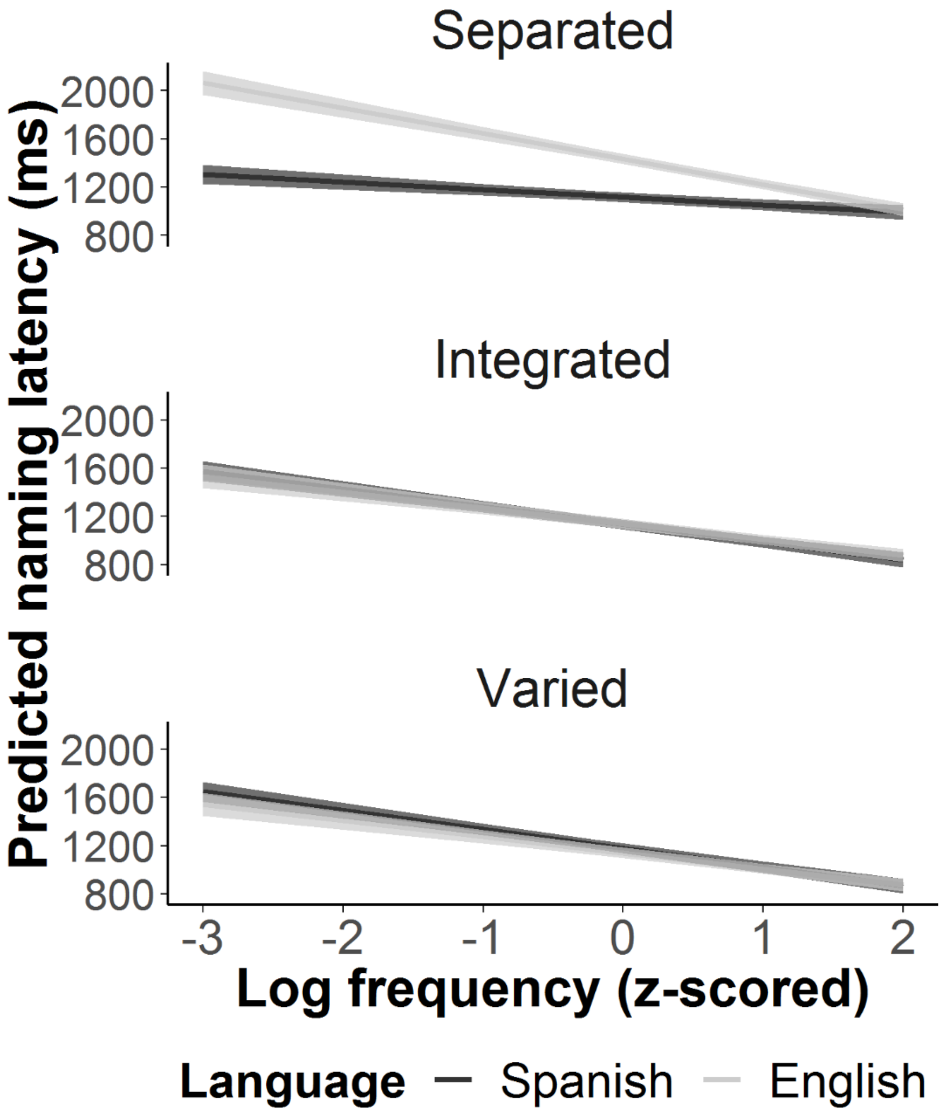

Figure 5. Predicted picture naming latencies displayed via a three-way interaction of context, language, and z-scored log word frequency. Negative values on the x-axis indicate lower frequency words. Shaded areas indicate standard errors of the means.

Although the data reported so far can be explained in terms of how bilinguals use their languages, it is possible that individual variability in language dominance might have washed out the frequency asymmetries for bilinguals in the integrated and varied contexts. We tested this possibility by creating a language dominance index by calculating the difference between Spanish and English picture naming accuracy (see Figure S1 in the Supplementary Materials). The dominance index was then included as a continuous predictor in the picture naming RT analysis, and allowed to interact with all other fixed effects. For our purposes, the key prediction in this analysis is that, if cumulative linguistic experience determines frequency asymmetries, then there should be a significant three-way interaction between dominance, language, and frequency for bilinguals in each context.

In the analysis (Table S6), language dominance did not reliably modulate the frequency effects for individuals in the integrated and varied contexts. In these groups, the three-way interaction between language, frequency, and dominance was not significant (Integrated: $\beta = -0.01$, $SE = 0.00$, $t = -1.31$, $p = .192$, 95% CI = [-0.01, 0.00]; Varied: $\beta = -0.01$, $SE = 0.01$, $t = -1.04$, $p = .297$, 95% CI = [-0.01, 0.00]). In contrast, for separated-context bilinguals, the three-way interaction was significant ($\beta = -0.01$, $SE = 0.00$, $t = -2.61$, $p = .010$, 95% CI = [-0.02, -0.00]). Follow-up simple slopes analyses revealed that the frequency asymmetry was largest for highly Spanish dominant individuals in the separated context (Spanish: $\beta = -0.02$, $SE = 0.01$, $t = -2.04$, $p = .043$, 95% CI = [-0.03, -0.00]; English: $\beta = -0.07$, $SE = 0.01$, $t = -7.98$, $p < .001$, 95% CI = [-0.09, -0.05]). However, for individuals who were less Spanish dominant, the magnitude of the frequency asymmetry was reduced (Spanish:

$\beta = -0.03$, $SE = 0.01$, $t = -3.54$, $p = .001$, 95% CI = [-0.05, -0.01]; English: $\beta = -0.06$, $SE = 0.01$, $t = -7.05$, $p < .001$, 95% CI = [-0.08, -0.04]).

Additionally, there was a significant two-way interaction between language block and dominance for low frequency words ($\beta = 0.04$, $SE = 0.01$, $t = 3.70$, $p < .001$, 95% CI = [0.02, 0.06]), such that individuals who were more Spanish dominant became slower to produce low frequency words in English ($\beta = 0.03$, $SE = 0.01$, $t = 2.03$, $p = .044$, 95% CI = [0.00, 0.05]). Taken together, these results suggest that, at least for highly proficient bilinguals such as those examined here, picture naming performance can reflect differences in how the languages are habitually used (i.e., in an independent or interdependent manner), and not just reduced lexical access due to the amount of experience in each language.

3.2. To What Extent Do Cognitive Control Strategies Reflect Bilinguals' Interactional Demands?

AX-CPT error rates. Table 7 shows the AX-CPT error rates and RTs across conditions for each group. Individuals across the three contexts of language use made on average more errors in the AY condition relative to the BY control condition. The mixed model analysis confirmed this pattern of results (Separated: $\beta = -1.18$, $SE = 0.30$, $z = -3.93$, $p < .001$, 95% CI = [-1.78, -0.59]; Integrated: $\beta = -1.79$, $SE = 0.34$, $z = -5.22$, $p < .001$, 95% CI = [-2.46, -1.12]; Varied: $\beta = -1.72$, $SE = 0.34$, $z = -5.05$, $p < .001$, 95% CI = [-2.39, -1.05]). This suggests that on AY sequences participants relied on contextual information (i.e., the A-cue) to anticipate upcoming probe responses and, as a result, had greater difficulty selecting the correct probe response.

Table 7. AX-CPT scores by interactional context

	Context					
Condition	Separated		Integrated		Varied	
	M	SD	M	SD	M	SD
AX error rate	0.09	0.07	0.10	0.07	0.09	0.06
AY error rate	0.28	0.16	0.26	0.18	0.26	0.16
BX error rate	0.30	0.30	0.15	0.15	0.13	0.11
BY error rate	0.14	0.15	0.08	0.12	0.07	0.10
AX latency (*ms*)	299.97	34.81	283.63	36.48	321.47	66.55
AY latency (*ms*)	414.42	60.85	425.22	70.10	466.41	86.64
BX latency (*ms*)	243.61	64.09	231.32	59.64	247.83	71.75
BY latency (*ms*)	286.15	91.21	251.54	48.38	274.09	23.40

Notes. Some data were excluded due to experimental or equipment error. We also excluded one participant in the integrated group who had an outlier performance score in the AY efficiency measure. This was determined through visual inspection of residual plots, and by calculating Cook's distance on the individual difference analyses. See Table S5 in the Supplementary Materials for confidence intervals and the valid N for each measure per group.

Relative to the BX condition, AY error rates were also higher for integrated (β = 1.08, *SE* = 0.32, *z* = 3.93, *p* = .001, 95% CI = [0.46, 1.71]) and varied-context bilinguals (β = 1.08, *SE* = 0.32, *z* = 3.93, *p* = .001, 95% CI = [0.46, 1.71]), although no difference in error rates was found for separated-context bilinguals (Table 8). This indicates that on BX sequences, integrated- and varied-context bilinguals used contextual information to minimize prepotent response tendencies, but that separated-context bilinguals relied more on probe information, likely triggering prepotent responses that required reactive inhibitory mechanisms.

Finally, relative to BY, error rates in BX trials were higher for separated-context (Table 8) and integrated-context bilinguals (β = -0.79, *SE* = 0.37, *z* = -2.16, *p* = .031, 95% CI = [-0.07, -1.52]), although the difference between BX and BY was not significant for varied-context bilinguals (β = -0.62, *SE* = 0.38, *z* = -1.64, *p* = .102, 95% CI = [0.12, -1.35]). This suggests that individuals in the varied context were the most efficient at taking advantage of the cue to override prepotent response tendencies on X-probes. Critically, no reliable between-group differences were observed for BY error rates (Separated vs. Integrated: β = 0.73, *SE* = 0.40, *z* = 1.81, *p* = .070, 95% CI = [-0.06, 1.51]; Separated vs. Varied: β = 0.69, *SE* = 0.41, *z* = 1.70, *p* = .089, 95% CI = [-0.11, 1.51]; Integrated vs. Varied: β = -0.04, *SE* = 0.43, *z* = -0.09, *p* = .929, 95% CI = [-0.87, 0.80]).

Table 8. Estimated coefficients from the mixed model on AX-CPT error rates

Fixed effects	Estimate	*SE*	*z*-values	*p*	Lower	Upper
(Intercept)	**-1.14**	**0.28**	**-4.13**	**< 0.001**	**-1.68**	**-0.60**
BY	**-1.08**	**0.32**	**-3.39**	**0.001**	**-1.71**	**-0.46**
AY	0.10	0.32	0.33	0.744	-0.52	0.73
Integrated	**-1.01**	**0.38**	**-2.67**	**0.008**	**-1.76**	**-0.27**
Varied	**-1.17**	**0.40**	**-2.92**	**0.003**	**-1.95**	**-0.38**
BY*Integrated	0.29	0.43	0.67	0.501	-0.55	1.13
AY*Integrated	**0.89**	**0.43**	**2.08**	**0.038**	**-0.05**	**-1.73**
BY*Varied	0.47	0.45	1.02	0.306	0.43	-1.36
AY*Varied	**1.00**	**0.46**	**2.18**	**0.029**	**0.10**	**1.90**

Random effects	Variance	*SD*	Covariance	
Intercept | item	0.1325	0.36		
Integrated | item	0.0816	0.29	0.28	
Varied | item	0.2124	0.46	-0.49	-0.97
Intercept | participant	1.2263	1.11		
BY | participant	0.5824	0.76	-0.48	
AY | participant	1.2570	1.12	-0.90	0.80

Notes. The BX condition and the separated-context bilingual group were set as the reference levels. Intercept reflects mean BX error rates (in log odds) for separated-context bilinguals.

AX-CPT RTs. Consistent with the results for error rates, AY trials yielded slower responses relative to BY control trials across the three contexts (Separated: β = 0.20, *SE* = 0.02, *t* = 9.86, *p* < .001, 95% CI = [0.16, 0.23]; Integrated: β = 0.24, *SE* = 0.02, *t* = 13.02, *p* < .001, 95% CI = [0.20, 0.27]; Varied: β = 0.25, *SE* = 0.02, *t* = 13.01, *p* < .001, 95% CI = [0.21, 0.28]), suggesting that the A-cue bias led to subsequent processing difficulties during correct probe responses. Unlike AY trials, BX trials yielded facilitatory responses, such that bilinguals in each context responded faster to BX trials relative to BY trials (Separated: Table 9; Integrated: β = 0.04, *SE* = 0.01, *t* = 2.98, *p* = .003, 95% CI = [0.06, 0.01]; Varied: β = 0.04, *SE* = 0.13, *t* = 2.93, *p* = .004, 95% CI = [0.07, 0.07]). This suggests that correct responses were achieved by anticipating X-probes upon detection of the B-cue. However, a significant group-by-AY interaction (Table 9) indicated that the magnitude of the BX vs. AY difference was greater for varied-context bilinguals ($\beta = 0.29$, $SE = 0.02$, $t = 14.32$, $p < .001$, 95% CI = [0.25, 0.32]) than for separated-context bilinguals ($\beta = 0.23$, $SE = 0.02$, $t = 10.44$, $p < .001$, 95% CI = [0.18, 0.26]), suggesting greater general reliance on proactive control for the former group and greater reliance on reactive control for the latter group. Follow-up group comparisons additionally revealed slower AY responses for varied-context bilinguals relative to separated-context bilinguals ($\beta = 0.05$, $SE = 0.02$, $t = 2.42$, $p = .018$, 95% CI = [0.01, 0.08]), but no reliable differences between integrated- and varied-context bilinguals ($\beta = 0.04$, $SE = 0.02$, $t = 1.91$, $p = .059$, 95% CI = [-0.00, 0.07]) or between separated- and integrated-context bilinguals ($\beta = 0.01$, $SE = 0.02$, $t = 0.95$, $p = .557$, 95% CI = [-0.03, 0.05]) were observed. This suggests that varied-context bilinguals had a greater tendency to rely on contextual information, whereas separated-context bilinguals were better able to minimize the effect of the context bias.

Similar to the BY error rates results, no reliable between-group differences were observed for BY RTs (Separated vs. Integrated: $\beta = -0.02$, $SE = 0.03$, $t = -0.79$, $p = .432$, 95% CI = [-0.07, 0.03]; Separated vs. Varied: $\beta = 0.01$, $SE = 0.03$, $t = 0.26$, $p = .797$, 95% CI = [-0.04, 0.06]; Integrated vs. Varied: $\beta = 0.03$, $SE = 0.02$, $t = 1.08$, $p = .284$, 95% CI = [-0.02, 0.07]), indicating that the results observed likely reflect strategy differences in cognitive control recruitment and not differences in general processing speed. In the next section, we proceed to analyze individual differences in picture naming performance using the AY and BX measures extracted from the AX-CPT.

Table 9. Estimated coefficients from the mixed model on AX-CPT RTs

Fixed effects	Estimate	*SE*	*t*-values	*p*	Lower	Upper
(Intercept)	**2.37**	**0.02**	**118.59**	**0.000**	**2.34**	**2.41**
BY	**0.03**	**0.02**	**2.03**	**0.043**	**0.00**	**0.06**
AY	**0.23**	**0.02**	**10.44**	**0.000**	**0.18**	**0.27**
Integrated	-0.04	0.03	-1.35	0.180	-0.09	0.02
Varied	-0.01	0.03	-0.44	0.662	-0.06	0.04
BY*Integrated	0.01	0.02	0.33	0.739	-0.03	0.04
AY*Integrated	0.05	0.03	1.64	0.104	-0.01	0.10
BY*Varied	0.01	0.02	0.37	0.713	-0.03	0.05
AY*Varied	**0.06**	**0.03**	**2.00**	**0.048**	**0.00**	**0.12**

Random effects	Variance	*SD*	Correlation	
Intercept | item	0.0001	0.01		
Intercept | participant	0.0073	0.09		
BY | participant	0.0000	0.01	-0.81	
AY | participant	0.0064	0.08	0.08	0.53
Residual	0.0173	0.13		

Note. The BX condition and the separated-context bilingual group were set as the reference levels. Intercept reflects mean BX log RTs for separated-context bilinguals.

3.4. How Does a Bilingual's Interactional Context Mediate the Relation between Cognitive Control and Lexical Access?

The AX-CPT results reported above suggest group differences in cognitive control strategies that align with the hypotheses that were laid out in the introduction. On average, separated-context bilinguals showed a tendency to minimally rely on context processing, favoring engagement of reactive control processes. In turn, bilinguals from the varied context showed greater reliance on contextual information, favoring engagement of proactive control processes. Finally, performance for bilinguals from the integrated context seemed to fall somewhere in between the other two groups. In the individual difference analyses below, we report results for the three AX-CPT measures that significantly predicted picture naming performance (i.e., AY error rates, AY efficiency, and BX efficiency). We also report results for separated and varied-context bilinguals only, since no reliable patterns of association between AX-CPT and picture naming performance were found for integrated-context bilinguals.

Cognitive control and picture naming accuracy. For bilinguals in the separated and varied contexts, the individual difference analyses revealed a pattern of association between AY error rates and picture naming accuracy (Figure 6A). In the mixed logistic regression, there was a significant interaction between AY error rates and language for separated-context bilinguals (β = 0.68, *SE* = 0.31, *z* = 2.20, *p* = .028, 95% CI = [0.08, 1.29]), indicating a negative association between error rates and Spanish accuracy (β = -0.50, *SE* = 0.25, *z* = -2.00, *p* = .046, 95% CI = [-0.98, -0.01]), although no pattern of association emerged with English accuracy (β = 0.19, *SE* = 0.18, *z* = 1.08, *p* = .282, 95% CI = [-0.16, 0.53]). This suggests that, for these bilinguals, appropriate suppression of a context-driven bias might be a favorable strategy for accessing words in the dominant L1. For varied-context bilinguals, the opposite pattern emerged. A significant interaction between AY error rates and language (Table 10) revealed a positive association between error rates and Spanish accuracy[18] (β = 0.58, *SE* = 0.18, *z* = 3.16, *p* = .002, 95% CI = [0.22, 0.94]) but no

[18]A separate correlational analysis also revealed a positive association between mean Spanish picture naming accuracy and AY error rates for varied-context bilinguals (*r*(28) = 0.50, *p* = .004, 95% CI = [0.18, 0.73]). A statistical power analysis was performed for sample size estimation based on this correlation. With an alpha = .05 and power = 0.80, the projected sample size needed with a similar effect size (-0.45) for a two-tailed test is approximately N = 33. The sample size used in the correlation (N = 30) approximates this number.

association with English accuracy (β = 0.05, *SE* = 0.20, *z* = 0.23, *p* = .820, 95% CI = [-0.35, 0.44]). This suggests that, for varied-context bilinguals, L1 lexical access might be best supported by a greater tendency to rely on context processing.

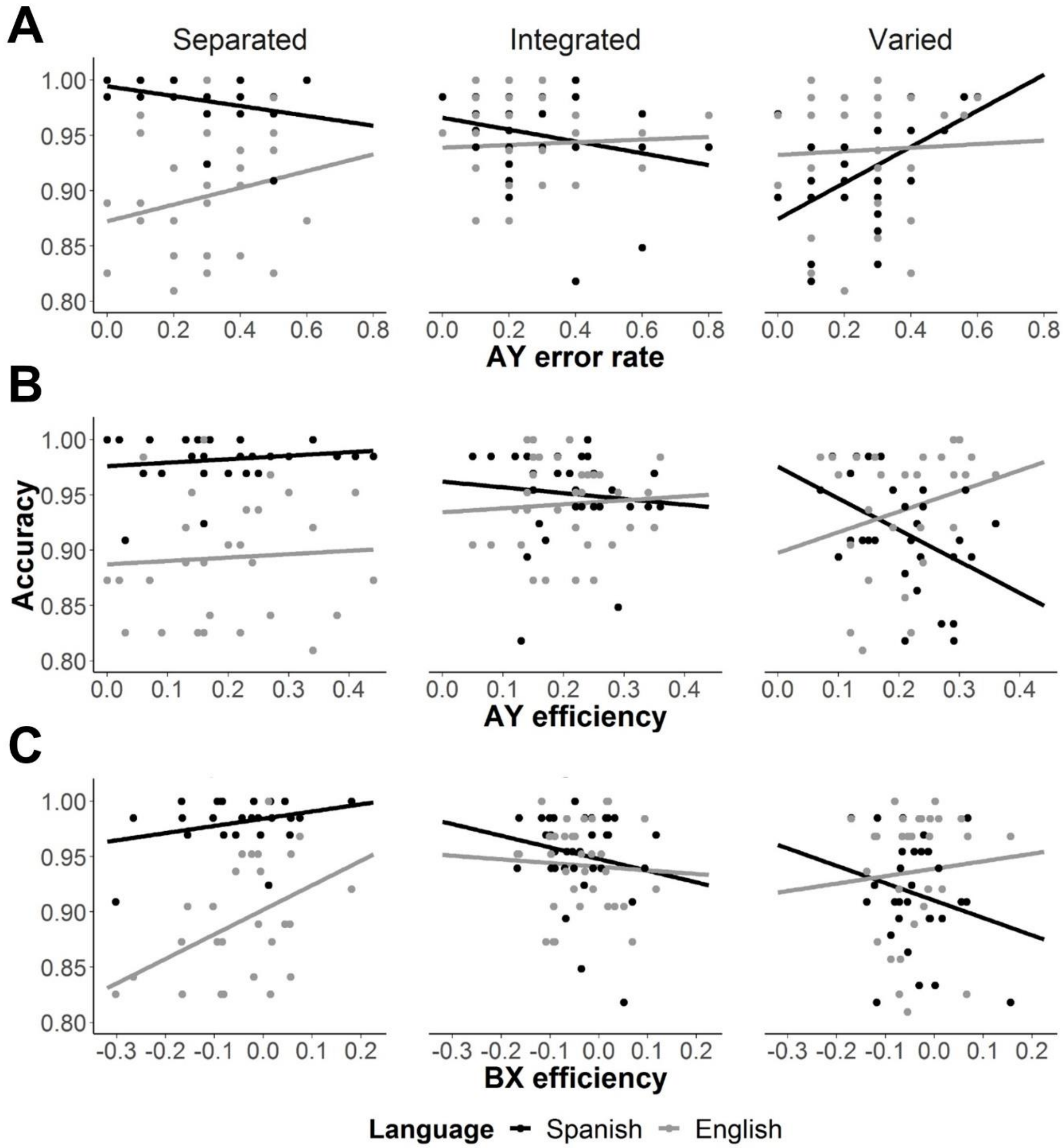

Figure 6. Relation between picture naming accuracy and AY error rates (A), AY efficiency (B), and BX efficiency (C) in Spanish and in English for individuals in each context. More positive values on the x-axes indicate the following: higher AY error rates (A); slower AY responses relative to BY trials (B); slower BX responses relative to BY trials (C).

Table 10. Estimated coefficients from mixed model of AY error rates on picture naming accuracy

Fixed effects	Estimate	*SE*	*z*-values	*p*	Lower	Upper
(Intercept)	**4.16**	**0.23**	**17.81**	**< 0.001**	**3.70**	**4.62**
Frequency	**1.41**	**0.18**	**7.85**	**< 0.001**	**1.06**	**1.76**
Language	-0.32	0.42	-0.76	0.449	-1.14	0.51
Separated	0.48	0.26	1.86	0.062	-0.02	0.98
Integrated	**0.64**	**0.25**	**2.59**	**0.009**	**0.16**	**1.13**
AY error	**0.31**	**0.13**	**2.38**	**0.018**	**0.05**	**0.57**
Separated*Language	**-2.69**	**0.45**	**-5.95**	**< 0.001**	**-3.58**	**-1.81**
Integrated*Language	**-0.77**	**0.39**	**-1.96**	**0.050**	**-1.55**	**0.00**
Language*AY error	**-0.50**	**0.17**	**-2.98**	**0.003**	**-0.82**	**-0.17**
Separated*AY error	**-0.46**	**0.20**	**-2.35**	**0.019**	**-0.85**	**-0.08**
Integrated*AY error	**-0.41**	**0.17**	**-2.42**	**0.015**	**-0.74**	**-0.08**
Separated*Language*AY error	**1.14**	**0.27**	**4.16**	**< 0.001**	**0.60**	**1.67**
Integrated*Language*AY error	**0.79**	**0.22**	**3.61**	**< 0.001**	**0.36**	**1.21**

Random effects	Variance	*SD*	Correlation	
Intercept | item	2.4535	1.57		
Separated | item	0.6530	0.81	-0.44	
Integrated | item	0.3669	0.61	0.14	-0.21
Intercept | participant	0.2348	0.48		
Language | participant	1.1733	1.08	0.15	

Notes. The varied-context bilingual group was set as the reference level. Intercept represents mean naming accuracy (in log odds) for varied-context bilinguals with an average frequency effect and average AY error rates.

In addition to the AY error rate results, a converging pattern of association emerged between the AY and BX efficiency measures and picture naming accuracy (Figures 4B and 4C). For varied-context bilinguals, a significant interaction between language and AY efficiency (Table 11) revealed an effect of AY efficiency on accuracy in Spanish (β = -0.49, *SE*, = 0.20, *z* = -2.41, *p* = .016, 95% CI = [-0.89, -0.09]) but not in English (β = 0.37, *SE* = 0.23,

z = 1.64, p = .101, 95% CI = [-0.07, 0.82]). This effect indicated that higher AY efficiency (i.e., a smaller RT difference between AY and BY) predicted higher accuracy in Spanish[19]. The interaction also revealed that, for individuals with high AY efficiency (1 SD below the mean), Spanish naming had higher accuracy than English naming (β = -1.17, SE = 0.53, z = 2.22, p = .027, 95% CI = [-2.21, -0.14]), but for individuals with low AY efficiency (1 SD above the mean), naming accuracy was similar in the two languages (β = 0.58, SE = 0.51, z = 1.12, p = .260, 95% CI = [-0.43, 1.59]). This suggests that, for bilinguals in the varied context, the ability to efficiently resolve context-driven interference might help maintain fluid lexical access in Spanish when immersed in an English-predominant environment that does not support the use of Spanish.

[19] The correlational analysis also revealed a negative association between mean Spanish picture naming accuracy and AY efficiency for varied-context bilinguals (r(28) = -0.45, p = .013, 95% CI = [-0.70, -0.11]).

Table 11. Estimated coefficients from the mixed model of AY efficiency on picture naming accuracy

Fixed effects	Estimate	*SE*	*z*-values	*p*	Lower	Upper
(Intercept)	**4.15**	**0.24**	**17.64**	**< 0.001**	**3.69**	**4.61**
Frequency	**1.41**	**0.18**	**7.87**	**< 0.001**	**1.06**	**1.76**
Language	-0.30	0.42	-0.71	0.480	-1.12	0.53
Separated	0.46	0.26	1.79	0.073	-0.04	0.96
Integrated	**0.66**	**0.25**	**2.61**	**0.009**	**0.17**	**1.16**
AY efficiency	-0.06	0.15	-0.39	0.694	-0.35	0.23
Separated*Language	**-2.64**	**0.45**	**-5.90**	**0.000**	**-3.52**	**-1.77**
Integrated*Language	**-0.82**	**0.40**	**-2.05**	**0.041**	**-1.60**	**-0.03**
Language*AY efficiency	**0.88**	**0.31**	**2.83**	**0.005**	**0.27**	**1.48**
Separated*AY efficiency	0.15	0.19	0.79	0.428	-0.22	0.52
Integrated*AY efficiency	-0.02	0.21	-0.10	0.918	-0.43	0.39
Separated*Language*AY efficiency	**-1.02**	**0.24**	**-4.17**	**< 0.001**	**-1.49**	**-0.54**
Integrated*Language*AY efficiency	**-0.67**	**0.26**	**-2.59**	**0.010**	**-1.17**	**-0.16**

Random effects	Variance	*SD*	Correlation	
Intercept | item	2.4555	1.57		
Separated | item	0.6530	0.81	-0.44	
Integrated | item	0.3699	0.61	0.15	-0.27
Intercept | participant	0.2671	0.52		
Language | participant	1.2015	1.10	0.05	

Notes. The varied-context bilingual group was set as the reference level. Intercept represents mean naming accuracy (in log odds) for varied-context bilinguals with an average frequency effect and an average AY efficiency score.

On the other hand, there was an effect of BX efficiency on picture naming accuracy for separated-context bilinguals (Figure 4C), such that lower BX efficiency (i.e., slower BX responses relative to BY) predicted better overall accuracy (Table 12). This suggests that for bilinguals in a separated context, less reliance on context-driven processing, and therefore greater reliance on reactive control processes, might be a beneficial control strategy when using one language at the expense of the other.

Table 12. Estimated coefficients from the mixed model of BX efficiency on picture naming accuracy

Fixed effects	Estimate	*SE*	*z*-values	*p*	Lower	Upper
(Intercept)	**4.62**	**0.26**	**17.90**	**< 0.001**	**4.12**	**5.13**
Frequency	**1.40**	**0.18**	**7.75**	**< 0.001**	**1.05**	**1.76**
Language	**-2.91**	**0.47**	**-6.22**	**< 0.001**	**-3.83**	**-2.00**
Varied	**-0.52**	**0.26**	**-1.99**	**0.046**	**-1.02**	**-0.01**
Integrated	0.18	0.30	0.60	0.549	-0.41	0.77
BX efficiency	**0.26**	**0.11**	**2.47**	**0.014**	**0.05**	**0.47**
Varied*Language	**2.64**	**0.47**	**5.66**	**< 0.001**	**1.73**	**3.56**
Integrated*Language	**1.79**	**0.51**	**3.53**	**< 0.001**	**0.79**	**2.78**
Language*BX efficiency	-0.01	0.23	-0.06	0.954	-0.47	0.44
Varied*BX efficiency	-0.31	0.17	-1.93	0.054	-0.64	0.01
Integrated*BX efficiency	**-0.41**	**0.17**	**-2.34**	**0.019**	**-0.75**	**-0.07**
Varied*Language*BX efficiency	0.36	0.38	0.94	0.346	-0.39	1.12
Integrated*Language*BX efficiency	0.19	0.38	0.50	0.617	-0.56	0.94

Random effects	Variance	*SD*	Correlation	
Intercept | item	2.0838	1.44		
Varied | item	0.6401	0.80	-0.14	
Integrated | item	1.2355	1.11	-0.08	0.84
Intercept | participant	0.2289	0.48		
Language | participant	1.2961	1.14	0.00	

Notes. The separated-context bilingual group was set as the reference level. Intercept represents mean naming accuracy (in log odds) for separated-context bilinguals with an average frequency effect and an average BX efficiency score.

Cognitive control and picture naming RTs. A converging pattern was observed between picture naming RTs and AY efficiency scores for bilinguals in the varied context (Figure 5A). The analysis yielded a significant interaction between language block and AY efficiency (Table 13). Follow-up simple effects analyses revealed that naming in the Spanish block was slower than naming in the English block for individuals with low AY efficiency (β = -0.05, = 0.02, t = -2.37, p = .019, 95% CI = [-0.09, -0.01]), although naming speed was similar in the two languages for individuals with high AY efficiency (β = 0.02, SE

= 0.02, t = 0.87, p = .385, 95% CI = [-0.02, 0.06]). This suggests that only bilinguals in the varied context with the most efficient context processing were able to maintain Spanish lexical retrieval speed on par with English.

Table 13. Estimated coefficients from the mixed model of AY efficiency on picture naming RTs

Fixed effects	Estimate	*SE*	*t*-values	*p*	Lower	Upper
(Intercept)	**3.03**	**0.01**	**232.15**	**< 0.001**	**3.01**	**3.06**
Lang	-0.01	0.01	-0.96	0.336	-0.04	0.01
Freq	**-0.05**	**0.01**	**-8.52**	**< 0.001**	**-0.06**	**-0.04**
Separated	**0.05**	**0.02**	**2.69**	**0.008**	**0.01**	**0.08**
Integrated	-0.01	0.02	-0.51	0.611	-0.04	0.02
AY efficiency	-0.00	0.01	-0.04	0.968	-0.03	0.03
Lang*Freq	0.00	0.01	0.32	0.750	-0.02	0.03
Lang*Separated	**0.11**	**0.02**	**6.53**	**< 0.001**	**0.08**	**0.14**
Lang*Integrated	0.01	0.02	0.91	0.365	-0.02	0.05
Freq*Separated	0.01	0.01	1.34	0.183	0.00	0.02
Freq*Integrated	0.00	0.01	0.29	0.775	-0.01	0.01
Lang*AY efficiency	**-0.03**	**0.01**	**-2.49**	**0.015**	**-0.06**	**-0.01**
Freq*AY efficiency	-0.00	0.00	-0.72	0.475	-0.01	0.00
Separated*AY efficiency	0.02	0.02	0.92	0.359	-0.02	0.05
Integrated*AY efficiency	-0.00	0.02	-0.01	0.991	-0.04	0.04
Lang*Freq*Separated	**-0.05**	**0.01**	**-5.07**	**< 0.001**	**-0.06**	**-0.03**
Lang*Freq*Integrated	-0.00	0.01	-0.16	0.870	-0.02	0.02
Lang*Freq*AY efficiency	**0.01**	**0.01**	**1.97**	**0.049**	**0.00**	**0.02**
Lang*Separated*AY efficiency	**0.02**	**0.01**	**3.69**	**< 0.001**	**0.01**	**0.04**
Lang*Integrated*AY efficiency	**0.04**	**0.01**	**5.22**	**< 0.001**	**0.02**	**0.05**
Freq*Separated*AY efficiency	0.00	0.00	0.44	0.662	-0.01	0.01
Freq*Integrated*AY efficiency	0.01	0.00	1.66	0.100	0.00	0.02
Lang*Separated*Freq*AY efficiency	-0.01	0.01	-1.53	0.127	-0.02	0.00
Lang*Integrated*Freq*AY efficiency	**-0.02**	**0.01**	**-2.04**	**0.041**	**-0.03**	**-0.00**

Random effects	Variance	*SD*	Correlation	
Intercept | item	0.0033	0.06		
Separated | item	0.0011	0.03	-0.43	
Integrated | item	0.0009	0.03	-0.05	0.34
Intercept | participant	0.0042	0.06		
Lang | participant	0.0029	0.05	-0.09	
Freq | participant	0.0001	0.01	-0.07	-0.15
Residual	0.0144	0.12		

Notes. Lang = Language; Freq = Frequency. The varied-context bilingual group was set as the reference level. Intercept represents mean log RTs for varied-context bilinguals with an average frequency effect and an average AY efficiency score.

Does cognitive control aid in the retrieval of lexical items that are more prone to retrieval difficulties (i.e., low frequency words)? We ask this question given recent claims that the relative engagement of cognitive control during language processing depends on whether such processing involves linguistic information that is effortful and conflict-prone (Hsu & Novick, 2016; Nozari & Novick, 2017). A significant three-way interaction between AY efficiency, language, and frequency for varied-context bilinguals (Table 13) indicated that the effect of AY efficiency on language block depended on frequency status. As Figure 7B shows, individuals with lower AY efficiency were slower in Spanish than in English when naming low frequency words ($\beta = -0.06$, $SE = 0.02$, $t = -2.59$, $p = .010$, 95% CI = [-0.11, -0.02]). On the other hand, naming speed was similar in the two language for individuals with high AY efficiency ($\beta = 0.02$, $SE = 0.02$, $t = 1.04$, $p = .316$, 95% CI = [-0.02, 0.07]). No modulations of the language effect were observed at high and low levels of AY efficiency when naming high frequency words (Low AY efficiency: $\beta = -0.03$, $SE = 0.02$, $t = -1.45$, $p = .149$, 95% CI = [-0.08, 0.01]; High AY efficiency: $\beta = 0.01$, $SE = 0.02$, $t = 0.48$, $p = .632$, 95% CI = [-0.04, 0.06]). This suggests that when lexical access is most effortful (i.e., when retrieving Spanish low frequency words in an English-predominant environment), cognitive control might facilitate the retrieval process. More generally, these results are in line with the adaptive control hypothesis (Green & Abutalebi, 2013) in that cognitive processes adapt to the demands of the environment. For bilinguals in the varied context, English lexical access is facilitated by the predominant English context. However, these individuals face the challenge of maintaining Spanish in a dynamic environment where there is a constant need for monitoring the appropriateness of using one of both languages.

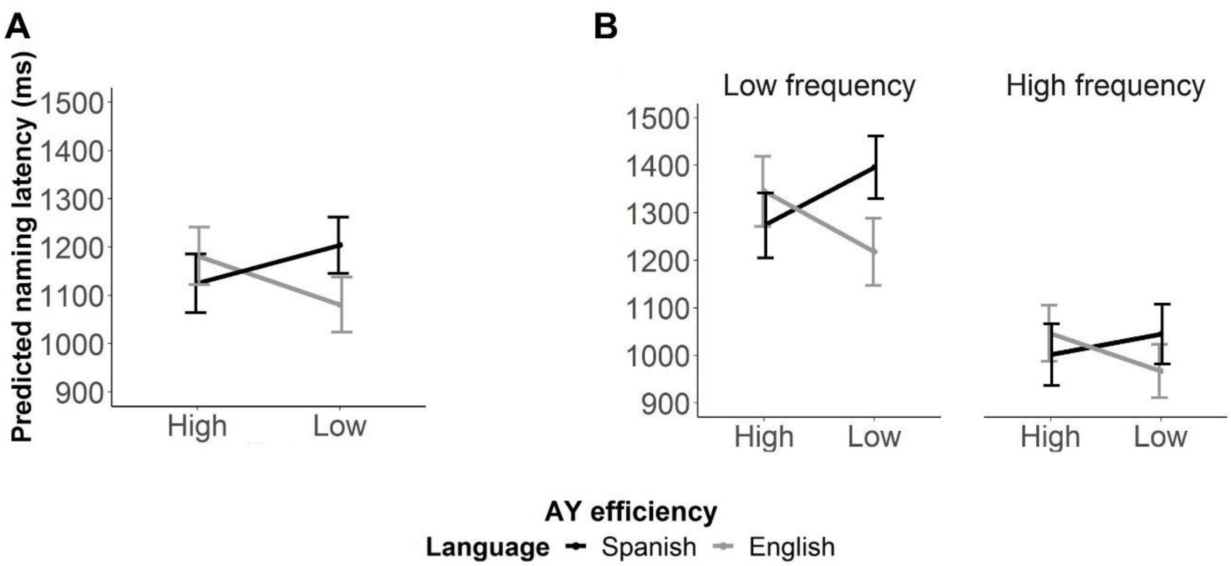

Figure 7. Two-way interaction between language and AY efficiency (A), and a three-way interaction between language, frequency, and AY efficiency (B), for bilinguals in the varied context. Error bars represent standard error of the mean.

4. Discussion

The present study sought to characterize the consequences of the context in which bilinguals use their two languages to better understand the way that bilingualism draws upon cognitive resources. By examining language and cognitive factors in tandem, this work gives insight into how bilinguals differ amongst themselves. Our findings suggest that the engagement of cognitive control depends on the demands of the language environment, at least once a critical threshold of proficiency has been achieved. This is consistent with the adaptive control hypothesis (Green & Abutalebi, 2013) and with recent empirical evidence indicating how different contexts of language use affect cognitive control ability (Gullifer et al., 2018; Hartanto & Yang, 2016; Ooi et al., 2018; Pot et al., 2018). Notably, all of the bilinguals in the present study reached a level of picture naming accuracy and verbal fluency that is indicative of high proficiency in both Spanish and English. In most past research, these bilinguals might have well been aggregated into a single bilingual group to be compared to monolingual speakers. The results we have presented show that aggregating data in a way that ignores the context of language use is likely to mask the relation between language and the cognitive control processes that support them.

4.1. Implications for Language Production

In the past literature on lexical access in bilinguals, there has been an ongoing debate as to whether the costs to production in each of the bilingual's two languages should be attributed to cross-language competition or to functionally lower frequency because the use of two languages necessarily reduces the time available to speak each language (see Kroll & Gollan, 2014, for a review). The typical pattern that has been observed in picture naming is slower RTs and larger frequency effects in the L2 relative to the L1 (e.g., Gollan et al., 2008). In the present study, the frequency asymmetries in picture naming followed the predictions of the frequency-lag account only for bilinguals in the separated context. Bilinguals in neither the integrated nor the varied context showed the predicted pattern. In particular, bilinguals in the varied context had reversed language dominance, with greater dominance in the L2 than the L1, yet the frequency effects were not asymmetric in the way that might be predicted. Other recent studies (e.g., Sullivan et al., 2017) have challenged the frequency-lag account on the grounds that trilinguals, who presumably divide their time even more finely than bilinguals, did not produce frequency effects that differed from bilinguals. While the frequency-lag account cannot adequately account for performance across the three bilingual contexts in the present study, neither can the competition-for-selection alternative easily provide a simple interpretation for the observed differences.

The finding that codeswitching experience appears to be associated with more symmetrical frequency effects thus adds a further dimension to this discussion: habitually switching between languages may have enduring consequences for the language control network. This is consistent with the control processes model (CPM; Green, 2018; Green & Wei, 2014), which posits that the dynamics of bilingual language control are directly mediated by the speaker's intention to use the languages in specific ways. Under the CPM, bilinguals in separated contexts engage language control *competitively* (i.e., where the activation of one language is suppressed at the expense of the other). In turn, bilinguals in integrated contexts engage language control *cooperatively* (i.e., where co-activation is maintained all the way through speech planning so that items from both languages can be used opportunistically). Critically, a given control state is hypothesized to result in a "habit of control" with repeated use (Green & Abutalebi, 2013). Therefore, for a bilingual who relies on a cooperative control state, the relevance of language membership is minimized; conversely, the relevance of language membership is maximized for a bilingual in a competitive control state.

One inadvertent consequence of a cooperative control state may be that information from lexical items (such as frequency) in one language can be mapped on to lexical equivalents in the other language due to greater cross-language overlap. In turn, under a competitive control state, information flow from one language to the other is restricted. In this scenario, lexical retrieval may be largely dependent on the functional use of each language, which may create a profile in which one language is more dominant than the other, even at high levels of proficiency.

Notably, most of the frequency asymmetries in bilinguals come from work on Spanish heritage speakers who are dominant in English (Kroll & Gollan, 2014; cf., Gollan et al., 2011; Ivanova & Costa, 2008[20]). In general, these individuals initially acquire Spanish as their L1 at home, but then become educated in English almost exclusively. Critically, the use of Spanish is often limited to the home environment, whereas English remains the dominant language in most contexts. We hypothesize that the heritage speakers from previous studies, together with the bilinguals from the separated context reported here, share the trait of predominantly using their languages in relative isolation. To the extent that this is true, we hypothesize that frequency asymmetries are more likely to emerge in bilinguals who primarily use their languages in a separated (i.e., competitive) fashion. However, we note that this explanation is speculative. Future work will need to assess variation in codeswitching experience, drawing on corpus-driven and experimental research that can help identify different contexts of language use that differentiate bilingual communities, including those who speak the same languages (Beatty-Martínez et al., 2018).

[20] Gollan et al. (2011) and Ivanova and Costa (2008) tested non-heritage Dutch-English and Spanish-Catalan bilinguals, respectively. However, based on the information provided in both studies, it is difficult to identify a context of language use for each of these groups. Additionally, picture naming performance in these studies was assessed in one language only (i.e., English in Gollan et al., and Spanish in Ivanova & Costa), making it difficult to examine frequency effects in a way that is comparable to the present study.

4.2. Implications for the Relation between Language Production and Cognitive Control

Some past studies have suggested that language proficiency determines the need for cognitive control in planning speech in the two languages (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Meuter & Allport, 1999). From this perspective, bilinguals with the largest language asymmetries (i.e., those in the separated context) should rely more on inhibitory control to adjust the dominant language when speaking the weaker language. However, in more recent work, better inhibitory control has been shown to positively predict speaking abilities both in the dominant and less dominant language (Linck et al., 2012; Pivneva, Mercier, & Titone, 2012). This is consistent with the pattern of associations found for bilinguals in the separated context, where greater reliance on reactive control processes (as indexed by lower AY error rates and slower BX vs. BY responses) predicted higher picture naming accuracy in both the L1 and in the L2. For these individuals, reliance on context monitoring may not be as crucial given that in a single-language context, bilinguals can use their lifelong experience to reliably predict which language will be used in a given domain (e.g., Spanish at home). However, in some cases, this expectation might not be met (e.g., when a foreign exchange student enters a conversation prompting a change in language among a group of Spanish-speaking friends), which may trigger a need for reactive control processes to suppress potential cross-language intrusions and guarantee fluent retrieval in the target language.

Unlike the results for bilinguals in the separated context, greater reliance on context processing (as indexed by AY error rates) seemed to be critical for bilinguals in the varied context, particularly for maintaining lexical accessibility in the L1. One interpretation is that proactive control processes involving monitoring are more likely to be recruited by bilinguals who are immersed in an environment where the types of conversational exchanges are diverse. This explanation is consistent with two recent studies examining the relation between social diversity and cognitive control ability. A study by Pot et al. (2018) reported that increased L2 usage across different social domains, together with self-reported switching behavior, predicted better flanker performance. Similarly, Gullifer et al. (2018) found that increased social diversity in language use (using a measure of language entropy) and increased reliance on proactive control (using an AY-BX index from the AX-CPT) predicted greater functional connectivity between brain regions that are typically associated with conflict detection and monitoring processes (Abutalebi et al., 2012; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004; Kerns, et al., 2004). Spanish-English bilinguals immersed in an L2 environment, such as the varied-context bilinguals tested in the present study, have to carefully monitor with whom they are able to use just one or both of their languages and with whom they might codeswitch. Such heterogeneity may hinder the ability to rely on a single form of conversational exchange across contexts, creating a greater need for proactive control processes.

The efficiency in which varied-context bilinguals overcame the context bias in the AX-CPT (as indexed by the AY efficiency measure) was also predictive of their picture naming performance, with greater accuracy and faster naming in Spanish for those who showed greater efficiency. A similar pattern was reported by Zirnstein and colleagues (2018), who tested a group of Mandarin-English bilinguals immersed in their L2 (English), and found that faster AY responses were associated with more efficient recovery of prediction errors while reading sentences in their L2, although this pattern of association depended on their ability to regulate the activation of their L1 (Mandarin). According to Morales and colleagues (2013), faster AY vs. BY responses could be achieved by relying on a proactive conflict-resolution strategy (i.e., preparing for potential probe conflict in advance) or by engaging response inhibition (i.e., suppressing a prepotent incorrect response in reaction to the probe). In an attempt to dissociate these two possibilities, we created a delta plot for AY and BY RTs following the procedures used by Morales and colleagues (see Figure 6). A visual inspection of the delta plot suggested that, for separated-context bilinguals, high AY efficiency was likely achieved via response inhibition, replicating the pattern reported by Morales and colleagues. However, for varied-context bilinguals, a different pattern emerged that was suggestive of a proactive conflict-resolution strategy, which is consistent with the previously described association between AY error rates and picture naming accuracy.

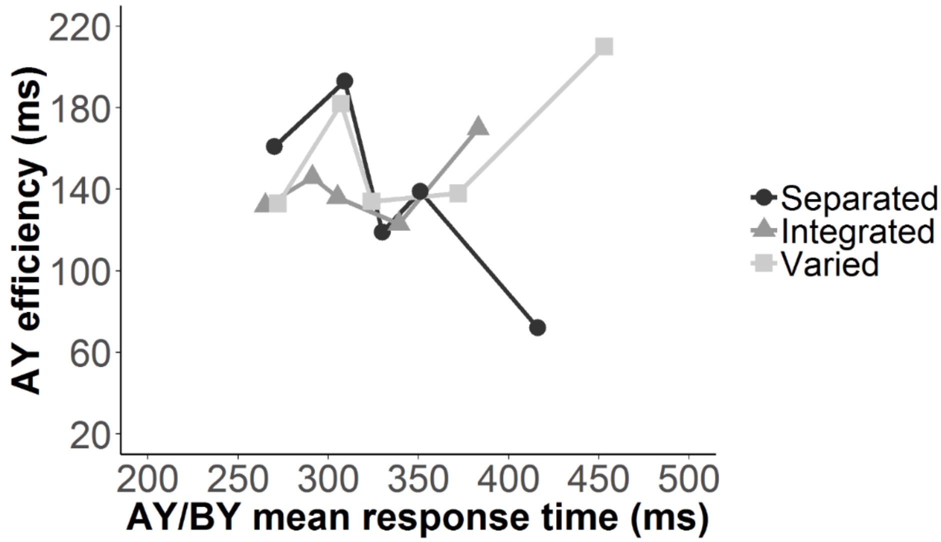

Figure 8. Delta plots showing the condition difference as a function of quintile scores across bilinguals in each context. Delta plots show an effect size (i.e., the difference between AY and BY RTs) as a function of response speed across participants (i.e., the average AY and BY RT for any given participant). This is achieved by ordering and dividing RTs for each participant into quintiles. More positive values on the y-axis indicate slower AY responses relative to BY trials. More positive values on the x-axis indicate individuals with slower RTs across the two conditions. Response inhibition is typically assumed to require time to unfold (De Jong, Liang, & Lauber, 1994; Ridderinkhof, Scheres, Oosterlaan, & Segeant, 2005). In this case, reduced interference effects (i.e., high AY efficiency) should emerge for individuals with slower overall responses. This pattern is observed for bilinguals in the separated context, which is also consistent with the pattern reported by Morales and colleagues (2013), who also tested Spanish-English bilinguals from the same community in Spain. On the other hand, bilinguals in integrated and varied contexts showed the opposite trend (i.e., high AY efficiency emerged for individuals with faster overall responses),

suggesting that, for these individuals, high AY efficiency is achieved via context monitoring procedures.

For bilinguals in an L2-predominant environment, L1 lexical access is likely more susceptible to interference, especially when attempting to retrieve low frequency words. Bilinguals in such a scenario can attempt to detect contextual cues that may signal upcoming conflict. This would create an opportunity to preemptively limit the activation of competing information, be it within-language or cross-language competitors. This explanation is consistent with recent research in cognitive control and language processing showing that prior conflict detection improves subsequent conflict-related language performance (e.g., Freund et al., 2016; Hsu & Novick, 2016; Navarro-Torres, Garcia, Chidambaram, & Kroll, 2019; Thothathiri, Asaro, Hsu, & Novick, 2018). A recent study by Navarro-Torres et al. (2019) tested a group of monolinguals and L2-immersed bilinguals, and found that Stroop-related conflict facilitated recovery from syntactic ambiguity in spoken sentence comprehension. However, unlike monolinguals, bilinguals initiated this recovery process more quickly (i.e., before encountering the ambiguity) by relying on linguistic cues that appeared early on in the spoken sentence, which could be used to anticipate the ambiguity. In the context of the present results, varied-context bilinguals with high proactive control efficiency may be able to exploit environmental cues to initiate cognitive control recruitment (e.g., when the desire or requirement is to speak in Spanish in the presence of English cues), which may subsequently facilitate the retrieval of less accessible information in the L1. We note, however, that the results in the present study are correlational, thus making it difficult to establish a causal relation[21]. Future research will be needed to identify the critical features of the immersion environment that reflect this aspect of coordination.

Many new questions are raised by the findings reported here. One of them concerns the role of codeswitching. By some accounts (e.g., Hofweber et al., 2016; Verreyt, Woumans, Vandelanotte, Szmalec, & Duyck, 2016; Yim & Bialystok, 2012), codeswitching is considered the feature of bilingual language use that is most critical in determining the way that the two languages are regulated. Proficient codeswitching appears to be relatively seamless, yet exquisite skill is required to observe the constraints that underlie acceptable performance. It may seem striking that the sentence and discourse level requirements associated with codeswitching did not, on their own, determine the degree to which lexical production engaged cognitive resources. Bilinguals from both integrated and varied contexts in the present work actively codeswitch. Yet, only the bilinguals from the varied context showed modulation of picture naming performance via proactive control. However, this result may not seem as striking considering that the integrated-context bilinguals come from a stable codeswitching community (Beatty-Martínez, 2019; Guzzardo Tamargo et al. 2018; Pousada, 2017). In these cases, proficient codeswitching behavior often follows conventionalized forms of language use that have been adopted by the community at large (Beatty-Martínez & Dussias, 2017; Guzzardo Tamargo et al., 2016; Herring; Deuchar,

[21] It is possible that the interactional context is responsible for inducing these adaptive changes in how cognitive control is utilized during lexical access. Conversely, individuals with particular cognitive control profiles may be more likely to thrive linguistically in specific contexts. We note that in either case, the interactional context would have an important meditative role in the relation between lexical access and cognitive control.

Parafita Couto, & Moro Quintanilla, 2010; Parafita Couto, Munarriz, Epelde, Deuchar, & Oyharçabal, 2015; Poplack, Zentz, & Dion, 2012; Valdés Kroff et al., 2016). Given sufficient regularity, these individuals may be able to exploit a variety of linguistic cues and adopt social-discourse strategies to anticipate (and signal) switches (Beatty-Martínez, 2019; Fricke & Kootstra, 2016; Fricke et al., 2016; Torres Cacoullos & Travis, 2018), thus minimizing the need for the recruitment of the control processes during language production (Green & Abutalebi, 2013). But again, the point is that no single feature of language experience appears to be sufficient to account for all of the findings.

5. Conclusions

The results we have reported represent a preliminary step towards characterizing aspects of bilingual experience that may be crucial for understanding how the two languages are used in ways that draw differentially on cognitive resources. What is important to consider is that, by most past accounts, the bilinguals from the three contexts examined here would be assumed to represent samples drawn from the same or similar underlying populations of proficient Spanish-English speakers. Yet we have shown that the different contexts in which the two languages are spoken have clear consequences for the way that cognitive control is engaged to enable language production. Like other recent studies, the current results suggest that bilinguals who are immersed in their L2, will vary in how well they regulate their native language. Those with lower levels of proactive control, as indexed by performance on the AX-CPT, appear to succumb to the pressures of the environment, becoming better able to retrieve the L2 but at the expense of the L1. Those with higher levels of proactive control engagement appear to maintain the privileged access typically associated with the native language. In the present study, it is difficult to interpret the source of these individual differences. Although it is possible that they are stable attributes of the individual bilinguals, it is also possible that they are a consequence, at least to some degree, of language experience.

What are the implications of the results we have presented for the controversies about the cognitive consequences of bilingualism? The present study was not designed to examine this issue directly, but we believe that there are important implications for considering the ways in which the questions about cognitive consequences have been framed. In the literature on the consequences of bilingualism for executive function, some have used the fact that different executive function tasks produce different outcomes with respect to bilingual effects as a basis on which to dismiss the entire enterprise (de Bruin et al., 2015a; Valian, 2015). However, given the pattern of associations observed between AX-CPT and picture naming performance, it may be possible to generate hypotheses about cognitive consequences more generally based on the three contexts examined here. Separated contexts might be more likely to induce changes in reactive control processes, while varied contexts might result in the strengthening of proactive control processes. In this regard, the findings of the between-group differences of BX error rates may be taken by some as evidence for cognitive advantages in varied-context bilinguals. Nevertheless, as it has been argued elsewhere (Gullifer et al., 2018), it is difficult to identify advantages in a complex task such as the AX-CPT, since lower error rates and/or faster responses are not necessarily indicative of better executive functions. It is also possible that each interactional context might strengthen a more diverse set of executive control processes, but that only a subset of these processes are critical for those aspects of language examined here. Regardless, the current results are, if anything, a call to action to better understand how these tasks function and to increase, rather than reduce, the complexity of the exercise by including language processing tasks that also differ in the way that they may draw on aspects of domain general cognition (Baum & Titone, 2014).

An important feature of the present results is that they are behavioral. In the controversy about when you see the consequences of bilingualism and when you do not, there has been consideration given to the fact that measures of brain activity are often more likely to reveal these effects than standard behavioral measures of RT and accuracy. Many recent studies have shown dissociations between behavioral and neurocognitive measures that suggest that brain activity often provides a more sensitive index of both early language processes (e.g., Bice & Kroll, 2015; McLaughlin et al., 2004) and of the consequences of life long bilingualism (e.g., Gullifer et al., 2018; Kousaie & Phillips, 2017) and short-term intensive language training (e.g., Zhang et al., 2015). The current results do not address this issue since measures of brain activity were not included, but critically, they suggest that it is possible to observe reliable differences in behavior that reflect the consequences of language use. In future research, it will be important to better understand the mappings between brain activity and behavior for the contextual differences we have identified here.

Contrary to the view that failures to replicate the consequences of bilingualism are due to noise (Luk & Bialystok, 2013; Valian, 2015), the results we report suggest that systematic variation in language use may determine the pattern of consequences that are observed.

Bilingualism is a complex life experience. Characterizing that complexity, particularly with respect to language use, will be critical to fully understanding the cognitive consequences of life in two languages.

PART B

1. Introduction

The study by Beatty-Martinez, Navarro-Torres et al. (2020) suggests that how cognitive control is engaged to enable fluent lexical access depends on bilinguals' interactional context. Recall that better picture naming performance in bilinguals from the separated context emerged in those who adopted reactive control tendencies, whereas no association between naming performance and cognitive control was found in bilinguals from the integrated context. In turn, for bilinguals in the varied context, those with proactive control tendencies were the ones who generated better picture naming performance, particularly in the native L1, which is the language no longer supported by the environment. This is arguably due to the need to actively engage language regulation in a manner that maximizes fluency in the dominant language of the environment (English) while still retaining high fluency in Spanish. However, although the types of analyses reported in that study are rich and complex, they do not directly assess whether language regulation, as defined in Chapter 4, mediates (or whether it is separate from) the observed cognitive control effects in a similar manner to the interactive effects observed in the L2-immersed Mandarin-English bilinguals from the Zirnstein et al. (2018) study.

In this portion of the chapter, I follow up on the results reported in Beatty-Martínez, Navarro-Torres et al. (2020) and provide a reanalysis on a subset of the data across the three contexts. One possibility examined in this reanalysis is that the involvement of language regulation yields an interaction between fluency in the dominant language (a measure not included in the original analyses to examine individual differences in picture naming performance) and AX-CPT performance. More specifically, the effect of cognitive control (either in the form of reactive or proactive control reliance) may depend on bilinguals' ability to up-regulate accessibility of the dominant language, as measured as higher dominant-language fluency. I particularly focus on errors rates from the AY condition, which yielded significant effects for both separated and varied contexts, because they allow us to examine proactive and reactive control engagement in a predictive context (see Section 2.2 in Part A of this chapter), which may better reflect the types of interactional demands that arise in the varied context.

To reiterate, the three contexts of bilingualism examined in Beatty-Martínez et al. (2020) vary not only in the types of habitual conversational exchanges that take place, but they also vary with respect to the degree of support that the environment provides for each language. For instance, for separated-context bilinguals in Granada, Spain, the cues for using L1 and L2 are quite unambiguous, such that the L1, Spanish, is the dominant language of the community and the separate use of the two languages is associated with different contexts. For integrated-context bilinguals in San Juan, In Puerto Rico, support for both languages is relatively egalitarian, although the L1 Spanish is still the predominant language. Finally, for varied-context bilinguals in the US, it is the L2 English that is consistently supported by the environment, with very limited (and often no) support for the L1. Thus, the prediction is that bilinguals from the varied context should be more likely to reveal an interaction as measured by their English fluency. In this context, those with better regulatory skill may be more apt at engaging engage proactive control to facilitate lexical access. Although it is feasible to also hypothesize that bilinguals from the separated context engage language regulation under some circumstances, the fact is that they live in an environment where the dominant L1 is reliably supported, making this context less likely to reveal regulatory effects via an interaction. Lastly, bilinguals from the integrated context may be the least likely to engage regulation, as they live in an environment where the two languages are relatively well supported and can therefore minimize the interactional costs that could emerge when using either language.

2. Method

2.1. Participants

The participants included in this study are the same as those recruited in the Beatty-Martínez, Navarro-Torres et al. (2020) study (see Section 2.1 in Part A of this chapter for a full characterization of the participants).

2.2. Materials

The three measures included in the reanalysis (i.e., category verbal fluency, picture naming, and AX-CPT) were identical to those reported in Beatty-Martínez, Navarro-Torres et al. (2020). See Part A (Section 2.2.) of this chapter for details on the materials.

2.3. Procedure

The procedure was identical to the one reported in Beatty-Martínez, Navarro-Torres et al. (2020). See Part A (Section 2.3.) of this chapter for details on the procedure.

2.4. Analysis

For picture naming and AX-CPT, although both response times (RTs) and accuracy were reported in Beatty-Martínez, Navarro-Torres et al., only naming accuracy and error rates are included in the analyses reported here. Verbal fluency scores were extracted using the average number of exemplars produced across categories in Spanish and in English.

Statistical analyses were performed to examine individual differences in picture naming accuracy using logistic mixed-effects models in the lme4 software package (version 1.1-18-1; Bates et al., 2015) in the R programming environment (version 3.5.1; R Development Core Team, 2014). Given the addition of verbal fluency data, each group was analyzed separately to minimize the probability of overfitting the data relative to the sample size of each group. Each analysis included a contrast coded fixed effect of language block (Spanish = -0.5, English = 0.5), log-transformed word frequency values (measured as a continuous factor), AY error rates from the AX-CPT, verbal fluency scores in either Spanish or English (each analyzed in separate models). Each factor was allowed to interact with one another except for word frequency, which was included as a covariate to control for frequency effects. The random effects structure contained random intercepts for subjects and items and by-participant random slopes for frequency. Continuous fixed-effects were z-scored to make the intercept in the models reflect average performance.

Significant interactions were examined by refitting a model with a dummy coded categorical factor to examine simple slopes at each level of the categorical factor, or by rescaling continuous factors one SD above/below the mean to examine simple effects of categorical factors at high and low values of the continuous factor. For example, a significant interaction between language block and fluency might indicate that the effect of language changes at high or low levels of fluency (which can be assessed by rescaling fluency within the model 1 SD above or below the mean) but can also indicate that there are continuous effects of fluency significantly different from zero in one of the language blocks.

3. Results

3.1. Separated context

Tables 14 and 15 provide the model estimates for separated-context bilinguals with Spanish and English fluency, respectively. The mixed logistic models confirmed a continuous effect of word frequency, indicating that naming accuracy improved with increasingly higher frequency words. After controlling for the effect of frequency, the effect of language was also significant, indicating that on average Spanish yielded higher naming accuracy than English. Further, there was a significant interaction between language and AY error rates, indicating that those who had reduced AY error rates had higher naming accuracy in the L1 Spanish (Spanish fluency model: $\beta = -0.48$, $SE = 0.23$, $z = -2.05$, $p = .040$; English fluency model: $\beta = -0.43$, $SE = 0.21$, $z = -2.08$, $p = .037$). These results replicate those reported by Beatty-Martinez, Navarro-Torres et al. (2020).

In the Spanish fluency model, fluency interacted with language. Follow-up simple effects analyses revealed that higher Spanish fluency predicted lower naming accuracy in English (Figure 9 top panel; $\beta = -0.30$, $SE = 0.12$, $z = -2.55$, $p = .011$). In the English fluency model, there was also a significant interaction between fluency and language, indicating that higher English fluency predicted higher naming accuracy in English (Figure 9 bottom panel; $\beta = 0.24$, $SE = 0.11$, $z = 2.12$, $p = .034$), but not in Spanish ($\beta = -0.16$, $SE = 0.21$, $z = -0.77$, $p = .440$). No reliable interactions between fluency and AY error rates were observed in either model.

Table 14. Model estimates for separated context with Spanish fluency

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	4.37	0.26	16.84	<.001
Frequency	1.23	0.20	6.23	<.001
Language	-2.65	0.40	-6.61	<.001
AY errors	-0.15	0.14	-1.04	.299
Spanish fluency	-0.02	0.12	-0.15	.884
Language*AY errors	0.66	0.24	2.75	.006
Language*Spanish fluency	-0.57	0.20	-2.84	.005
AY errors*Spanish fluency	-0.04	0.26	-0.15	.878
Language*AY errors*Spanish fluency	-0.26	0.45	-0.58	.561

Random effects	Variance	*SD*
Intercept | item	1.849	1.360
Intercept | participant	0.132	0.364
Frequency | participant	0.001	0.024

Table 15. Model estimates for the separated context with English fluency

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	4.35	0.26	16.79	<.001
Frequency	1.23	0.20	6.25	<.001
Language	-2.60	0.39	-6.58	<.001
AY errors	-0.12	0.13	-0.95	.342
English fluency	0.02	0.15	0.15	.883
Language*AY errors	0.62	0.21	2.89	.004
Language*English fluency	0.41	0.21	2.10	.054
AY errors*English fluency	-0.03	0.16	-0.16	.869
Language*AY errors*English fluency	0.23	0.25	0.91	.361

Random effects	Variance	*SD*
Intercept | item	1.835	1.355
Intercept | participant	0.167	0.408
Frequency | participant	0.000	0.006

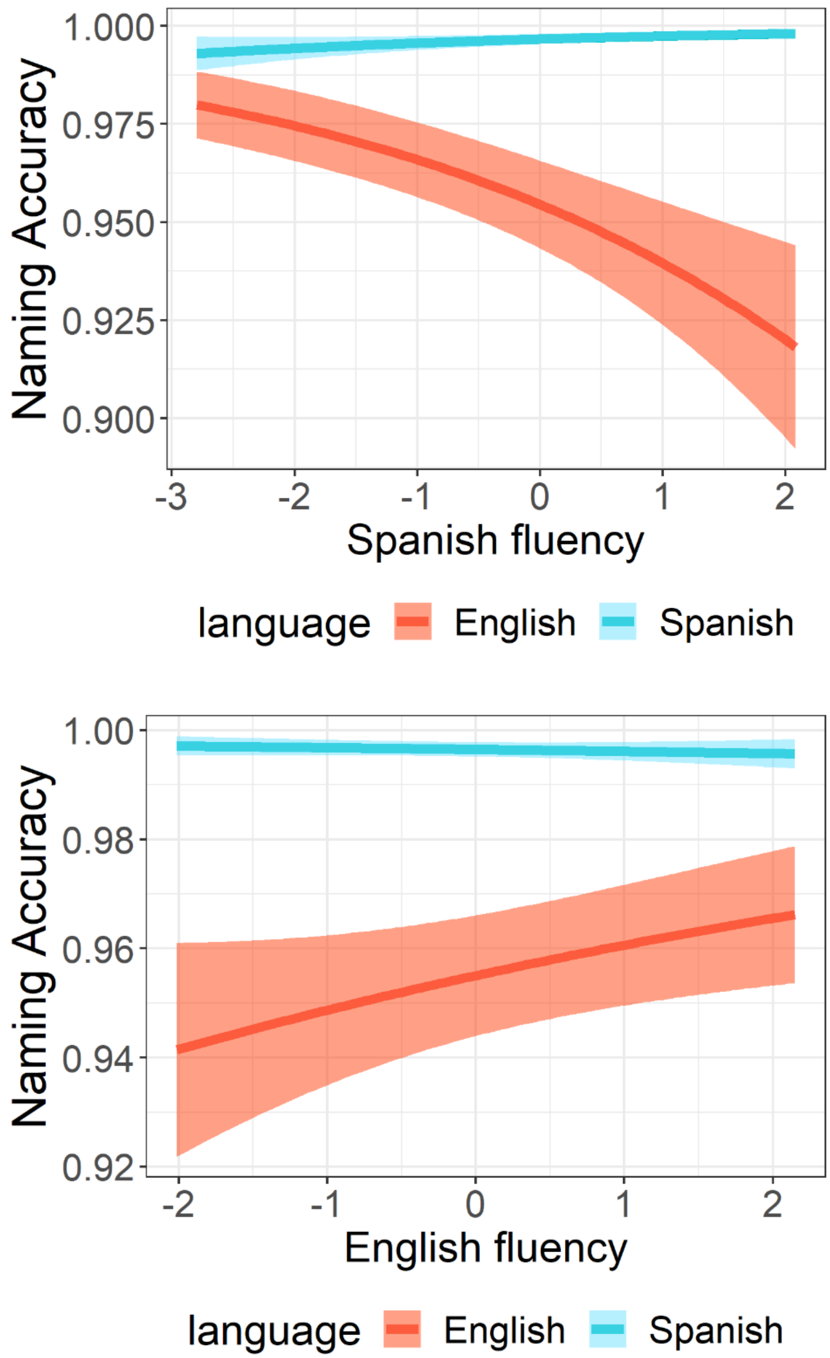

Figure 9. Estimated effect of Spanish (top) and English (bottom) fluency on naming accuracy for separated-context bilinguals in Spain. More positive values on the x-axis indicate higher fluency.

3.2. Integrated context

Tables 16 and 17 provide the model estimates for integrated-context bilinguals with Spanish and English fluency, respectively. After significantly accounting for word frequency variance, the effect of language was significant, indicating that naming accuracy was higher in Spanish than in English. Although no reliable effects involving AY error rates were found, there were several significant effects involving fluency in Spanish and in English.

In the Spanish fluency model, there was a significant main effect of Spanish fluency, indicating that those with higher Spanish fluency generated better naming accuracy regardless of language (Figure 10 top panel). This is unlike the separated-context effect of Spanish fluency, which predicted a negative association between Spanish fluency and English naming accuracy. However, in the English fluency model, the effect of English fluency was similar to the one observed in the separated-context analysis, as revealed by a significant interaction between English fluency and language. The follow-up analysis further confirmed that higher English fluency was associated with higher naming accuracy in English (Figure 10 bottom panel; β = 0.36, SE = 0.17, z = 2.15, p = .031) but not in Spanish (β = -0.28, SE = 0.17, z = -1.63, p = .104).

Table 16. Model estimates for the integrated context with Spanish fluency

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	5.05	0.33	15.07	<.001
Frequency	1.54	0.25	6.05	<.001
Language	-1.46	0.52	-2.83	.005
AY errors	-0.13	0.12	-1.11	.269
Spanish fluency	0.29	0.12	2.34	.019
Language*AY errors	0.29	0.16	1.82	.069
Language*Spanish fluency	-0.35	0.23	-1.55	.122
AY errors*Spanish fluency	0.16	0.10	1.59	.111
Language*AY errors*Spanish fluency	0.02	0.14	0.12	.901

Random effects	Variance	*SD*
Intercept | item	3.313	1.820
Intercept | participant	0.196	0.442
Frequency | participant	0.002	0.049

Table 17. Model estimates for the integrated context with English fluency

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	4.85	0.32	15.03	<.001
Frequency	1.47	0.25	5.81	<.001
Language	-1.30	0.46	-2.80	.005
AY errors	-0.13	0.14	-0.92	.358
English fluency	0.12	0.13	0.91	.365
Language*AY errors	0.30	0.25	1.19	.232
Language*English fluency	0.63	0.19	3.36	.001
AY errors*English fluency	0.05	0.11	0.42	.672
Language*AY errors*English fluency	-0.02	0.14	-0.17	.863

Random effects	Variance	*SD*
Intercept | item	3.274	1.810
Intercept | participant	0.258	0.508
Frequency | participant	0.006	0.077

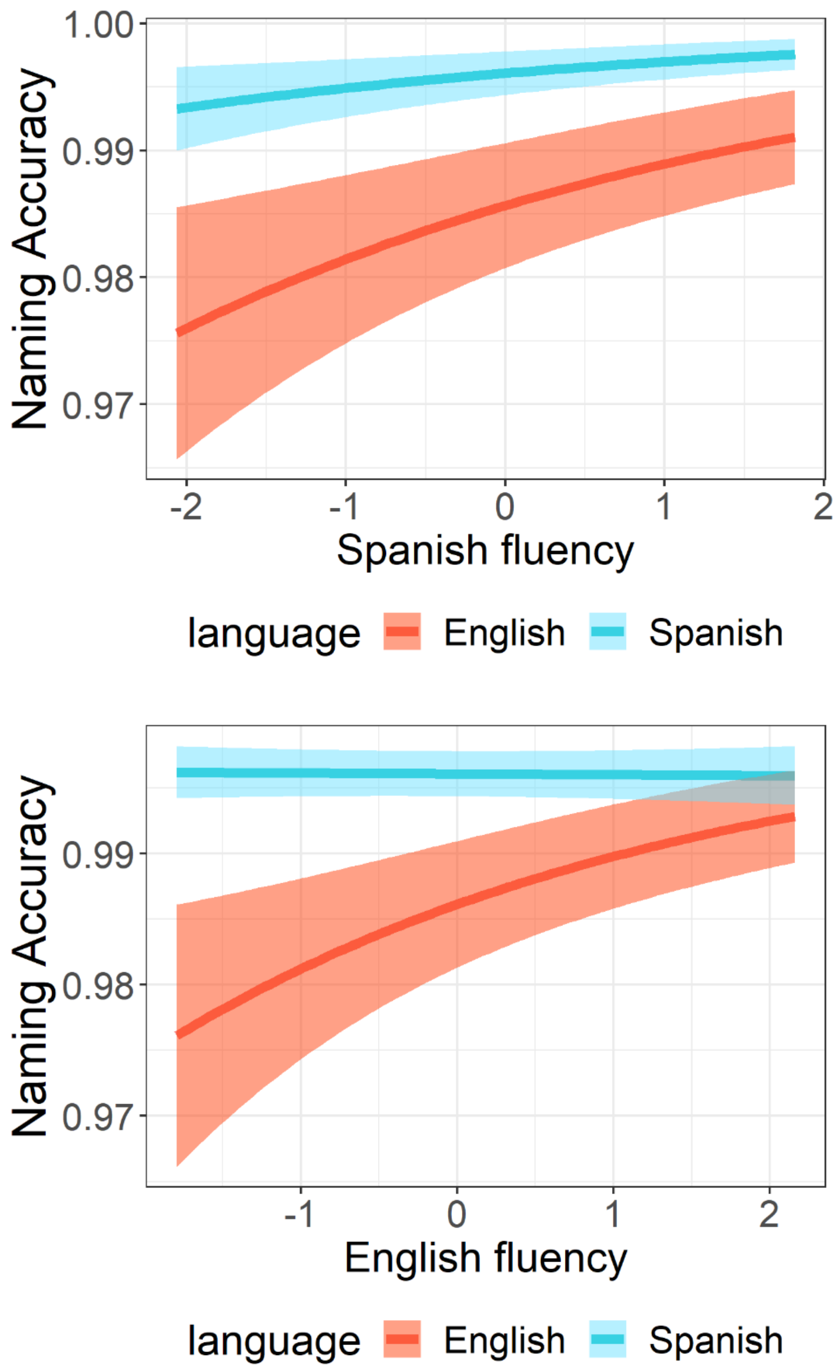

Figure 10. Estimated effect of Spanish (top) and English (bottom) fluency on naming accuracy for integrated-context bilinguals in Puerto Rico. More positive values on the x-axis indicate higher fluency.

3.3. Varied context

Tables 7 and 8 provide the model estimates for varied-context bilinguals with Spanish and English fluency, respectively. After accounting for word frequency, the effect of language was not significant, indicating similar accuracy in both languages. In the English fluency model, there was also a significant interaction between AY error rates and language, indicating that higher AY error rates (i.e., those with increased tendencies to rely on proactive control) were associated with higher naming accuracy in Spanish (β = 0.50, SE = 0.10, z = 4.77, p = <.001) but not in English (β = -0.10, SE = 0.14, z = -0.74, p = .461). This pattern replicates the continuous effect reported by Beatty-Martínez, Navarro-Torres et al. (2020). In the Spanish fluency model, the form of the interaction was different, indicating that for those with highest AY error rates, naming accuracy in Spanish was significantly higher than in English (β = -0.82, SE = 0.40, z = -2.07, p = .039), but for those with the lowest AY error rates, naming performance was similar in the two languages (β = -0.00, SE = 0.40, z = -0.01, p = .995).

In the Spanish fluency model, there was an effect of Spanish fluency on overall naming accuracy similar to the one reported in the integrated-context group: higher Spanish fluency was associated with higher naming accuracy in both languages (Figure 11). Beyond this result, however, no interactions between Spanish fluency, AY error rates, and language were found.

Unlike the other two groups, and as predicted, the English fluency model yielded a significant three-way interaction between language, English fluency, and AY error rates. A statistical breakdown of this interaction revealed a two-way interaction between language and English fluency for those with low AY error rates (i.e., those with greater reactive control tendencies; β = 0.89, SE = 0.26, z = 3.46, p = .001). Follow-up simple effects analyses indicated that for this subgroup, higher English fluency was associated with more accurate naming in English (Figure 12 left panel; β = 0.83, SE = 0.24, z = 3.51, p = .001) but not in Spanish (β = -0.07, SE = 0.15, z = -0.45, p = .650). In turn, for individuals with the highest AY error rates (i.e., those with the highest proactive control tendencies), English fluency predicted overall increases in naming accuracy regardless of the language (Figure 12 right panel; β = 0.53, SE = 0.17, z = 3.06, p = .002). Together, these results confirm the presence of an interaction consistent with the regulation hypothesis, but its manifestation emerges only in bilinguals who are L2 immersed.

Table 18. Model estimates for the varied context with Spanish fluency

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	3.83	0.23	16.81	<.001
Frequency	1.32	0.19	6.77	<.001
Language	-0.41	0.35	-1.17	.242
AY errors	-0.04	0.11	-0.41	.679
Spanish fluency	0.50	0.10	4.77	<.001
Language*AY errors	-0.41	0.18	-2.25	.024
Language*Spanish fluency	-0.07	0.18	-0.40	.688
AY errors*Spanish fluency	0.06	0.09	0.73	.466
Language*AY errors*Spanish fluency	-0.13	0.15	-0.82	.409

Random effects	Variance	*SD*
Intercept | item	0.112	0.335
Intercept | participant	0.023	0.153
Frequency | participant	0.112	0.335

Table 19. Model estimates for the varied context with English fluency

Fixed effects	Estimate	*SE*	*z*-values	*P*
(Intercept)	3.86	0.22	17.20	<.001
Frequency	1.30	0.19	6.78	<.001
Language	-0.23	0.35	-0.67	.500
AY errors	0.14	0.11	1.33	.184
English fluency	0.45	0.10	4.38	<.001
Language*AY errors	-0.49	0.16	-3.01	.003
Language*English fluency	0.41	0.17	2.45	.014
AY errors*English fluency	0.08	0.12	0.61	.539
Language*AY errors*English fluency	-0.48	0.21	-2.27	.023

Random effects	Variance	*SD*
Intercept | item	1.969	1.403
Intercept | participant	0.112	0.335
Frequency | participant	0.023	0.153

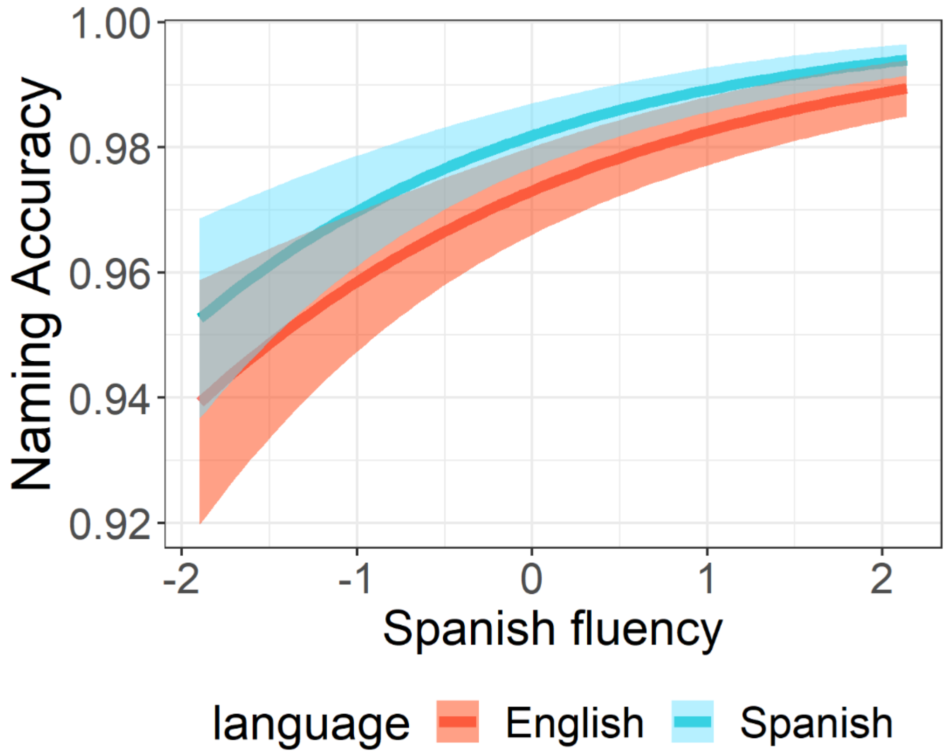

Figure 11. Estimated effect of Spanish fluency for varied-context bilinguals in the US. More positive values on the x-axis indicate higher Spanish fluency.

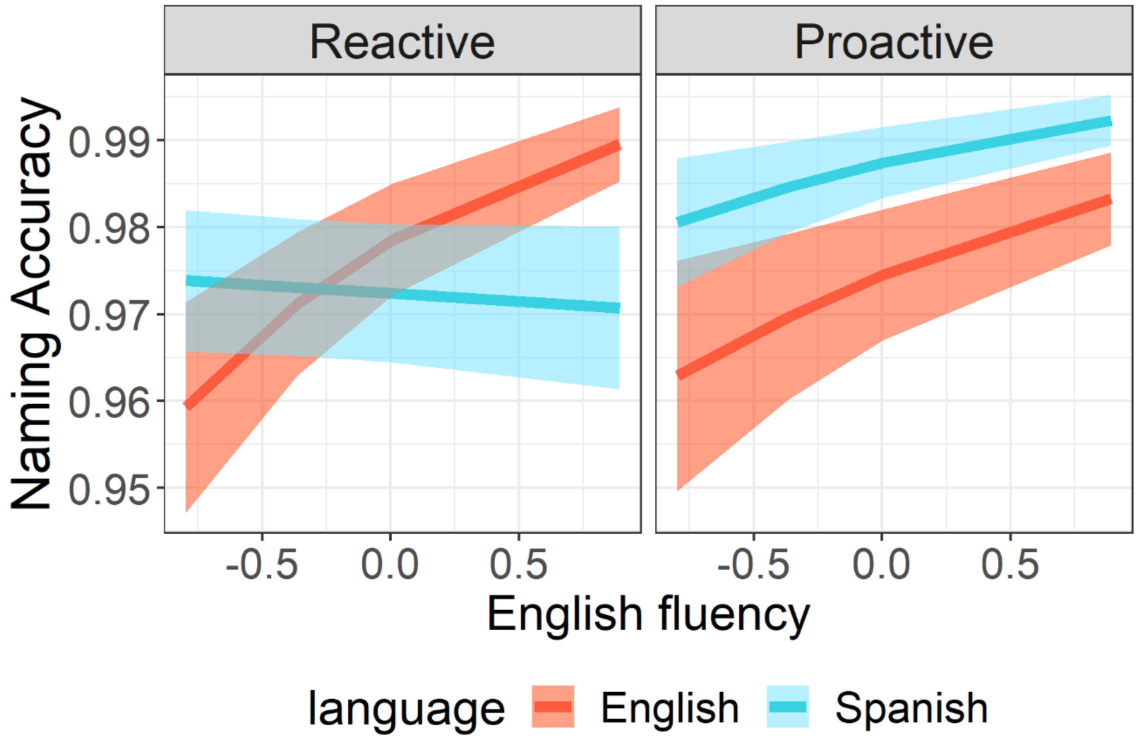

Figure 12. Estimated interaction between language, AY error rates, and English fluency for varied-context bilinguals in the US. The left figure represents individuals with the lowest AY error rates (1 SD below the mean). The right figure represents individuals with highest error rates (1 SD above the mean). More positive values on the x-axis indicate higher English fluency.

4. Discussion

The reanalysis implemented here sought to reexamine variation in language production across the three contexts of bilingualism presented in Beatty-Martínez, Navarro-Torres et al. (2020), with the goal of exploring the hypothesis that active regulation involves an interactive process between language-related and domain-general resources that mediates fluent lexical access, and that such an interaction is at least partially separable from proficiency. The results are in line with the hypothesis, particularly those observed in the varied-context group, who revealed an interaction between AX-CPT and fluency in the L2 English, their dominant language following L2 immersion. In addition, although neither of the other contexts yielded a pattern consistent with a regulatory effect, several noteworthy patterns of association between fluency and naming accuracy emerged.

4.1. Fluency, competition, and cooperation

For separated and integrated-context bilinguals, both Spanish and English fluency predicted naming accuracy, but the form of the association varied by language and context in a manner that is consistent with their patterns of language use, as well as with proficiency effects. For both groups, fluency in English (the language with less support from the environment, especially in the separated-context group), was positively associated with English naming accuracy, consistent with a proficiency effect. But when examining Spanish fluency (the predominant language of the environment in both groups), it was either negatively associated with English naming, as in the case of the separated-context group, or positively associated with naming in both languages, as in the case of the integrated-context group. Despite these effects not interacting with AX-CPT, such patterns are likely not indicative of proficiency in the strict sense. Rather, they may reflect how these bilinguals habitually use their languages, which may also be reflective of how the environment supports those practices and they shape the language network.

Such an interpretation is consistent with the control processes model (CPM) proposed by Green and Wei (2014), and later revisited in Green (2018), which posits that the dynamics of bilingual language control are directly mediated by the intention of how to to use the languages. Under the CPM, bilinguals in separated contexts engage their languages *competitively*, where activation of one language is suppressed at the expense of the other. In turn, bilinguals in integrated contexts engage their languages *cooperatively*, where co-activation is maintained all the way through speech planning so that items from both languages can be retrieved opportunistically. Critically, a given control state is hypothesized to result in a "habit of control" with repeated use (Green & Abutalebi, 2013). Therefore, for a bilingual who relies on a cooperative control state, the relevance of language membership is minimized. Conversely, the relevance of language membership is maximized for a bilingual in a competitive control state.

One consequence of cooperative states may be that information from lexical items in one language can be readily mapped on to lexical equivalents in the other language due to greater cross-language overlap and integration. The finding that Spanish fluency in the integrated-context (as well as in the varied-context) group predicted greater naming accuracy gains across languages supports this idea. In turn, under a competitive state, information flow from one language to the other is restricted. In this scenario, lexical retrieval may be limited to the functional use of each language. In the case of the separated context, where the support for English is minimized, fluency gains in the Spanish language network may reduce accessibility in the less active L2, regardless of the speaker's proficiency level more generally.

An issue that remains to be determined is whether these effects exclusively reflect control states, as proposed by the CPM, or whether they also reflect the regulatory mechanisms that characterize bilinguals in separated and integrated contexts (see Beatty-Martínez & Titone, 2021 for discussion on this issue). It is possible that regulation entails not only activation-strength management, but also integration and differentiation processes between the two languages. In research examining memory and learning more generally, such models have been proposed to account for the various states of consolidation that may be necessary to integrate novel information from competing memory traces, such as in the case of retrieval-induced-forgetting phenomena (Hulbert & Norman, 2015; Norman et al., 2007). However, this interpretation is speculative at best. For now, the point is made that measures such as category fluency are capable of tapping into different structural and functional aspects of the language network, including proficiency, as well as how the language network is organized following the habitual patterns of language use and the environmental demands imposed on those usage patterns.

4.2. Fluency, regulation, and L2 immersion

Despite the intriguing fluency results for the separated and integrated-context groups, neither yielded interactions involving fluency and cognitive control ability. This may be in part because neither context offers a challenging-enough environment that requires active regulation, but may also be because both separated and integrated contexts generally offer cues that reliably signal when and how to use which language in a manner that may not be available to bilinguals immersed in an L2 environment. This idea is supported by research on codeswitching indicating that habitual codeswitchers tend to switch their languages following distributional regularities from the community (Beatty-Martínez & Dussias, 2017; Beatty-Martínez et al., 2020; Valdés Kroff, 2016), which in turn can be used as cues when comprehending codeswitched speech (Fricke, Kroll, & Dussias, 2016; Tomic & Valdés Kroff, 2021; Valdés Kroff et al., 2017) or when parsing grammatical structures (Beatty-Martínez et al., 2017; Guzzardo Tamargo, Valdés Kroff, & Dussias, 2016). Thus, the absence of an interactive effect in the integrated group is perhaps relatively straightforward: since these individuals maximize opportunistic and cooperative use of the languages, and because the environment allows them to do so, they are able to effectively minimize reliance on language regulation.

In the case of the separated context, regulation of the dominant L1 may be engaged under more restrictive circumstances, such as when speakers are required to use the L2 in a reactive fashion. Consider a case where a young Spanish-English bilingual from Granada engages in conversation with other locals who prefer Spanish, but a foreign-exchange student with limited Spanish proficiency joins the conversation, acting as a salient cue to engage English. The issue, however, is that these scenarios may not be as a recurrent and diverse as in an L2 immersed context. If so, it may be possible to induce regulatory effects in bilingual phenotypes from separated contexts, with the prediction that reactive control reliance would result in more fluent speech (a prediction that is examined in Chapter 6 with Spanish-English bilingual heritage speakers). Unlike these two contexts, varied contexts such as the one examined here should more readily yield interactions that reflect active regulation of the dominant language, together with a proactive style of cognitive control engagement that can facilitate conversational exchanges in the absence of reliable cues regarding language choice.

There are several studies that support the latter prediction to different degrees. Perhaps the most immediate source of converging evidence comes from Zirnstein et al. (2018), who found evidence for active engagement of regulation in a group of Mandarin-English bilinguals from China who were L2 immersed at the time of testing. Similar to the results reported here, dominant-language fluency (Mandarin), together with cognitive control ability (also measured using the AX-CPT), interactively predicted their ability to engage prediction mechanisms in the L2 when reading sentences with highly unexpected words following the semantic context (e.g., sentences such as *after their meal, they forgot to leave* <u>*tip*</u>, which has an expected lexical continuation, whereas *after their meal, they forgot to leave a* <u>*ten*</u>, which has a less expected continuation). Such dynamics would allow them to efficiently generate predictions in real time while also being able to recover from those predictions when they were not met as the sentence unfolded.

Another source of evidence involving language processing comes from an eye-tracking study by Navarro-Torres et al. (2019), where they examined recovery from syntactic ambiguity, and how such recovery was impacted by cognitive control engagement, in a group of L2 immersed bilinguals at the University if Edinburgh. The authors found that the L2 immersed bilinguals, relative to a monolingual group, engaged cognitive control more proactively to disengage incorrect interpretations as they parsed the ambiguous sentence. Although this study did not directly assess language regulation per se, it is consistent with the notion that L2 immersion may induce a shift to a proactive control style by virtue of having to engage regulation.

In addition, there is recent research focusing on cognitive control that provides support for proactive control reliance in contexts of L2 immersion. Zhang et al. (in press) examined shifts in cognitive control tendencies (using the AX-CPT) in two groups of Mandarin-English bilinguals, one immersed in the L1 (China) and the other immersed in the L2 (US). Using a language switching training paradigm, the authors observed that the L1-immersed group shifted from a reactive to a more proactive control strategy following training. However, unlike the L1-immersed group, training in the L2-immersed group did not impact their cognitive control tendencies, since at baseline they already displayed greater reliance on proactive control.

Finally, like the study by Zhang et al., data from neuroscience also provides evidence suggesting that proactive control reliance emerges in bilingual phenotypes who live in L2-immersion-like environments. For instance, Gullifer et al. (2018) observed in a diverse bilingual group from Montreal that greater self-reported diversity in language use (measured as the relative entropy, or uncertainty, of using one language most of the time vs. using both languages equally) across social contexts was associated with greater functional connectivity with regions associated with proactive control engagement (i.e., goal maintenance and conflict monitoring). In a concurrent analysis, they also found that the functional connectivity of those same brain networks was also predicted by cognitive control tendencies (as measured by the AX-CPT), with greater reliance on proactive control predicting increased functional connectivity between those regions.

5. Conclusions

The results of the data reported here suggest that bilingual language regulation is part of a complex and interactive set of processes that mediate how efficiently bilinguals plan speech. The manifestation of such a mechanism, however, is dependent on the demands imposed by the language environment in which the speaker is immersed. Together with other recent converging studies, the available evidence to date suggests that varied contexts (here defined as those environments where there is a high degree of variation in how and when the languages are used across social contexts) may come to tune domain-general cognitive resources by actively engaging in language regulation, which in turn may impact bilinguals' ability to function in each language. In the following chapter, I further examine the regulation hypothesis, and its relation to proficiency, in another bilingual phenotype; namely, heritage speakers who, like the varied-context speakers examined here, are also L2-dominant but under different circumstances.

CHAPTER 6: CATCHING REGULATION ON THE FLY

1. Introduction

As alluded to throughout the dissertation, there is great interest in understanding how bilingual experience creates long-term adaptive changes in the processes associated with cognitive and brain functioning more generally (Baum & Titone, 2014; Bialystok, 2017). Part of this interest stems from the finding that bilingual experience seems to impact how cognitive and brain pathology is manifested later in life (Avelado et al., 2021; Bialystok, 2021a). One possibility is that these consequences result from the need to actively regulate the L1 over the lifespan, an ability that seems to decline with increasing aging (Mendez, 2019; Mendez, Chavez, & Akhlaghipour, 2019).

Although the converging evidence supporting this idea is still in its infancy, the findings presented in Chapter 5 suggest that this may be the case, particularly in bilinguals who are immersed in highly dynamic environments where the decision of when and how to use the languages is not straightforward. One challenge, however, is that most of the evidence to date is correlational, where the effects observed in a given task are assumed to reflect long-term adaptive changes that gradually shape cognitive and brain functioning (an issue that is also applicable to the results presented in Chapter 5).

There are a few recent studies that attempt to circumvent this issue by asking how cognitive resources are recruited within the same bilingual individual under different conditions that impose different demands on those resources. For instance, Wu and Thierry (2013; see also sections 2.3 and 1.2. in Chapters 3 and 5, respectively) observed in Welsh-English bilinguals more efficient conflict resolution (measured electrophysiologically) in a Flanker task when performance took place in the presence of pseudo-randomly interleaved words from both languages (resembling a dual-language context) relative to a context where words from one language only appeared (resembling a single-language context). Together with other recent studies using variations of this paradigm (Bosma & Pablos, 2020; Timmer et al., 2021; Yiao et al., 2020a, 2020b), these findings are consistent with the adaptive control hypothesis (Green & Abutalebi, 2013) in that there is greater need to engage cognitive resources when the two languages are present in a competitive manner.

In a similar vein, other studies ask the question of how cognitive resources are dynamically engaged during language processing. For instance, there is evidence that, relative to monolinguals, bilinguals experience conflict-resolution gains in visuo-perceptual tasks immediately after experiencing competition induced by the presence of within-language phonological competitors (Blumenfeld & Marian, 2013). Similarly, individuals recover more efficiently from syntactic ambiguity when such ambiguity is immediately preceded by non-linguistic (e.g., Stroop and Flanker-related) conflict (Hsu & Novick, 2016; Hsu et al., 2020; Thothathiri et al., 2018), but bilinguals immersed in an L2 environment have been shown to initiate such recovery more quickly as the ambiguous sentence unfolds (Navarro-Torres et al., 2019). Such findings offer evidence that help establish causality between the demands imposed by language and cognitive control engagement.

1.2. The present study

A question thar remains open, however, is the extent to which language regulation, as defined in Chapter 4, and as empirically assessed in Chapter 5, can also be examined in a causal manner that allows us to dissociate it from other related domains, such as proficiency. In the present study, I explore this issue by investigating patterns of association between language production, cognitive control, and fluency, in a group of Spanish-English bilinguals who acquired Spanish as the home language and, by virtue of being immersed in the societal language (English), have become L2 dominant.

A notable feature of the current study is that, unlike the study in Chapter 5, where evidence supporting the language regulation hypothesis relied on comparing patterns of association across bilingual groups, it implements a within-subjects design where individuals perform the category fluency task in the dominant L2 (English) twice under different regulatory conditions. That is, participants first complete an initial block of verbal fluency in English, followed by a second block where participants are asked to perform the task in Spanish, and a final block where participants are asked to perform the task in English again but with new semantic categories that differed from the baseline English block (see Figure 13 for an illustration). This design stems from evidence showing that when dominant-language performance is preceded by performance in the less dominant language, speech is disrupted (Van Assche et al., 2013) and more distributed brain activity involving cognitive control regions is observed (Guo et al., 2012; Misra et al., 2012; Rossi et al., 2021).

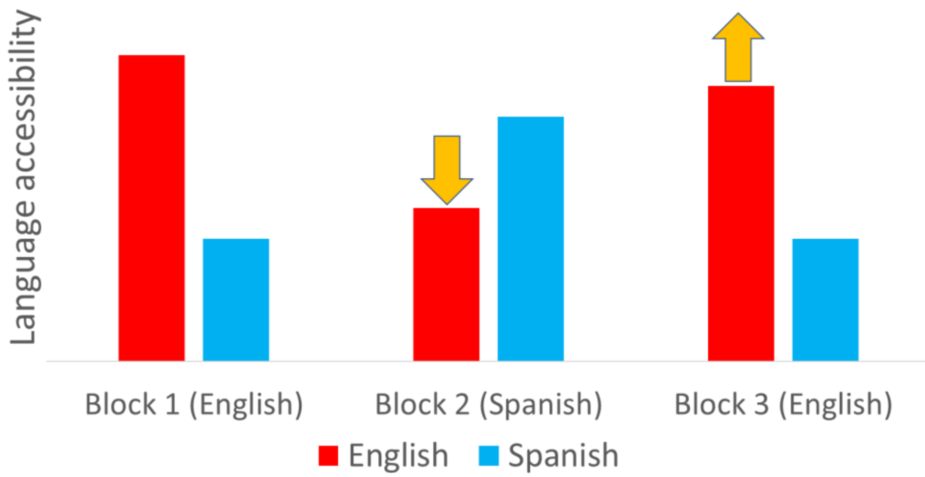

Figure 13. Hypothesized regulatory effect in the verbal fluency task as implemented here. Performance in Block 1 is hypothesized to assess lexical access as it relates to general proficiency in English (e.g., vocabulary size), since no regulatory demands are imposed in this context. In Block 2, where participants perform the task in Spanish, down-regulation of English becomes important to enable fluent speech in the less dominant language. In the third block, where participants are asked to perform in English again, participants need to actively up-regulate English to reestablish baseline lexical accessibility in the language.

As such, by examining performance in Block 1 vs. Block 3, we can ask whether the context in which verbal fluency is performed differentially yields evidence for a language regulatory mechanism. The predictions are relatively straightforward: since Block 1 performance does not impose the need for regulation, any effects that account for significant variance in picture naming performance should be in the form of proficiency (i.e., independently predicting naming performance in English). In Block 2, participants need to down-regulate English to enable speech in Spanish, but since Spanish is the less proficient language, it is possible that Spanish fluency also yields predictive effects in the form of proficiency. In the case of Block 3, however, fluency in English and AX-CPT may interactively predict naming performance, since, as described in Figure 9, recovery from the inhibitory effects in Block 2 are critical to up-regulate English.

2. Method

2.1. Participants

Participants included a total of 34 Spanish-English bilinguals (29 female) raised in the the greater Southern California area in the US. Participants were recruited as part of an online study using the Gorilla Experiment Builder platform (www.gorilla.sc; Anwyl-Irvine et al., 2018) and received either course credit via the University of California, Irvine subject pool or $10/hour. All participants gave informed consent, and the procedures had the approval of the Institutional Review Board at the University of California, Irvine.

Table 20 provides descriptive statistics and participant characteristics. Participants reported acquiring Spanish before English and having greater exposure to Spanish relative to English before age seven. Despite reporting similar levels of exposure and proficiency in the two languages, participants reported using English more frequently than Spanish, produced more category exemplars in English than in Spanish, and displayed higher picture naming accuracy in English than in Spanish. These patterns are consistent with a general characterization of heritage speakers, who, as previously described, typically acquire a home language (Spanish) first but eventually become linguistically dominant in the societal language (English).

Finally, in the AX-CPT, participants committed more errors in the AY condition relative to the BY control condition, consistent with the general pattern of performance in young adults, although they also displayed high error rates in the BX condition comparable to the AY condition. This pattern of error rates (high AY and BX error rates) resembles the one found in separated-context bilinguals from Chapter 5 (see Table 7 and Section 3.2. in Part A of Chapter 5), and may indicate that on average, some participants had difficulty engaging proactive and reactive control both in predictive (AY condition) and non-predictive (BX condition) contexts.

Table 20. Participant characteristics for heritage speakers.

Measure	M	SD
Age, years	21.7	3.2
Active AoA (years)		
Spanish	1.5	0.9
English	5.8	4.8
Current exposure (%)		
Spanish	49.1	21.9
English	49.3	21.0
Exposure before age seven (%)		
Spanish	67.3	8.2
English	31.5	7.5
Current use (1-5 scale)		
Spanish	1.6	0.6
English	3.4	0.2
Proficiency (1-5 scale)		
Spanish	3.9	1.0
English	4.1	0.3
Verbal fluency		
English (block 1)	47.3	10.1
Spanish (block 2)	30.6	10.5
English (block 3)	45.1	10.9
Picture naming accuracy (%)		
Spanish	74.3	15.8
English	95.4	4.1
AX-CPT error rates (%)		
AX	14.6	13.2

AY	29.7	24.8
BX	30.3	35.7
BY	14.7	20.3

Note: Means and standard deviations for participants' language history characteristics. Active AoA = age at which participants began actively using the language. Proficiency and use ratings were made on a 5-point scale ranging from 0 (not proficient) to 5 (highly proficient). Fluency scores were calculated as the average number of exemplars produced across categories.

2.2. Materials

Category verbal fluency task. In this version of the task, participants were asked to generate as many exemplars as possible that belong to a semantic category within a 30-second time limit. The task included 12 categories that were counterbalanced and evenly distributed across three language blocks (i.e., English, Spanish, and English; see Figure 8). The categories were *animals*, *clothing*, *musical instruments*, *vegetables*, *body parts*, *colors*, *fruits*, *furniture*. *professions, sports, school supplies,* and *family members.* Participants were asked to avoid producing repetitions and names of people or places. Responses were recorded via the participant's recording system from their computer in a quiet space.

Picture naming task. The materials used for the picture naming task were identical to those used in Chapter 5.

AX-CPT task. The materials used for the AX-CPT were identical to those used in Chapter 5.

2.3. Procedure

All tasks were completed online using the Gorilla Experiment Builder platform. Participants were asked to ensure to conduct the study in a quiet area with the least number of possible interruptions. When the task involved producing words, participants were asked to speak loudly and clearly into their computers' recording device. At the beginning of each task, participants were carefully briefed on the experimental procedure, and they completed a practice run for each task to ensure that they understood the instructions. Participants first performed a baseline block of verbal fluency in English, followed by a Spanish block in the same task. Participants then proceeded to perform the Spanish block of the picture naming task, followed by blocks of English verbal fluency and English picture naming, respectively. This order was chosen following the theoretical motivations outlined in the Present Study section (i.e., to include a baseline block in English and a subsequent block preceded by Spanish performance). For each block, written instructions indicating the language to be used appeared on the screen. After completing the language tasks, participants performed the AX-CPT, followed by a language history questionnaire via Qualtrics.

2.4. Analysis

For picture naming, although both response times (RTs) and accuracy were collected, here only accuracy is reported. A response was considered accurate if it matched the intended target name. Where appropriate, alternative dialectal variations were also considered accurate. Error rates and average RTs were also recorded for the AX-CPT, but only error rates for the AY condition are included for the purposes of the study. For category fluency, performance was assessed by calculating the average number of exemplars produced across categories in Spanish and in English.

Statistical analyses were performed on picture naming accuracy using logistic mixed-effects models in the lme4 software package (version 1.1-18-1; Bates, Mächler, Bolker, & Walker, 2015) in the R programming environment (version 3.5.1; R Development Core Team, 2014). The analysis included a contrast coded fixed effect of language block (Spanish = -0.5, English = 0.5), log-transformed word frequency values (measured as a continuous factor), AY error rates from the AX-CPT, and verbal fluency scores. Three separate models were created that included fluency scores from either English (block 1), Spanish (block 2), or English (block 3) as fixed effects. For each model, all factors were allowed to interact with one another except for word frequency, which was included as a covariate to control for frequency effects. The random effects structure contained random intercepts for subjects and items and by-participant random slopes for frequency. Continuous fixed-effects were z-scored to make the intercept in the models reflect average performance.

3. Results

3.1. Block 1 (English)

Table 21 provides the estimates for the model with English fluency from block 1. The significant word frequency effect indicated that participants had higher naming accuracy overall with increasing word frequency target words. After controlling for frequency, the effect of language was significant, indicating that participants had on average higher picture naming accuracy in the English block than in the Spanish block. Aside from these two effects, however, no significant effects involving AY errors and English fluency were observed. Next, I report results for the Spanish fluency model.

Table 21. Model estimates for heritage speakers with English fluency (Block 1)

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	3.41	0.25	13.82	<.001
Frequency	1.18	0.19	6.22	<.001
Language	-2.50	0.37	-6.86	<.001
AY errors	-0.05	0.17	-0.28	.777
English fluency	0.15	0.16	0.93	.352
Language*AY errors	-0.25	0.20	-1.27	.203
Language*English fluency	0.00	0.18	-0.01	.990
AY errors*English fluency	-0.19	0.14	-1.35	.177
Language*AY errors*English fluency	0.15	0.15	1.02	.310

Random effects	Variance	*SD*
Intercept | item	2.015	1.419
Intercept | participant	0.281	0.531
Frequency | participant	0.062	0.249

3.2. Block 2 (Spanish)

Table 22 provides the estimates for the model with Spanish fluency from Block 2. After accounting for significant effects of frequency and language, the analysis yielded a two-way interaction between language block and Spanish fluency. As Figure 14 shows, higher Spanish fluency was associated with higher naming accuracy in Spanish (β = 0.33, *SE* = 0.15, z = 2.21, p = .027) but not in English (β = 0.01, *SE* = 0.19, z = 0.08, p = .938). No significant effects involving AY error rates were observed. We now proceed to the final model including English fluency performance from Block 3.

Table 22. Model estimates for heritage speakers with Spanish fluency (Block 2)

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	3.33	0.25	13.50	<.001
Frequency	1.21	0.19	6.20	<.001
Language	-2.46	0.36	-6.82	<.001
AY errors	0.00	0.14	0.02	.983
Spanish fluency	0.18	0.15	1.19	.235
Language*AY errors	-0.29	0.17	-1.73	.083
Language*Spanish fluency	0.32	0.16	1.96	.050
AY errors*Spanish fluency	0.09	0.15	0.60	.546
Language*AY errors*Spanish fluency	-0.46	0.29	-1.60	.110

Random effects	Variance	*SD*
Intercept | item	2.109	1.452
Intercept | participant	0.329	0.573
Frequency | participant	0.059	0.242

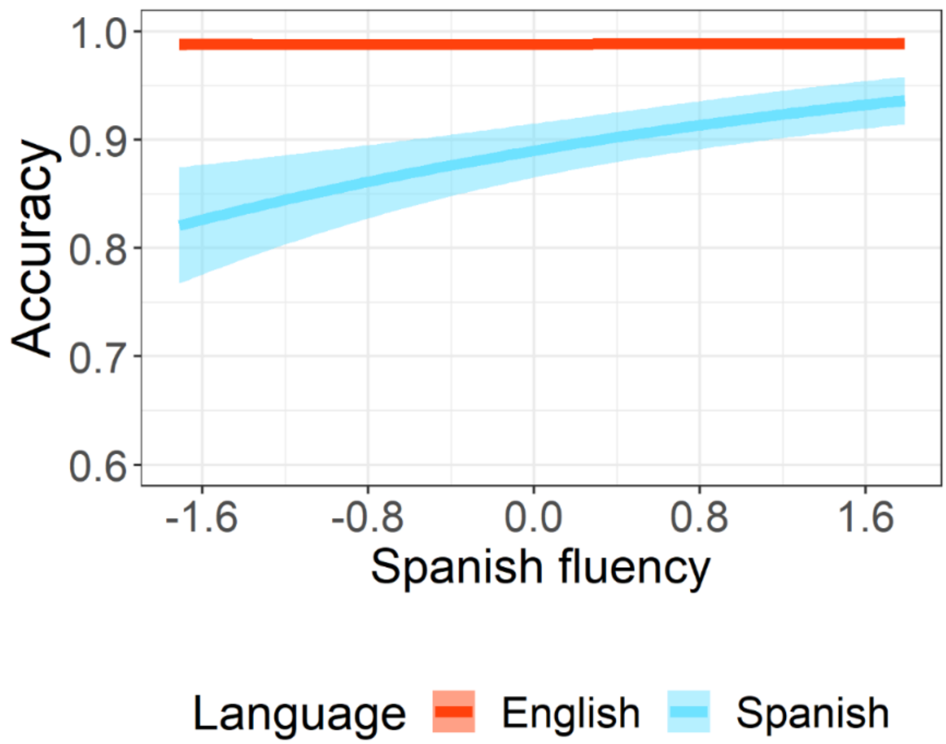

Figure 14. Estimated effect of Spanish fluency (Block 2) on naming accuracy in heritage speakers. More positive values on the x-axis indicate higher Spanish fluency.

3.3. Block 3 (English)

Table 23 provides the estimates for the model with English fluency from Block 3. After accounting for significant effects of frequency and language, the analysis yielded two-way interactions between language and AY error rates, as well as language and English fluency. Importantly, the analysis additionally yielded a three-way interaction between language, English fluency, and AY error rates. Follow-up analyses further revealed significant interactions between language and AY error rates for individuals with both low (β = 0.50, SE = 0.18, z = 1.98, p = .045) and high (β = -0.49, SE = 0.19, z = -2.52, p = .012)

English fluency, though the direction of these interactive effects were in opposite directions. As Figure 15 shows, for low English fluency bilinguals (i.e., those with low regulatory abilities), higher AY error rates were associated with higher Spanish naming accuracy (β = 0.70, SE = 0.30, z = 2.35, p = .019), but for individuals with high English fluency (i.e., those with high regulatory abilities), higher AY error rates were associated with lower Spanish naming accuracy (β = -0.56, SE = 0.18, z = -3.15, p = .002). Together, these interactive results are consistent with the basic premise of the language regulation hypothesis and suggest that efficient regulation of the dominant language can facilitate lexical accessibility in the less dominant heritage language, although this effect only emerges for individuals who adopt greater reactive control tendencies.

Table 23. Model estimates for heritage speakers with English fluency (block 3)

Fixed effects	Estimate	*SE*	*z*-values	*p*
(Intercept)	3.33	0.23	14.31	<.001
Frequency	1.20	0.19	6.32	<.001
Language	-2.42	0.35	-6.85	<.001
AY errors	0.07	0.14	0.45	.649
English fluency	-0.23	0.18	-1.29	.196
Language*AY errors	0.01	0.18	0.03	.972
Language*English fluency	0.52	0.21	2.48	.013
AY errors*English fluency	-0.41	0.20	-2.05	.040
Language*AY errors*English fluency	-0.49	0.24	-2.01	.045

Random effects	Variance	*SD*
Intercept | item	1.849	1.360
Intercept | participant	0.132	0.364
Frequency | participant	0.001	0.024

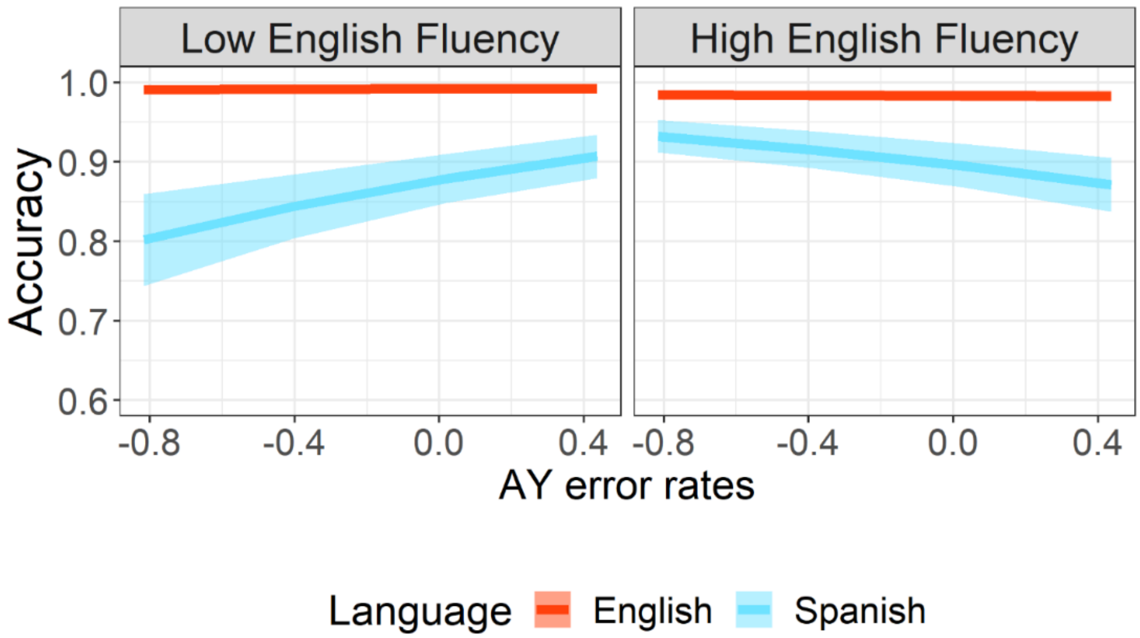

Figure 15. Estimated effect for the interaction between language, AY error rates, and English fluency (Block 3) on naming accuracy in heritage speakers. Low English fluency (left) depicts individuals with scores 1 SD below the mean, whereas high English fluency (right) depicts individuals with scores 1 SD above the mean. More positive values on the x-axis indicate higher AY error rates.

4. Discussion

This study sought to identify a potential causal role for language regulation on lexical access in a group Spanish-English bilingual heritage speakers. By manipulating the demands in which participants performed the fluency task in English (i.e., whether they generated exemplars in English under conditions of low or high regulatory engagement), the results showed that only dominant-language fluency in a high-regulatory context (in Block 3), together with the AX-CPT, interactively predicted naming accuracy in Spanish, the heritage language. Interestingly, the interactive effect revealed two distinct profiles of performance. In the subgroup with low regulatory ability, Spanish picture naming accuracy was highest for those adopting a proactive control-oriented strategy on the AY condition. In turn, in the group with high regulatory ability, Spanish naming accuracy was highest in those who adopted reactive control tendencies. These results are difficult to explain purely in terms of proficiency ability, since no effects of Block 1 English fluency on naming accuracy were observed, although an effect of Spanish fluency from Block 2 on Spanish naming accuracy was found, consistent with a proficiency effect.

One possible interpretation for these results is that each subgroup reflects different types of heritage-speaker phenotypes with distinct features regarding language experience and language use. Low regulatory bilinguals may navigate environments that more closely resemble integrated contexts, where active regulation is reduced, though still possible. Consider a bilingual with friends and family members who typically engage the two languages, be it via habitual codeswitching patterns, or via more innovative forms of integration, such as congruent lexicalization (i.e., when the two languages share a structure that can be lexicalized by features of either language; e.g., consider the Spanish *printear*, which emerges from the integration between the verb *print*, while applying the morphemic rules of Spanish used with verbs in Spanish*;* Green, 2018; Muysken, 1997). On the other hand, for heritage speakers with high regulatory skill, their language use may be similar to the group in the separated context from Spain. That is, they may habitually engage in contexts where the languages are kept separate and used competitively (consider a proficient heritage speaker where Spanish is used in limited contexts, such as home with some family members who predominantly speak Spanish).

There is some evidence supporting these interpretations. For instance, Hofweber, Marinis, and Treffers-Dallers (2016) found that German-English bilinguals from South Africa, who live in a language contact situation (similar to the colonial situation between Puerto Rico and the US; see Section 3.1 in Chapter 3) and who habitually engage in congruent lexicalization, tended to display greater cognitive control efficiency under conditions in which monitoring demands are high (a process that is hypothesized to engage proactive control). Further, a subsequent study by Hofweber, Marinis, and Treffers-Dallers (2020) found that bilinguals who typically engaged the two languages in an alternational manner (when codeswitching involves long stretches of each language) displayed more efficient cognitive control under conditions of low monitoring, where reactive control strategies become more relevant.

If these interpretations are correct, then a key prediction can be readily generated. It may be possible to induce regulatory effects in bilinguals from both integrated and separated contexts, but the form of the interaction reflecting regulation would be different for each. That is, in contexts requiring active language regulation, integrated contexts should favor proactive control reliance to facilitate lexical access, whereas separated contexts may favor reactive control reliance to achieve the same outcome. Further, it may be possible to identify potential sources of these effects using a combination of tools from network science (Tiv et al., 2020) to examine how these bilinguals use their languages for given topic across distinct social networks, Information Theory (Gullifer & Titone, 2020b;

Feldman et al., 2021) to determine which of those social network structures yield greater diversity in language use, as well as using self-reported data to determine which forms of codeswitching behaviors are most prevalent in each phenotype.

For instance, high regulatory heritage speakers may be more likely to prefer using Spanish in home and family contexts while predominantly using English in work and school contexts. Switching languages may be possible in less formal contexts, such as social contexts and (under restricted circumstances) family context, but with the tendency to rely on alternations rather than congruent lexicalization. In turn, low regulatory heritage speakers may display a more distributed use of the two languages across each of these contexts, with some contexts (e.g., home, social, and work contexts) exhibiting increased congruent lexicalization. Although these interpretations are speculative at best, they suggest that an adequate characterization of bilingualism in heritage speakers will require an approach that accounts for a host of factors regarding habits of language use, as well as how the social network structure of different heritage-speaking phenotypes supports those habits of language use.

5. Conclusions

The results reported in this chapter further support the language regulation hypothesis by suggesting that language regulation can be induced experimentally to yield an interactive effect between regulation of the dominant language and cognitive control engagement. These effects may be observed even in bilingual phenotypes who may not habitually engage regulation, suggesting that the mechanisms involving language regulation are a fundamental feature of the bilingual brain, similar to the cross-language interactive effects that have been previously established in the literature. It will be critical for future research to more precisely identify the contexts in which active regulation is more likely to emerge across different bilingual phenotypes, and the function that regulation plays in each of those contexts.

SUMMARY AND CONCLUSIONS

The overarching goal of this dissertation was to examine the unique ways in which proficient bilinguals come to engage the cognitive mechanisms that enable fluid and efficient speech. This aim stemmed from four major discoveries reviewed in Chapter 1 that have characterized research on bilingualism over the last two decades; namely, that the bilinguals' two languages are inherently interconnected, that such interconnectedness creates cross-language bidirectional influences, that negotiating the competition that arises from such interactions has consequences for cognitive and brain functioning more generally, and that those consequences are shaped by the contexts in which bilinguals come to use their languages.

In Chapter 3, my colleagues and I proposed in a recently published paper that a more nuanced understanding of the nature of these discoveries, and how they are interrelated, will require the implementation of ideas and research practices grounded in a context of scientific discovery, where the process of exploration and falsification leads to new insights and allows for new questions to emerge. Prescriptions and replication efforts need to be in service of this process to enable incremental progress. Critical to ensure such progress is the application of variety in ideas, practices, and tools. In the context of research on bilingualism, the application of variety entails an understanding of the variation that arises across different bilingual phenotypes, together with research questions and tools that enable the identification of systematic interactions in the data.

The culmination of these ideas, together with the four established discoveries, led to the proposal in Chapter 4 that proficient bilingualism is characterized by the engagement of a language regulatory mechanism that interactively coordinates language-related and domain-general resources to maximize fluidity and processing in the languages in accordance with the demands imposed by the language context. Exploiting the presence of such a mechanism requires an approach that focuses on variation in language processing and how it draws from cognitive resources in different contexts of bilingualism.

Chapter 5 takes an initial step in assessing the regulation hypothesis by examining individual differences in the relation between lexical access, language regulation, and cognitive control ability in three groups of proficient Spanish-English bilinguals. Part A of the chapter, which includes a second recently published paper, shows that lexical access (as measured in a picture naming task) and its relation to cognitive control (as measured by the AX-CPT, a measure of proactive vs. reactive control engagement) is mediated by the interactional context in which bilinguals are immersed in: those immersed in an L1 environment who typically use their languages separately (i.e., bilinguals in Granada, Spain) achieve better naming performance when their preference is to rely on reactive control. For those immersed in an L1 environment but who typically use the languages in an integrated and opportunistic manner, naming performance was not associated with cognitive control tendencies. In turn, those who are immersed in an L2 environment where the choice of when and how to use the L1 is constrained achieve better naming performance when adopting a proactive style of control.

In Part B of Chapter 5, I follow up on the results reported in Part A by providing a reanalysis that directly assessed the extent to which language regulation (as measured by a category verbal fluency task) mediated the cognitive control effects on picture naming performance. These results further reveal that the effects of cognitive control on naming performance depend on regulation of the dominant language, particularly for L2 immersed bilinguals.

In Chapter 6, I examine individual differences in language regulation and how it mediates the effect of cognitive control on naming performance in a group of Spanish-English bilingual heritage speakers who, like the L2 immersed bilinguals from the varied context, are also L2 dominant. Using a within-subjects design, the results further confirm that the predicted interaction between dominant-language fluency and AX-CPT only emerge when verbal fluency is performed in a context that induces a high regulatory state (i.e., when preceded by performance in the less proficient Spanish). This interaction further revealed the presence of two bilingual phenotypes within the heritage speaker population with distinct patterns of association between cognitive control and naming performance in Spanish: those with low regulatory ability, who displayed better Spanish naming performance when the preferred style of cognitive control is proactive, and those with high regulatory ability, who also displayed better Spanish naming performance when the preferred control style is reactive.

What conclusions can we draw from these findings regarding discovery science in research on bilingualism? This question is particularly relevant in the context of the debate surrounding the cognitive consequences of bilingualism, which has yielded inconsistent findings regarding the group comparisons between monolinguals and bilinguals on measures of cognitive control. I comment on three aspects. First, an adequate causal account of the consequences of bilingualism for cognitive and brain functioning can likely only be achieved by identifying how the brain tunes the different cognitive processes and brain networks that support language as they relate to language processing, language learning, and language use. The language regulation hypothesis, as presented here and in other emerging research, is an initial attempt to provide a context to exploit this issue. As pointed out in Chapter 3, what is critical is not the identification of a single brain region or process, but an understanding of how the relation between distinct processes changes in response to different demands imposed by language. This can only be achieved via the exploration of different tasks that tap into forms of language processing and cognitive engagement. Such an approach can enable the identification of interactive effects that likely drive most of the simple main effects that are typically reported in the literature.

Second, research showing how different bilingual contexts of language use mediate the relation between language and cognition (a critical component of the studies reported here) highlight the need to examine the range of variation that emerges within and across populations of speakers, and not just focus on traditional group comparisons. It is not that group comparisons are not useful, but that we are in a better position to understand their significance if we identify how communities of speakers, who likely share more in common than individuals from different communities, differentially adapt to the specifics of their environment. Indeed, we may be better able to identify the presence of underlying mechanisms more generally if we identify how they are uniquely deployed across well-defined samples.

Finally, I conclude by making the point that for the foreseeable future, replication efforts in the field are likely more beneficial if they are primarily geared towards the exploration of variety in methods, questions, tools, and samples. It is the identification of converging signals across methods and populations that enables the progressive and self-correcting dynamics of the discovery process, and that have characterized science throughout history. This is because science is ultimately about the unknown (i.e., the exploration of the conceptual ground). Therefore, for direct-replication research to eventually become fruitful, replication efforts are likely to benefit more by reconceptualizing replicability in the form of a science of consilience.

REFERENCES

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., ... & Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, *22*, 2076-2086.

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*, 242-275. doi:10.1016/j.jneuroling.2006.10.003

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*, 689-698. doi:10.1017/S1366728916000225

Adamou, E. (2010). Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish. Language in Society, 39, 147–171.

Adler, R. M., Valdés Kroff, J. R., & Novick, J. (2018). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning Memory and Cognition*. *46,* 741-759. https://doi.org/10.1037/xlm0000755

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Ambridge, B., & Brandt, S. (2013). Lisa filled water into the cup: The roles of entrenchment, pre-emption and verb semantics in German speakers' L2 acquisition of English locatives. *Zeitschrift für Anglistik und Amerikanistik*, *61*(3), 245-263.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of memory and language*, *53*(1), 60-80.

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, *58*(2), 262-285.

Anderson, J. A. E., Hawrylewicz, K., & Grundy, J. G. (2020). Does bilingualism protect against dementia? A meta-analysis. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-020-01736-5

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., … Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology*, *5*, 398. https://doi.org/10.3389/fpsyg.2014.00398

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics, 5*, 1–21. doi: 10.1146/annurev-linguistics-011718-011820

Austin, G., Pongpairoj, N., & Trenkic, D. (2015). Structural Competition in Second Language Production: Towards a Constraint-Satisfaction Model. *Language Learning*, *65*(3), 689-722.

Aveledo, F., Higueras, Y., Marinis, T., Bose, A., Pliatsikas, C., Meldaña-Rivera, A., ... & Goicochea-Briceño, H. (2021). Multiple sclerosis and bilingualism: Some initial findings. *Linguistic Approaches to Bilingualism*, *11*(4), 551-577.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). The CELEX database. *Nijmegen: Center for Lexical Information, Max Planck Institute for Psycholinguistics, CD-ROM*.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 176–189. https://doi.org/10.1016/S0022-5371(70)80048-2

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, *75*, 959–963. https://doi.org/10.1002/ana.24158

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml/2012.11.001

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. https://doi.org/10.1093/scan/nsw154

Barrett, M. (2018, May 13). Penicillin's first patient. Retrieved from: https://mosaicscience.com/story/penicillin-first-patient-history-albert-alexander-AMR-DRI/

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. *Mechanisms of language acquisition*, 157-193.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. arXiv:1406.5823

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, *35*, 857–894. doi:10.1017/S0142716414000174

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*, 402–409. doi:10.1016/j.actpsy.2013.01.010

Beatty-Martínez, A. L. (2019). Revisiting Spanish grammatical gender in monolingual and bilingual speakers: Evidence from event-related potentials and eye-movements. Unpublished Ph.D. dissertation, The Pennsylvania State University, State College, PA, USA.

Beatty-Martínez, A. L., & Titone, D. A. (2021). The Quest for Signals in Noise: Leveraging Experiential Variation to Identify Bilingual Phenotypes. *Languages*, *6*(4), 168.

Beatty-Martínez, A.L., Bruni, M.R., Bajo, M.T., Dussias, P.E (2020). Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers. *Psychophysiology*, *58*(3):e13737 https://doi:10.1111/psyp.13737

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173–189. https://doi.org/10.1016/j.jml.2017.04.002

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology*, *11*, 1699. https://doi.org/10.3389/fpsyg.2020.01699

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology-Learning Memory and Cognition*, *46*, 1022–1047. https://doi.org/10.1037/xlm0000770

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language, 95,* 173-189. doi: 10.1016/j.jml.2017.04.002

Beatty-Martínez, A. L., Valdés-Kroff, J. R., & Dussias, P. E., (2018). From the field to the lab: A converging methods approach to the study of codeswitching. In M. Deuchar & D. Kašćelan (Eds.)., Interdisciplinary Perspectives on Code-Switching [Special Issue]. *Languages 3*(2), 19. doi:10.3390/languages3020019.

Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *110*, 461–468. https://doi.org/10.1016/j.jecp.2011.05.005

Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., & Craik, F. I. M. (2015). Publication bias and the validity of evidence: What's the connection? *Psychological Science, 26,* 944-946. https://doi.org/10.1177/0956797615573759

Bialystok, E. (2017). The Bilingual Adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*, 233–262. doi:10.1037/bul0000099

Bialystok, E. (2021a). Bilingualism: Pathway to Cognitive Reserve. *Trends in Cognitive Sciences*.

Bialystok, E. (2021b). Cognitive effects of bilingualism: An evolving perspective. *Bilingualism Across the Lifespan*, 9-28.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, *21*, 522–538. doi: 10.1016/j.jneuroling.2007.07.001

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*(1), 40-49.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. *NeuroReport*, *26*, 966–971. doi: 10.1097/WNR.0000000000000453

Bice, K., & Kroll, J. F. (2019). English only? Monolinguals in linguistically diverse contexts have an edge in language learning. *Brain & Language, 196,* 104644, https://doi.org/10.1016/j.bandl.2019.104644

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. *NeuroReport, 26,* 966-971. https://doi.org/ 10.1097/WNR.0000000000000453

Billig, M. (2019). More examples, less theory. Cambridge University Press.

Blanco-Elorrieta, E., Emmorey, K., & Pylkkänen, L. (2018). Language switching decomposed through MEG and evidence from bimodal bilinguals. *Proceedings of the National Academy of Sciences*, *115*, 9708. https://doi.org/10.1073/pnas.1809779115

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Sciences*, *22*, 1117-1126. https://doi.org/10.1016/j.tics.2018.10.001

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*(5), 547-567.

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science*, *363*(eaav3218). https://doi.org/10.1126/science.aav3218

Bobb, S. C., & Wodniecka, Z. (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*, 568–585. doi: 10.1080/20445911.2013.792822

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, *22*(5), 1052-1067.

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Language experience modulates bilingual language control: The effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychologica*, *193*, 160–170. https://doi.org/10.1016/j.actpsy.2018.11.004

Bosma, E., & Pablos, L. (2020). Switching direction modulates the engagement of cognitive control in bilingual reading comprehension: An ERP study. *Journal of Neurolinguistics*, *55*, 100894. https://doi.org/10.1016/j.jneuroling.2020.100894

Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652. doi:10.1037//0033-295X.108.3.624.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Science, 8*, 539-546.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624.

Bowers, K.S., Regeher, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*, 72- l10. https://doi.org/10.1016/0010-0285(90)90004-N

Brace, C. L. (1986). Egg on the face, f in the mouth, and the overbite. *American Anthropologist*, *88*, 695–697. https://doi.org/https://www.jstor.org/stable/679490

Braver, T. S. (2012). The variable nature of cognitive control: A dual-mechanisms framework. *Trends in Cognitive Sciences, 16*, 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J., Kaye, J. A.,... Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*, 746– 763. doi: 10.1037//0096-3445.130.4.746

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513–531. https://doi.org/10.1037/0003-066X.32.7.513

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of cognition*.

Brysbaert, M. (2020). Power considerations in bilingualism research: Time to step up our game. *Bilingualism: Language and Cognition*, 1-6. https://doi.org/10.1017/S1366728920000437

Bultena, S., Dijkstra, T., & van Hell, J. G. (2015). Switch cost modulations in bilingual sentence processing: Evidence from shadowing. *Language, Cognition and Neuroscience*, *30*(5), 586-605.

Calhoun, V. & Bandettini C. (2020, May 22). Defending fMRI, brain mapping, and discovery science. [Blog post]. Retrieved from: http://www.thebrainblog.org/2020/05/22/defending-fmri-brain-mapping-and-discovery-science/

Casas, M. P. (2016). Codeswitching and identity among Island Puerto Rican bilinguals. In R. E. Guzzardo Tamargo, C. M. Mazak, & M. C. Parafita Couto (Eds.), *Spanish-English codeswitching in the Caribbean and the US* (pp. 37–60). Amsterdam, The Netherlands: John Benjamins.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of phonetics*, *40*(2), 249-268.

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, *41*, 520–533. doi: 10.1016/j.wocn.2013.09.006

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. doi: 10.1016/j.brainres.2007.01.137

Clopper, C. G. (2014). Sound change in the individual: Effects of exposure on cross-dialect speech processing. *Laboratory Phonology, 5*, 69–90. https://doi.org/10.1515/lp-2014-0004

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 308–325). New York: Oxford University Press.

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

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135-149.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection?. *Journal of Memory and language*, *41*(3), 365-397.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491–511. doi:10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32,* 1057-1074. doi: 10.1037/0278-7393.32.5.1057

de Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences, 9*(3), 33 doi: 10.3390/bs9030033

de Bruin, A., Treccani, B., & Della Sala, S. (2015a). Cognitive advantage in bilingualism: an example of publication bias? *Psychological Science*, *26*, 99–107. doi:10.1177/0956797614557866

de Bruin, A., Treccani, B., & Della Sala, S. (2015b). The connection is in the data: We should consider them all. *Psychological Science*, *26*(6), 947–949. https://doi.org/10.1177/0956797615583443

Declerck, M. (2020). What about proactive language control?. *Psychonomic Bulletin & Review*, *27*(1), 24-35.

Declerck, M., Eben, C., & Grainger, J. (2019). A different perspective on domain-general language control using the flanker task. *Acta psychologica*, *198*, 102884.

Declerck, M., Koch, I., & Philipp, A. M. (2015). The minimum requirements of language control: evidence from sequential predictability effects in language switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 377.

Declerck, M., Meade, G., Midgley, K. J., Holcomb, P. J., Roelofs, A., & Emmorey, K. (2021). Language control in bimodal bilinguals: Evidence from ERPs. *Neuropsychologia*, *161*, 108019.

De Jong, R., Liang, C. C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus- response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 731-750.

Deacon, T. W. (2010). A role for relaxed selection in the evolution of the language capacity. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 9000–9006. https://doi.org/10.1073/pnas.0914624107

DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020). Understanding bilingual brain function and structure changes? U Bet! A unified bilingual experience trajectory model. https://doi.org/10.31234/osf.io/gkysn

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour*, *3*, 692–701. https://doi.org/10.1038/s41562-019-0609-3

Dijkstra, A. (2005). Bilingual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York, NY: Oxford University Press.

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

Donnelly, S., Brooks, P. J., & Homer, B. D. (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*, 1122-1147. https://doi.org/10.3758/s13423-019-01567-z

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or reality? In *Experimental Psychology 61,* 234-251. https://doi.org/10.1027/1618-3169/a000243

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *10*, 101-116. doi: 10.1017/S1366728906002847P

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 13763–13768. https://doi.org/10.1073/pnas.231499798

Emmorey, K., Borinstein, H. B., Thompson, R., & Gollan, T. H. (2008). Bimodal bilingualism. *Bilingualism, Language, and Cognition, 11*, 43–61. https://doi.org/10.1017/S1366728907003203

Emmorey, K., Li, C., Petrich, J., & Gollan, T. H. (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. https://doi.org/10.1037/xlm0000734

Edwards, B.G., Barch, D.M., & Braver, T.S. (2010). Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Frontiers in Human Neuroscience, 4,* 1–12. doi:10.3389/fnhum.2010.00032

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages following frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery and Psychiatry*, *68*, 650–652. http://dx.doi.org/10.1136/jnnp.68.5.650

Feldman, L., Srinivasan, V., Fernandes, R., & Shaikh, S. (2021). Insights into codeswitching from online communication: Effects of language preference and conditions arising from vocabulary richness. *Bilingualism: Language and Cognition,* 1-7. https://doi.org/10.1017/S1366728921000122

Ferrigno, S., Cheyette, S. J., Piantadosi, S. T., & Cantlon, J. F. (2020). Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. *Science Advances*, *6*(26), eaaz1002. https://doi.org/10.1126/sciadv.aaz1002

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem?. *Bilingualism: Language and Cognition*, *9*(2), 153-166.

Finn, A. S., & Hudson Kam, C. L. (2015). Why segmentation matters: Experience-driven segmentation errors impair “morpheme” learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1560.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226-248.

Freund, M., Gordon, B., & Nozari, N. (2016). Conflict-based regulation of control in language production. In A. Papafragou, D. Grodner, D. Mirman, & J.C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1625-1630). Austin, TX: Cognitive Science Society.

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language*, *91*, 181-201. doi:10.1016/j.jml.2016.04.003

Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language*, *89*, 110–137. doi:10.1016/j.jml.2015.10.001

Fricke, M., Zirnstein, M., Navarro-Torres, C. A., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition 22,* 200-207. https://doi.org/10.1017/S1366728918000482

García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016a). The neuroanatomy of bilingualism: How to turn a hazy view into the full picture. *Language, Cognition and Neuroscience*, *31*, 303–327. https://doi.org/10.1080/23273798.2015.1068944

García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M.

(2016b). “Hazy” or “jumbled”? Putting together the pieces of the bilingual puzzle. *Language, Cognition and Neuroscience*, *31*, 353–360. https://doi.org/10.1080/23273798.2015.1135247

Garraffa, M., Beveridge, M., & Sorace, A. (2015). Linguistic and Cognitive Skills in Sardinian–Italian Bilingual Children. *Frontiers in Psychology*, *6*, 1898. https://doi.org/10.3389/fpsyg.2015.01898

Garraffa, M., Obregon, M., & Sorace, A. (2017). Linguistic and Cognitive Effects of Bilingualism with Regional Minority Languages: A Study of Sardinian–Italian Adult Speakers. *Frontiers in Psychology*, *8*, 1907. https://doi.org/10.3389/fpsyg.2017.01907

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. New York: Cambridge University Press. doi:10.1017/CBO9780511790942

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*, 254-267. https://doi.org/10.1037/0033-295X.98.2.254

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 640-665. doi: 10.1037/a0014981

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What’s Easier: Doing What You Want, or Being Told What to Do? Cued versus Voluntary Language and Task Switching.

Journal of Experimental Psychology: General, *143*, 2167–2195. doi:10.1037/a0038006

Gollan, T. H., Montoya, R. I., Cera, C. M., & Sandoval, T. C. (2008). More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*, 787–814. doi: 10.1016/j.jml.2007.07.001

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological science*, *22*(9), 1155-1164.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency Drives Lexical Access in Reading but not in Speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General*, *140*, 186–209. doi:10.1037/a0022256

Anwyl-Irvine, A.L., Massonié J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods.* https://doi.org/10.3758/s13428-019-01237-x

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese–English bilinguals. *NeuroImage*, *56*(4), 2300-2309.

Gould, S. J., & Vrba, E. S. (1982). Exaptation-a missing term in the science of form. *Paleobiology, 8*, 4–15. http://www.jstor.com/stable/2400563

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210–223. https://doi.org/10.1016/0093-934X(86)90016-7.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-81. https://doi.org/10.1017/S1366728998000133

Green, D. W. (2018). Language control and code-switching. *Languages, 3,* 8. doi: 10.3390/languages3020008

Green, D. W. (2019). Language control and attention during conversation: An exploration. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 427–446). John Wiley & Sons. https://doi.org/10.1002/9781119387725.ch21

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515–530. doi:10.1080/20445911.2013.796377

Green, D. W., & Abutalebi, J. (2015). Language control and the neuroanatomy of bilingualism: in praise of variety. *Language, Cognition and Neuroscience*, *3798*(October), 1–5. https://doi.org/10.1080/23273798.2015.1084428

Green, D. W., Crinion, J., & Price, C. J. (2006). Convergence, degeneracy, and control. *Language Learning*, *56*, 99–125. https://doi.org/https://doi.org/10.1111/j.1467-9922.2006.00357.x

Green, D. W., & Kroll, J. F. (2019). The neurolinguistics of bilingualism: Plasticity and control. In G. I. de Zubicaray & N. O. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics*. https://doi.org/10.13153/diam.8.2006.233

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition, and Neuroscience*, *29*, 499–511. doi:10.1080/23273798.2014.882515

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*, 325-340. https://doi.org/10.1177/0267658316678286.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2. *Linguistic Approaches to Bilingualism*, *7*, 199–229. https://doi.org/10.1075/lab.15011.gru

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavioral Research Methods,* 45, 765-771.

Gullifer, J. W., & Titone, D. (2020a). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General.* Advance online publication. https://doi.org/10.1037/xge0000933

Gullifer, J. W., & Titone, D. (2020b). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23,* p. 283-294. https://doi.org/10.1017/S1366728919000026

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (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*, 123–134. doi: 10.1016/j.neuropsychologia.2018.04.037

Gullifer, J., & Titone, D. (2019). The impact of a momentary language switch on bilingual reading: Intense at the switch but merciful downstream for L2 but not L1 readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* doi: 10.1037/xlm0000695.

Gurunandan, K., Arnaez-Telleria, J., Carreiras, M., & Paz-Alonso, P. M. (2020). Converging evidence for differential specialisation and plasticity of language systems. *Journal of Neuroscience, 40*, 9715-9724. http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0851-20.2020

Guzzardo Tamargo, R. E., Loureiro-Rodríguez, V., Acar, E. F., & Vélez Avilés, J. (2018). Attitudes in progress: Puerto Rican youth's opinions on monolingual and code-switched language varieties. *Journal of Multilingual and Multicultural Development*, 1-18. doi:10.1080/01434632.2018.1515951

Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, *89*, 138–161. doi: 10.1016/j.jml.2015.12.002

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, *26*, 902-934. doi: 10.1080/01690965.2010.509946

Hardy, M. A. (1993). Regression with dummy variables. Newbury Park, CA: Sage

Hartanto, A., & Yang, H. (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. doi: 10.1016/j.cognition.2016.01.016

Hartanto, A., Toh, W. X., & Yang, H. (2018). Context counts: The different implications of weekday and weekend video gaming for academic performance in mathematics, reading, and science. *Computers & Education*, *120*, 51-63. https://doi.org/10.1016/j.compedu.2017.12.007

Hartanto, A. & Yang, H. (2020). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General, 149*, 609-633. https://doi.org/10.1037/xge0000672.

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, *20*(2), 219-234. https://doi.org/10.1017/S1366728915000164

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263-277.

Heisenberg, W. (1958). *Physics and Philosophy*. New York: Harper.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133*, 638–650. doi:10.1037/0033-2909.133.4.638

Hernandez, A. E., & Kohnert, K. J. (1999). Aging and language switching in bilinguals. *Aging, Neuropsychology, and Cognition*, *6*(2), 69-83.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in cognitive sciences*, *9*(5), 220-225.

Herring, J. R., Deuchar, M., Parafita Couto, M. C., & Moro Quintanilla, M. (2010). "I saw the madre": Evaluating predictions about code-switched determiner-noun sequences using Spanish-English and Welsh-English data. *International Journal of Bilingual Education and Bilingualism, 13*, 553–573. doi: 10.1080/13670050.2010.488286

Hervais-Adelman, A. G., Moser-Mercer, B., & Golestani, N. (2011). Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Frontiers in Psychology*, *2*(234), 1-8. doi: 10.3389/fpsyg.2011.00234

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*, 117-128. doi: 10.1017/S136672890300107X

Hockett, C. F. (1985). Distinguished lecture: F. *American Anthropologist 87*, 263–28. doi: 10.1525/aa.1985.87.2.02a00020

Hofweber, J., Marinis, T., & Treffers-Daller, J. (2016). Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism, 6*, 648-668. doi: 10.1075/lab.15052.hof

Hofweber, J., Marinis, T., Treffers-Daller, J., Sekerina, I. A., Spradlin, L., & Valian, V. (2019). Predicting executive functions in bilinguals using ecologically valid measures of code-switching behaviour. *Bilingualism, Executive Functions, and Beyond. Questions and Insights*, 181-205.

Hofweber, J., Marinis, T., & Treffers-Daller, J. (2020). How different code-switching types modulate bilinguals' executive functions: A dual control mode perspective. *Bilingualism: Language and Cognition*, *23*(4), 909-925.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script?. *Cognition*, *106*(1), 501-511.

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science*, *27*, 572-582. doi:10.1177/0956797615625223

Hsu, N. S., Jaeggi, S. M., & Novick, J. M. (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language, 166*, 63–77. https://doi.org/10.1016/j.bandl.2016.12.006

Hsu, N. S., Kuchinsky, S. E., & Novick, J. M. (2020). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience*, 1–29. https://doi.org/10.1080/23273798.2020.1836379

Hulbert, J. C., & Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cerebral Cortex*, *25*(10), 3994-4008.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, *2*(8), e124.

Ioannidis, J. P. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*(6), 645-654.

Iyer, A. S., Jones, F. K., Nodoushani, A., Kelly, M., Becker, M., Slater, D., Mills, R., Teng, E., Kamruzzaman, M., Garcia-Beltran, W. F., Astudillo, M., Yang, D., Miller, T. E., Oliver, E., Fischinger, S., Atyeo, C., Iafrate, A. J., Calderwood, S. B., Lauer, S. A., … Charles, R. C. (2020). Dynamics and significance of the antibody response to SARS-CoV-2 infection. *MedRxiv: The Preprint Server for Health Sciences*, 1–31. https://doi.org/10.1101/2020.07.18.20155374

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*, 277–288. doi: 10.1016/j.actpsy.2007.06.003

Jaccard, J., & Turrisi, R. (2003). *Quantitative Applications in the Social Sciences: Interaction effects in multiple regression*. Thousand Oaks, CA: SAGE. doi: 10.4135/9781412984522

Jared, D. (2015). Literacy and literacy development in bilinguals. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 165–182). Oxford University Press.

Jiao, L., Grundy, J. G., Liu, C., & Chen, B. (2020a). Language context modulates executive control in bilinguals: Evidence from language production. *Neuropsychologia*, *142*, 107441. https://doi.org/10.1016/j.neuropsychologia.2020.107441

Jiao, L., Liu, C., de Bruin, A., & Chen, B. (2020b). Effects of language context on executive control in unbalanced bilinguals: An ERPs study. *Psychophysiology*, *57*, 1–16. https://doi.org/10.1111/psyp.13653

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition*, *204*, 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kavé, G., & Goral, M. (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech?. *Aging, Neuropsychology, and Cognition*, *24*, 508-527. doi: 10.1080/13825585.2016.1226249.

Kerns, J. G., Cohen, J. D., Macdonald, A. M., Johnson, M. K., Stenger, A., Aizenstein, H., et al. (2005). Decreased conflict- and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. American *Journal of Psychiatry, 162*, 1833–1839. doi: 10.1176/appi.ajp.162.10.1833.

Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition*, *209*, 104575. https://doi.org/10.1016/j.cognition.2020.104575

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., ... & Sowden, W. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443-490.

Kleinman, D., & Gollan, T. H. (2016). Speaking Two Languages for the Price of One: Bypassing Language Control Mechanisms via Accessibility-Driven Switches. *Psychological Science*, *1*, 1–15. doi:10.1177/0956797616634633

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, *1*, 238. doi: 10.3389/fpsyg.2010.00238

Królikowska, M., Bierings, E., Beatty-Martínez, A. L., Navarro-Torres C. A., Dussias, P. E., & Parafita Couto, M. C. (2019, January). *Gender-Assignment Strategies within the Bilingual Determiner Phrase: Four Spanish-English Communities Examined*. Poster presented at the 3rd conference on Bilingualism in the Hispanic and Lusophone world (BHL), Leiden, NL.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language*, *33*(2), 149-174.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (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*, 119–135. doi:10.1017/S1366728906002483

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta psychologica*, *128*(3), 416-430.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*, 497-514. doi: 10.1080/20445911.2013.799170

Kroll, J. F. & Dussias, P.E. (2013). The comprehension of words and sentences in two languages. In T. Bhatia, W. Ritchie (Eds), *The Handbook of Bilingualism and Multilingualism* (2nd ed.) (pp. 216–243). Malden, MA: Wiley-Blackwell. doi:10.1002/9781118332382.ch9

Kroll, J. F., & Navarro-Torres, C. A. (2018). Bilingualism. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, *3*, 1-29.

Kroll, J. F., & Tokowicz, N. (2005). Models of Bilingual Representation and Processing: Looking Back and to the Future. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). Oxford University Press.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes-Kroff, J. R. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), *The psychology of learning and motivation*, Volume 56 (pp. 229- 262). San Diego: Academic Press. doi:10.1016/B978-0-12-394393-4.00007-8

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production*. New York, NY: Oxford University Press. doi:10.1093/oxfordhb/9780199735471.013.001

Kousaie, S., & Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia*, *94*, 23-35. doi: 10.1016/j.neuropsychologia.2016.11.013.

Kubota, M., Chevalier, N., & Sorace, A. (2020). Losing access to the second language and its effect on executive function development in childhood: The case of "returnees." *Journal of Neurolinguistics*, *55*, 100906. https://doi.org/10.1016/j.jneuroling.2020.100906

Kullmann, D. M. (2020). Editorial. *Brain*, *143*, 1045. https://doi.org/10.1093/brain/awaa082

Kurkela, J. L. O., Hämäläinen, J. A., Leppänen, P. H. T., Shu, H., & Astikainen, P. (2019). Passive exposure to speech sounds modifies change detection brain responses in adults. *NeuroImage, 188*, 208–216. https://doi.org/10.1016/j.neuroimage.2018.12.010

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. doi: 10.18637/jss.v082.i13

Labov, W. (1984). Field methods of the project in linguistic change and variation. In J. Baugh & J. Scherzer (Eds.), *Language in Use* (pp. 28-52). Upper Saddle River, New Jersey: Prentice-Hall.

Labov, W. (2001). Principles of linguistic change: Vol. II: Social factors (Blackwell, Oxford).

Laudan, L. (1971). William Whewell on the consilience of inductions. *The Monist*, 368-391. http://www.jstor.org/stable/27902225

Lauro, J., & Schwartz, A. I. (2019). Cognate effects on anaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(3), 381–396. https://doi.org/10.1037/xlm0000601

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*, 394–425. https://doi.org/10.1037/bul0000142

Leibniz, G. W. (1951). The horizon of human doctrine. In P. P. Wiener (Ed.), *Selections (pp.* 73-77). New York: Scribner's Sons (Original work published 1690).

Leivada, E., Westergaard, M., Duñabeitia, J., & Rothman, J. (2020). On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Bilingualism: Language and Cognition*, 1-14. https://doi.org/10.1017/S1366728920000358

Leonard, M.K., Baud, M.O., Sjerps, M.J., & Chang, E.F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communication*, *7*: 13619. https://doi.org/10.1038/ncomms13619

Li, S., Botezatu, M. R., Zhang, M., & Guo, T. (2021). Different inhibitory control components predict different levels of language control in bilinguals. *Memory & Cognition*, *49*(4), 758-770.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, *58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Li, X., Ng, K. K., Wong, J. J. Y., Lee, J. W., Zhou, J. H., & Yow, W. Q. (2021). Bilingual language entropy influences executive functions through functional connectivity and signal variability. *Brain and Language*, *222*, 105026.

Linck, J. A. (2016). Analyzing individual differences in second language research. *Cognitive Individual Differences in Second Language Processing and Acquisition*, *3*, 105. doi: 10.1075/bpa.3.06lin

Linck, J. A., Kroll, J. F., & Sunderman, G. (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*, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Linck, J., Schwieter, J., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*, 651–662. doi: 10.1017/S136672891100054X

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2020). The differential role of executive functions in the cognitive control of language switching. *Languages*, *5*(2), 19.

Liu, T., Wu, S., Tao, H., Zeng, G., & Zhou, F. (2020). Prevalence of IgG antibodies to SARS-CoV-2 in Wuhan: Implications for the ability to produce long-lasting protective antibodies against SARS-CoV-2. *MedRxiv*, *169*, 1–30. https://doi.org/10.1101/2020.06.13.20130252

Lorca-Puls, D. L., Gajardo-Vidal, A., White, J., Seghier, M. L., Leff, A. P., Green, D. W., Crinion, J. T., Ludersdorfer, P., Hope, T. M. H., Bowman, H., & Price, C. J. (2018). The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings.

Neuropsychologia, 115, 101–111. https://doi.org/10.1016/j.neuropsychologia.2018.03.014

Lorenzo, F., & Moore, P. (2009). European language policies in monolingual southern Europe: Implementation and outcomes. *European Journal of Language Policy*, *1*, 121–135. https://doi.org/10.3828/ejlp.2009.3

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605–621. doi: 10.1080/20445911.2013.795574

Luk, G., Anderson, J. A., Craik, F. I., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and cognition*, *74*(3), 347-357.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*, 29–41. doi: 10.1016/j.cognition.2009.08.014

Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at language learning: Promoting transparency, replication, and a synthetic ethic in the language sciences.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research*, *50*, 940–967. doi: 10.1044/1092-4388(2007/067)

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013).

Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*, 574-588. http://dx.doi.org/10.1016/j.jml.2013.08.001

Markowitz, D. M., & Hancock, J. T. (2014). Linguistic traces of a scientific fraud: The case of Diederik Stapel. *PloS one*, *9*(8), e105937.

Mashburn, C. A., Tsukahara, J. S., & Engle, R. W. (2020). Individual differences in attention control. *Working Memory: The state of the science*, 175.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of memory and language*, *94*, 305-315.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean?. *American Psychologist*, *70*(6), 487.

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. Nature Neuroscience. 7, 703–704. doi:10.1038/nn1264

Mendez, M. F. (2019). Bilingualism and dementia: cognitive reserve to linguistic competency. *Journal of Alzheimer's Disease*, *71*(2), 377-388.

Mendez, M. F., Chavez, D., & Akhlaghipour, G. (2019). Bilingualism delays expression of Alzheimer's clinical syndrome. *Dementia and geriatric cognitive disorders*, *48*(5-6), 281-289.

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, *17*(1), 89-117.

Mertzen, D., Lago, S., & Vasishth, S. (2020, February 28). The benefits of preregistration for hypothesis-driven bilingualism research. https://doi.org/10.31234/osf.io/nm3eg

Meuter, R. F. (2005). Language selection in bilinguals. *Handbook of bilingualism: Psycholinguistic approaches*, 349-370.

Meuter, R., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40. doi:10.1006/jmla.1998.2602

Miceli, S. (2019, Sept 3). Reproducibility and replicability in research. Retrieved from: https://www.nationalacademies.org/news/2019/09/reproducibility-and-replicability-in-research

Ministerio de Educación, Cultura y Deporte. (2012). Estudio europeo de competencia lingüística EECL: Informe español. Madrid: Secretaría General Técnica. https://www.comunidad.madrid/sites/default/files/doc/educacion/sgea_eval_eecl_nacional_vol1_2011.pdf. (accessed 5 January 2020).

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective, & Behavioral Neuroscience*, *11*, 32-43. doi:10.3758/s13415-010-0009-7

Misra, M., Guo, T., Bobb, S., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, *67*, 224–237. doi:10.1016/ j.jml.2012.05.001

Morales, J., Gómez-Ariza, C.J., & Bajo, M.T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology, 25*, 531-546. doi:10.1080/20445911.2013.807812

Morales, J., Yudes, C., Gómez-Ariza, C. J., & Bajo, M. T. (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia*, *66*, 157–169. doi:10.1016/ j.neuropsychologia.2014.11.014

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P., & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations?. *Cognition*, *118*(2), 286-292.

Mukadam, N., Sommerlad, A., & Livingston, G. (2017). The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: A systematic review and meta-analysis. *Journal of Alzheimer's Disease*, *58*, 45-54. https://doi.org/ 10.3233/JAD-170131

Muysken, P. (1997). Code-switching processes: Alternation, insertion, congruent lexicalization. *Language choices: Conditions, constraints, and consequences*, *20*, 361-380.

Nair, V., Rayner, T., Siyambalapitiya, S., & Biedermann, B. (2021). Domain-general cognitive control and domain-specific language control in bilingual aphasia: A systematic quantitative literature review. *Journal of Neurolinguistics*, *60*, 101021. https://doi.org/10.1016/j.jneuroling.2021.101021

Navarro-Torres, C. A., Dussias, P. E., & Kroll, J. F. (*under review)*. When Exceptions Matter: Bilinguals regulate their dominant language to exploit structural constraints. *Language, Cognition and Neuroscience.*

Navarro-Torres, C. A., Garcia, D. L., Chidambaram, V., & Kroll, J. F. (2019). Cognitive Control Facilitates Attentional Disengagement during Second Language Comprehension. *Brain sciences*, *9*(5), 95. doi: 10.3390/brainsci9050095

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), Visual information processing (pp. 283–308). New York, NY: Academic Press

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science, 31*, 548–567. https://doi.org/10.1177/0956797620903113

Noor, S, Friston, K.J., Ekert, J.O., Price, C.J., & Green, D.W. (2020). Neuromodulatory control and language recovery in bilingual aphasia: an active inference approach. *Behavioral Sciences*, *10*, 161. doi:10.3390/bs10100161

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological review*, *114*(4), 887.

Nozari, N., & Novick, J. (2017). Monitoring and control in language production. *Current Directions in Psychological Science*, *26*, 403-410. doi:10.1177/0963721417702419

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism*, *21,* 867-879. doi:10.1017/S1366728918000020

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Ophir, E., Nass, C., Wagner, A. D., & Posner, M. I. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 15583–15587. doi:10.1073/pnas.0903620106

Otwinowska, A., Marecka, M., Casado, A., Durlik, J., Szewczyk, J., Opacki, M., & Wodniecka, Z. (2021). Does L2 proficiency impact L2-L1 transfer while reading L1 collocations? Evidence from behavioral and ERP data. *Frontiers in psychology*, *12*.

Paap, K. R., Anders-Jefferson, R., Mason, L., Alvarado, K., & Zimiga, B. (2018). Bilingual advantages in inhibition or attentional control: More challenges. *Frontiers in Psychology*, *9*, 1409. https://doi.org/10.3389/FPSYG.2018.01409

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*, 232-258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Mason, L., & Anders-Jefferson, R. (2021). Predictions about the Cognitive Consequences of Language Switching on Executive Functioning Inspired by the Adaptive Control Hypothesis Fail More Often than Not. *Brain Sciences*, *11*(9), 1217.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., Mehler, J., (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex*, *13*, 1047–3211. https://doi.org/10.1093/cercor/13.2.155

Palma, P., & Titone, D. (2020). Bilingual lexical access and reading. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution* (pp. 159–183). Cambridge University Press. https://doi.org/10.1017/9781316535967.008

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. Journal of Cognitive Neuroscience, 12, 2728–2744. https://doi.org/10.1162/jocn.2009.21393

Paradis, M., Goldblum, M-C., & Abidi, R. (1982). Alternate antagonism with paradoxical translation in two bilingual aphasic patients. *Brain and Language*, *15*, 55-69. https://doi.org/10.1016/0093-934X(82)90046-3

Parafita Couto, M. C., Munarriz, A., Epelde, I., Deuchar, M., & Oyharçabal, B. (2015). Gender conflict resolution in Spanish–Basque mixed DPs. *Bilingualim, Language, and Cognition, 18*, 304–323. doi: 10.1017/S136672891400011X

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*, 1–23. https://doi.org/10.1037/h0093840

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current opinion in neurobiology*, *15*(2), 202-206.

Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K., & Ferraro, V. (2001).

Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, *28*, 453–496. https://doi.org/10.1017/S0305000901004718

Pfaff, C. W. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, 55, 291-318. doi:10.2307/412586

Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology*, *3*, 57. doi:10.3389/fpsyg.2012.00057

Pliatsikas, C., & Luk, G. (2016). Executive control in bilinguals: A concise review on fMRI studies. *Bilingualism*, *Language, and Cognition, 19*, 699–705. https://doi.org/10.1017/S1366728916000249

Pliatsikas, C., DeLuca, V., & Voits, T. (2020). The many shades of bilingualism: Language experiences modulate adaptations in brain structure. *Language Learning*, *70*(S2), 133-149.

Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, *23*(1), 4–20. Cambridge Core. https://doi.org/10.1017/S1366728919000245

Pollán, M., Pérez-Gómez, B., Pastor-Barriuso, R., Oteo, J., Hernán, M. A., Pérez-Olmeda, M., Sanmartín, J. L., Fernández-García, A., Cruz, I., Fernández de Larrea, N., Molina, M., Rodríguez-Cabrera, F., Martín, M., Merino-Amador, P., León Paniagua, J., Muñoz-Montalvo, J. F., Blanco, F., Yotti, R., Gutiérrez Fernández, R., … Vázquez de la Villa, A. (2020). Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. *The Lancet*, *396*, 535–544. https://doi.org/10.1016/S0140-6736(20)31483-5

Poplack, S. (1983). Bilingual competence: Linguistic interference or grammatical integrity? In E. Olivares (Ed.), *Spanish in the US Setting: Beyond the Southwest* (Arlington: National Clearinghouse for Bilingual Education), 107–131.

Poplack, S. (1987). Contrasting patterns of code-switching in two communities" in aspects of multilingualism. In E. Wande, J. Anward, B. Nordberg, L. Steensland, and M. Thelander (Eds.), *Proceedings of the Fourth Nordic Symposium on Bilingualism*, (Uppsala: Borgström), 51–77.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of codeswitching. *Linguistics*, *18*, 581–618. doi:10.1515/ling.1980.18.7-8.581

Poplack, S., Zentz, L., & Dion, N. (2012). Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition*, *15*, 203-225. doi:10.1017/S136672891100020

Pot, A., Keijzer, M., & de Bot, K. (2018). Intensity of multilingual language use predicts cognitive performance in some multilingual older adults. *Brain Sciences*, *8,* 92. doi:10.3390/brainsci8050092

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

Poulisse, N. (1997). Language production in bilinguals. *Tutorials in bilingualism: Psycholinguistic perspectives*, 201-224.

Poulisse, N., & Bongaerts, T. (1994). First language use in second

Pousada, A. (Ed.). (2017). *Being bilingual in Borinquen: Student voices from the University of Puerto Rico.* Newcastle upon Tyne, NE6 2PA, UK: Cambridge Scholars.

Pulido, M., & Dussias, P. (2020). Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning. *Bilingualism: Language and Cognition, 23*(3), 652-667. doi:10.1017/S1366728919000622

Pulido, M. F. (2021). Individual chunking ability predicts efficient or shallow L2 processing: Eye-tracking evidence from multiword units in relative clauses. *Frontiers in Psychology*, *11*, 4004. https://doi.org/10.3389/fpsyg.2020.607621

R Development Core Team. (2014). R: A language and environment for statistical computing. Retrieved from http://www.R-project.org.

Ridderinkhof, R. K., Van den Wildenberg, W. P. M., Wijnen, J., & Burle, B. (2004). Response inhibition in conflict tasks is revealed in delta plots. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 369-377). New York, NY: Guilford Press.

Rodriguez-Fornells, A., Kramer, U., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology*, *2*, 388. doi:10.3389/fpsyg.2011.00388

Rohrer, J. M., & Arslan, R. C. (2021). Precise answers to vague questions: Issues with interactions. *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/25152459211007368

Rossi, E., Newman, S., Kroll, J. F., & Diaz, M. T. (2018). Neural signatures of inhibitory control in bilingual spoken production. *Cortex*, *108*, 50-66.

Sabourin, L., Stowe, L. A., & De Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *22*(1), 1-29.

Salig, L. K., Valdés Kroff, J. R., Robert, S. L., & Novick, J. M. (2021). Moving from bilingual traits to states: Understanding cognition and language processing through moment-to-moment variation. *Neurobiology of Language*, 1–62. Advanced publication. https://doi.org/10.1162/nol_a_00046

Salzman, D. C., Murasugi, C. M., Britten, K. H., & Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of Neuroscience, 12*, 2331–2355.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*, 231-252. doi:10.1017/S1366728909990514

Santillán, J. & Khurana, A. (2017). Developmental associations between bilingual experience and inhibitory control trajectories in Head Start children. *Developmental Science*, e12624 http://doi.org/10.1111/desc.12624

Schroeder, S. R. (2018). Do bilinguals have an advantage in theory of mind? A meta-analysis. *Frontiers in Communication*, *3*, 36. https://doi.org/10.3389/fcomm.2018.00036

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, *22*, 261–273. https://doi.org/10.1080/00335557043000203

Simons, D. J., Holcombe, A. O., & Spellman, B. A. (2014). An introduction to registered replication reports at perspectives on psychological science. *Perspectives on Psychological Science*, *9*(5), 552-555.

Van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning*, *62*, 148-171.

Sanoudaki, E., & Thierry, G. (2015). Language non-selective syntactic activation in early bilinguals: the role of verbal fluency. *International Journal of Bilingual Education and Bilingualism*, *18*(5), 548-560.

Sebastián-Gallés, N., Martí, A., Cuetos, F., & Carreiras, M. (2000). *LEXESP: Base de datos informatizada del español*. Barcelona: Edicions de la Universitat de Barcelona.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*(772) 1-10. doi:10.3389/fpsyg.2014.00772

Shitova, N., Roelofs, A., Schriefers, H., Bastiaansen, M., & Schoffelen, J.-M. (2017). Control adjustments in speaking: Electrophysiology of the Gratton effect in picture naming. *Cortex*, *92*, 289–303. doi:10.1016/j.cortex.2017.04.017

Simpson, D. (2005). Phrenology and the neurosciences: Contributions of F. J. Gall and J. G.

Spurzheim. *ANZ Journal of Surgery*, *75*, 475–482. https://doi.org/10.1111/j.1445-2197.2005.03426.x

Squires, L. (2014). Social differences in the processing of grammatical variation. University of Pennsylvania Working Papers in Linguistics, 20, 178–188.

Stout, D., & Chaminade, T. (2012). Stone tools, language, and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 75–87. https://doi.org/10.1098/rstb.2011.0099

Sullivan, M. D., Poarch, G. J., & Bialystok, E. (2017). Why is lexical retrieval slower for bilinguals? Evidence from picture naming. *Bilingualism: Language and Cognition*, 1-10. doi:10.1017/S1366728917000694

Sundh, J., Collsiöö, A., Millroth, P. & Juslin, P. (2021). Precise/not precise (PNP): A Brunswikian model that uses judgment error distributions to identify cognitive processes. *Psychonomic Bulletin & Review, 28,* 351–373. https://doi.org/10.3758/s13423-020-01805-9

Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *108*, 834-853. https://doi.org/10.1016/j.neubiorev.2019.12.014

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 1-15. doi:10.1017/S1366728917000682

Szűcs, D., & Ioannidis, J. P. (2019). Sample size evolution in neuroimaging research: an evaluation of highly-cited studies (1990-2012) and of latest practices (2017-2018) in high-impact journals. *NeuroImage*, *221*, 117164. https://doi.org/10.1016/j.neuroimage.2020.117164

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence*. Bilingualism: Language and Cognition, 16*, 367– 382. https://doi.org/10.1017/S1366 72891 2000302

Tanner, D., & van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289–301. https://doi.org/10.1016/j.neuro psych ologia.2014.02.002

Teubner-Rhodes, S. E., Mishler, A., Corbett, R., Andreu, L., Sanz-Torrent, M., Trueswell, J. C., & Novick, J. M. (2016). The effects of bilingualism on conflict monitoring, cognitive control, and garden-path recovery. *Cognition*, *150*, 213-231. doi:10.1016/j.cognition.2016.02.011

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530-12535.

Timmermeister, M., Leseman, P., Wijnen, F., & Blom, E. (2020). No Bilingual Benefits Despite Relations Between Language Switching and Task Switching. *Frontiers in Psychology*, *11*, 1832.

Thothathiri, M., Asaro, C. T., Hsu, N. S., & Novick, J. M. (2018). Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition*, *178*, 162-177.

Tillett, R. L., Sevinsky, J. R., Hartley, P. D., Kerwin, H., Crawford, N., Gorzalski, A., Laverdure, C., Verma, S. C., Rossetto, C. C., Jackson, D., Farrell, M. J., Van Hooser, S., & Pandori, M. (2020). Genomic evidence for reinfection with SARS-CoV-2: a case study. *The Lancet Infectious Diseases*, *3099*, 1–7. https://doi.org/10.1016/S1473-3099(20)30764-7

Timmer, K., Wodniecka, Z., & Costa, A. (2021). Rapid attentional adaptations due to language (monolingual vs bilingual) context. *Neuropsychologia*, *159*, 107946.

Tiv, M., Gullifer, J. W., Feng, R. Y., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics*, *56*, 100913. https://doi.org/10.1016/j.jneuroling.2020.100913

Tomić, A., & Kroff, J. R. V. (2021). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*, 1-12.

Torres Cacoullos, R., & Travis, C. E. (2018). *Bilingualism in the Community: Code-switching and Grammars in Contact*. Cambridge University Press.

U.S. Census Bureau. (2015). Community Facts: State College borough, PA. Retrieved from: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=CF

Valdés Kroff, J. R. (2016). Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing.

Valdés Kroff, J. R., Dussias, P. E., Gerfen, C., Perrotti, L., & Bajo, M. T. (2017). Experience with code-switching modulates the use of grammatical gender during sentence processing. Linguistic Approaches to Bilingualism. doi:10.1075/lab.15010.va.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*, 3–24. doi:10.1017/S1366728914000522

Van Assche, E., Duyck, W., & Gollan, T.H. (2013). Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1781-1792. doi:10.1037/a0032859

van der Slik, F., Schepens, J., Bongaerts, T., & van Hout, R. (2021). Critical Period Claim Revisited: Reanalysis of Hartshorne, Tenenbaum, and Pinker (2018) Suggests Steady Decline and Learner-Type Differences. *Language Learning*.

Verreyt, N., Woumans, E. V. Y., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*, 181-190. doi: 10.1017/S1366728914000352Published

Vetter, P., Bola, Ł., Reich, L., Bennett, M., Muckli, L., & Amedi, A. (2020). Decoding natural sounds in early "visual" cortex of congenitally blind individuals. *Current Biology*, *30*, 3039-3044.e2. https://doi.org/10.1016/j.cub.2020.05.071

Vingron, N., Furlani, N., Mendelson, O., Thomas, M., & Titone, D. (2021). Bilingual Non-Selective Activation Modulates Attention During Visual Object Search. *Journal of Vision*, *21*(9), 2716-2716.

von Bastian, C. C., Blais, C., Brewer, G. A., Gyurkovics, M., Hedge, C., Kałamała, P., … Wiemers,

E. A. (2020, July 27). Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations. https://doi.org/10.31234/osf.io/x3b9k

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, *7*(6), 632-638.

Weyman, K. M., Shake, M., & Redifer, J. L. (2020). Extensive experience with multiple languages may not buffer age-related declines in executive function. *Experimental Aging Research*, *46*, 1-20. https://doi.org/10.1080/0361073X.2020.1753402

Wu, J., Liang, B., Chen, C., Wang, H., Fang, Y., Shen, S., Yang, X., Wang, B., Chen, L., Chen, Q., Wu, Y., Liu, J., Yang, X., Li, W., Zhu, B., Zhou, W., Wang, H., Li, S., Lu, S., … Fei Deng, -Prof. (2020). SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19 Correspondence. *MedRxiv*, 2020.07.21.20159178. https://doi.org/10.1101/2020.07.21.20159178

Wu, Y. J., & Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *The Journal of Neuroscience*, *33*, 13533–13537. doi:10.1523/JNEUROSCI.4760-12.2013

Yim, O., & Bialystok, E. (2012). Degree of conversational code-switching enhances verbal task switching in Cantonese–English bilinguals. *Bilingualism: Language and Cognition*, *15*, 873-883. doi:10.1017/S1366728912000478

Yuan, Q., Ma, F., Zhang, M., Chen, M., Zhang, Z., Wu, J., ... & Guo, T. (2021). Neural interaction between language control and cognitive control: Evidence from cross-task adaptation. *Behavioural Brain Research*, *401*, 113086.

Zhang, H., Kang, C., Wu, Y., Ma, F., & Guo, T. (2015). Improving proactive control with training on language switching in bilinguals. *NeuroReport*, *26*, 354–359. doi:10.1097/WNR.0000000000000353

Zhang, H., Diaz, M. T., Guo, T., & Kroll, J. F. (in press). Language immersion and language training: Two paths to enhanced language regulation and cognitive control. *Brain & Language*.

Zinszer, B. D., Malt, B. C., Ameel, E., & Li, P. (2014). Native-likeness in second language lexical categorization reflects individual language history and linguistic community norms. *Frontiers in psychology*, *5*, 1203.

Zirnstein, M., Bice, K., & Kroll, J. (2019). Variation in language experience shapes the consequences of bilingualism. *Bilingualism, Executive Function, and Beyond: Questions and Insights [Studies in Bilingualism, 57]*, 35-47.

Zirnstein, M., Van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition, 176*, 87-106. doi:10.1016/j.cognition.2018.03.001

APPENDIX A

Items from the Bilingual Switching Questionnaire (Rodriguez-Fornells et al., 2012):

1. I do not remember or I cannot recall some English words when I am speaking in this language.
2. I do not remember or I cannot recall some Spanish words when I am speaking in this language.
3. I tend to switch languages during a conversation (for example, I switch from Spanish to English or vice versa).
4. When I cannot recall a word in English, I tend to immediately produce it in Spanish.
5. When I cannot recall a word in Spanish, I tend to immediately produce it in English.
6. I do not realize when I switch the language during a conversation (e.g., from English to Spanish) or when I mix the two languages; I often realize it only if I am informed of the switch by another person.
7. When I switch languages, I do it consciously.
8. It is difficult for me to control the language switches I introduce during a conversation (e.g., from English to Spanish).
9. Without intending to, I sometimes produce the Spanish word faster when I am speaking in English.
10. Without intending to, I sometimes produce the English word faster when I am speaking in Spanish.
11. There are situations in which I always switch between the two languages.

12. There are certain topics or issues for which I normally switch between the two languages

APPENDIX B

List of picture names with their corresponding lexical frequency values. English lexical frequency norms were derived from CELEX (Baayen, Piepenbrock, & Gulikers, 1995). The lexical frequencies of the Spanish names were obtained from the LEXESP database (Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000) using the NIM search engine (Guasch, Boada, Ferré, & Sánchez-Casas, 2013).

Language Block	Picture Name	English Translation	Frequency per million
Spanish	abrigo	coat	23.8
Spanish	anzuelo	hook	2.8
Spanish	árbol	tree	34.8
Spanish	arco iris	rainbow	0.4
Spanish	aspiradora	vacuum	0.5
Spanish	bolso	bag	13.5
Spanish	bomba	bomb	26.1
Spanish	bufanda	scarf	4.1
Spanish	caballo	horse	62.9
Spanish	calabaza	pumpkin	2.5
Spanish	cama	bed	135.7
Spanish	cangrejo	crab	0.2
Spanish	casa	house	626.6
Spanish	casco	helmet	17.4
Spanish	cerilla	match	3.6
Spanish	clavo	nail	5.3
Spanish	coche	car	122.2
Spanish	corona	crown	25.2
Spanish	dados	dice	12.3
Spanish	dedo	finger	50.6
Spanish	diente	teeth	6.93
Spanish	dinero	money	205.9
Spanish	émbolo	plunger	0.7
Spanish	escalera	stairs	38.0
Spanish	fresa	strawberry	2.8
Spanish	helado	ice cream	13.3
Spanish	hoja	leaf	25.6

Spanish	hueso	bone	15.1
Spanish	lápiz	pencil	6.9
Spanish	lata	can	10.3
Spanish	león	lion	29.8
Spanish	libro	book	193.3
Spanish	luna	moon	52.2
Spanish	martillo	hammer	5.3
Spanish	muletas	crutches	2.8
Spanish	niño	boy	194.9
Spanish	oreja	ear	21.9
Spanish	pájaro	bird	20.6
Spanish	papalote	kite	0.4
Spanish	payaso	clown	4.1
Spanish	pierna	leg	24.5
Spanish	pistola	gun	26.7
Spanish	plancha	iron	5.15
Spanish	planta	plant	38.2
Spanish	pollo	chicken	11.7
Spanish	puente	bridge	35.5
Spanish	puerta	door	276.6
Spanish	pulpo	octopus	1.6
Spanish	queso	cheese	11.0
Spanish	rana	frog	6.2
Spanish	recogedor	dust pan	0.2
Spanish	reloj	clock	50.5
Spanish	rompecabezas	puzzle	4.6
Spanish	rueda	tire	22.7
Spanish	secador	hairdryer	0.7
Spanish	serrucho	saw	1.1
Spanish	silbato	whistle	1.6
Spanish	silla	chair	48.0
Spanish	tambor	drum	6.8
Spanish	tazón	bowl	1.4
Spanish	teclado	keyboard	5.0
Spanish	tenedor	fork	3.7
Spanish	uvas	grapes	5.7
Spanish	ventana	window	93.4
Spanish	vestido	dress	56.9
Spanish	zanahoria	carrot	2.31
English	airplane		5.7
English	ant		11.7
English	arm		210.4

English	axe	8.6
English	badge	9.2
English	ball	111.5
English	bat	14.4
English	bathtub	1.9
English	bee	16.7
English	bell	41.6
English	bottle	116.2
English	box	102.6
English	braid	1.5
English	brain	74.9
English	bread	74.1
English	broom	7.8
English	butterfly	175.4
English	button	26.2
English	cat	66.8
English	chain	48.6
English	cherry	7.4
English	comb	5.4
English	cow	40.3
English	crib	1.2
English	dog	115.1
English	eye	524.3
English	fish	163.5
English	foot	327.2
English	garlic	6.4
English	ghost	31
English	hand	725.3
English	hanger	1.8
English	hat	68.1
English	heart	164.1
English	key	86.3
English	king	99.7
English	knife	44.2
English	knot	14
English	lobster	3.4
English	lock	15.5
English	mailbox	1.8
English	mushroom	12.7
English	necklace	4
English	newspaper	121.6
English	nose	81.2

English	owl	7.2
English	peacock	3.9
English	popcorn	0.8
English	ring	49.1
English	shoe	79.2
English	slide	12.1
English	slippers	8.8
English	snail	4.5
English	spoon	15.4
English	star	100.8
English	steering wheel	0.2
English	suit	52.4
English	sun	152.4
English	swan	7.5
English	table	235.1
English	tent	43.9
English	umbrella	13.7
English	windmill	8.9

APPENDIX C

Supplementary Materials from Beatty-Martínez, Navarro-Torres et al. (2020)

Table S1

95% Confidence intervals and valid N for participant self-reported characteristics

Measure	Context					
	Separated		Integrated		Varied	
	95% CI	Valid N	95% CI	Valid N	95% CI	Valid N
Age, years	[22.3, 24.9]	29	[19.1, 20.8]	35	[20.1, 22.5]	30
English AoA, years	[5.1, 6.8]	29	[3.4, 5.1]	33	[4.2, 6.6]	30
English immersion, years	[0.5, 2.1]	29	[0.1, 1.6]	35	[6.5, 12.5]	30
Spanish exposure, %	[66.3, 76.9]	28	[57.6, 69.1]	33	[26.9, 37.6]	30
English exposure, %	[18.3, 32.6]	28	[28.3, 38.2]	33	[59.0, 70.6]	30
Spanish proficiency	[9.2, 9.8]	29	[8.9, 9.4]	35	[9.0, 9.7]	29
English proficiency	[7.8, 8.5]	29	[8.6, 9.1]	35	[8.7, 9.4]	30

Table S3

95% Confidence intervals and valid N for scores on the BSWQ subscale by context

Measure	Context					
	Separated		Integrated		Varied	
	95% CI	Valid N	95% CI	Valid N	95% CI	Valid N
Switching directionality:						
English into Spanish	[2.6, 3.2]	24	[3.0, 3.4]	35	[3.2, 3.6]	30
Spanish into English	[2.3, 3.1]	24	[3.3, 3.8]	35	[3.3, 3.7]	30
Contextual switching	[2.7, 3.5]	24	[3.7, 4.2]	35	[3.6, 4.1]	30
Unintended switching	[2.5, 2.9]	24	[2.8, 3.2]	35	[2.5, 3.0]	30

Table S4.

95% Confidence intervals of language production measures by task and context

Variable	Context		
	Separated 95% CI	Integrated 95% CI	Varied 95% CI
Verbal Fluency			
Spanish	[51.4, 57.3]	[41.4, 46.4]	[39.4, 46.4]
English	[41.7, 45.6]	[41.8, 46.0]	[44.5, 51.4]
Picture Naming			
Spanish accuracy	[97.4, 98.9]	[94.2, 96.2]	[89.7, 93.2]
English accuracy	[87.6, 91.6]	[93.0, 95.6]	[91.8, 95.8]
Spanish latency (*ms*)	[1041, 1201]	[1070, 1189]	[1104, 1241]
English latency (*ms*)	[1251, 1386]	[1019, 1140]	[1019, 1152]

Table S5

95% Confidence intervals and valid N for AX-CPT scores by context of language use

	Context					
Condition	Separated		Integrated		Varied	
	95% CI	Valid N	95% CI	Valid N	95% CI	Valid N
AX error rate	[.06, .11]	28	[.08, .12]	34	[.07, 11]	30
AY error rate	[.21, .33]	29	[.20, .33]	34	[.20, .31]	30
BX error rate	[.19, .41]	29	[.10, .20]	34	[.09, .17]	30
BY error rate	[.09, .20]	29	[.04, .12]	34	[.03, .11]	30
AX latency (*ms*)	[283, 309]	29	[271, 295]	34	[291, 338]	30
AY latency (*ms*)	[383, 421]	29	[366, 405]	34	[395, 456]	30
BX latency (*ms*)	[211, 263]	26	[209, 238]	34	[220, 272]	30
BY latency (*ms*)	[235, 291]	29	[228, 261]	34	[243, 290]	30

Table S6

Estimated coefficients from mixed model of language dominance on picture naming RTs

Fixed effects	Estimate	*SE*	*t*-values	*p*	Lower	Upper
(Intercept)	**3.08**	**0.01**	**243.85**	**< 0.001**	**3.05**	**3.10**
Lang	**0.10**	**0.01**	**7.21**	**< 0.001**	**0.07**	**0.13**
Freq	**-0.04**	**0.01**	**-8.00**	**0.001**	**-0.06**	**-0.03**
Integrated	**-0.05**	**0.02**	**-3.20**	**0.002**	**-0.08**	**-0.02**
Varied	**-0.04**	**0.02**	**-2.64**	**0.010**	**-0.08**	**-0.01**
Dom	0.01	0.01	0.47	0.640	-0.02	0.03
Lang*Freq	**-0.04**	**0.01**	**-4.20**	**< 0.001**	**-0.06**	**-0.02**
Lang*Integrated	**-0.10**	**0.02**	**-6.38**	**< 0.001**	**-0.13**	**-0.07**
Lang*Varied	**-0.11**	**0.02**	**-7.43**	**< 0.001**	**-0.14**	**-0.08**
Freq*Integrated	-0.01	0.01	-1.02	0.309	-0.02	0.01
Freq*Varied	-0.01	0.01	-1.15	0.252	-0.02	0.00
Lang*Dom	**0.03**	**0.01**	**2.99**	**0.004**	**0.01**	**0.05**
Freq*Dom	0.00	0.00	0.16	0.871	-0.01	0.01
Integrated*Dom	0.01	0.02	0.87	0.386	-0.02	0.05
Varied*Dom	-0.01	0.02	-0.35	0.725	-0.04	0.03
Lang*Freq*Integrated	**0.05**	**0.01**	**5.04**	**< 0.001**	**0.03**	**0.06**
Lang*Freq*Varied	**0.05**	**0.01**	**5.48**	**< 0.001**	**0.03**	**0.07**
Lang*Freq*Dom	**-0.01**	**0.00**	**-2.61**	**0.009**	**-0.02**	**-0.00**
Lang*Integrated*Dom	-0.01	0.01	-0.39	0.699	-0.03	0.02
Lang*Varied*Dom	-0.00	0.01	-0.25	0.803	-0.03	0.02
Freq*Integrated*Dom	-0.00	0.00	-0.85	0.398	-0.01	0.00
Freq*Varied*Dom	-0.00	0.00	-0.61	0.545	-0.01	0.01
Lang*Freq*Integrated*Dom	0.01	0.01	1.01	0.313	-0.01	0.02
Lang*Freq*Varied*Dom	0.01	0.01	1.06	0.289	-0.01	0.02

Random effects	Variance	*SD*	Correlation	
Intercept | item	0.0027	0.05		
Integrated | item	0.0015	0.04	-0.05	
Varied | item	0.0012	0.03	-0.19	0.66
Intercept | participant	0.0041	0.06		
Lang | participant	0.0024	0.05	-0.17	
Freq | participant	0.0001	0.01	-0.07	-0.05
Residual	0.0142	0.12		

Note. Lang = Language; Freq = Frequency; Dom = dominance index, centered around each group's means. Intercept represents average log RTs for separated-context bilinguals. Bold indicates coefficients that are significantly different from zero.

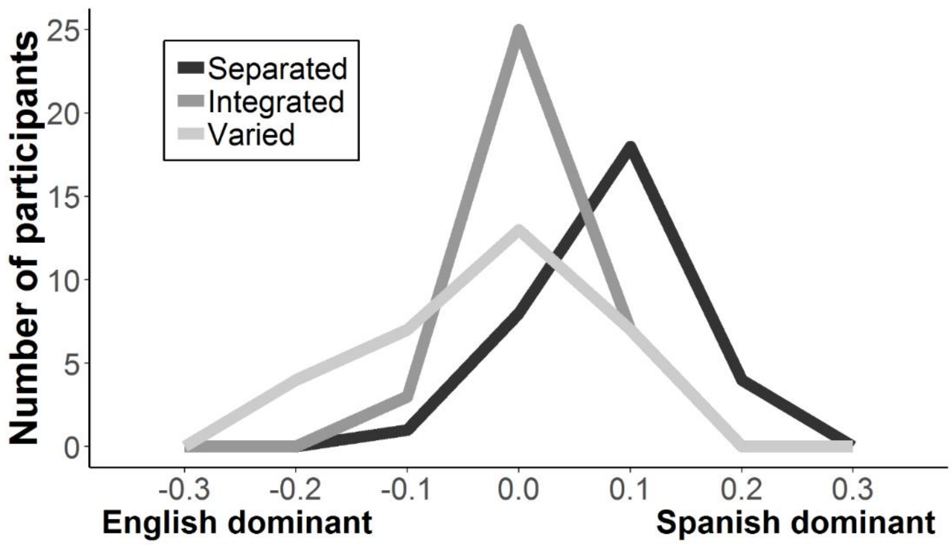

Figure S1. Histogram of the dominance index, which was calculated as the difference between Spanish and English picture naming accuracy. A value of zero on the x-axis represents individuals who obtained a balanced score in the two languages. Values greater than zero represent individuals with higher accuracy in Spanish relative to English.

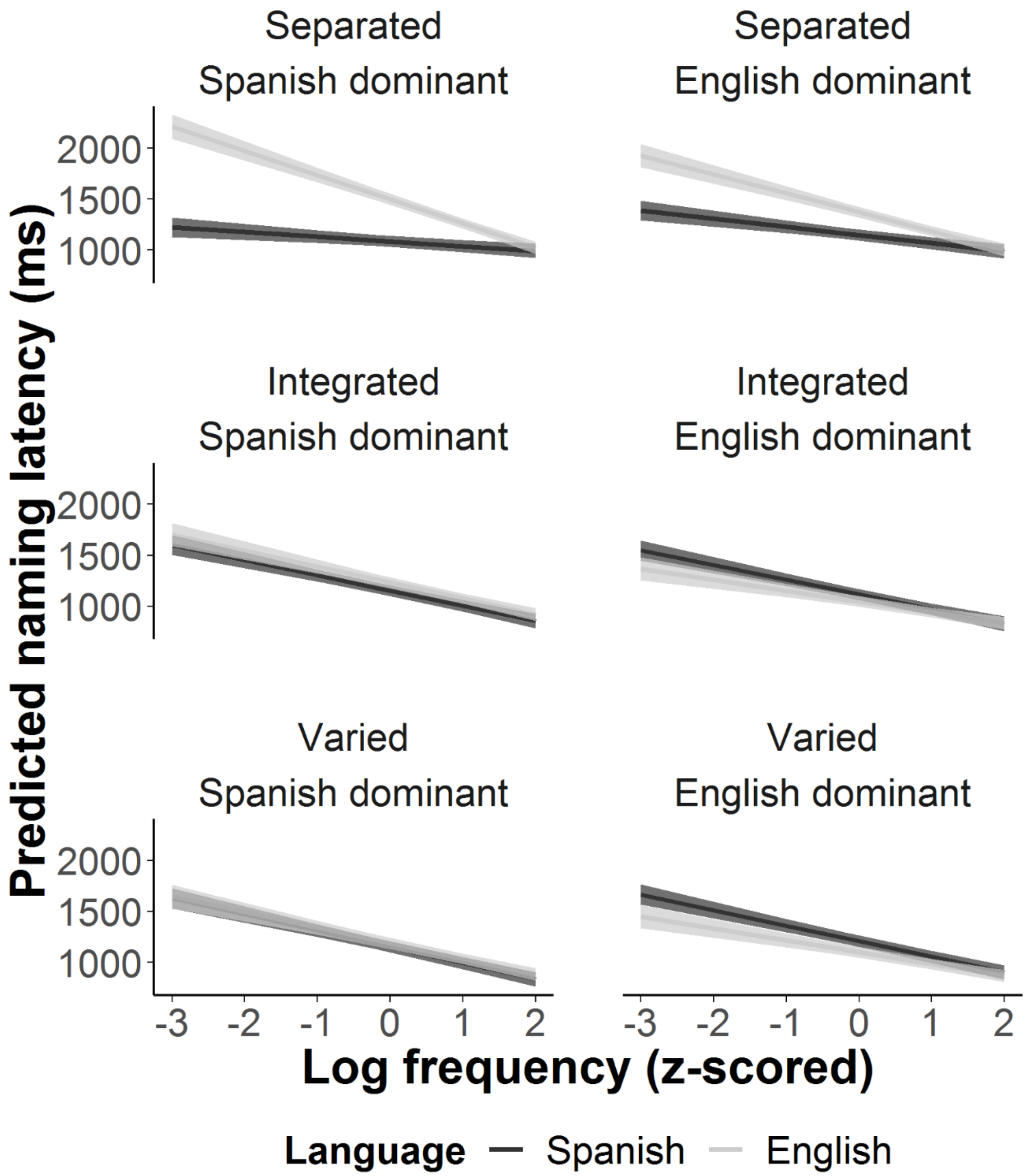

Figure S2. Interaction between language block and frequency for individuals who are Spanish dominant (left column) and English dominant (right column) across each context. Note that, for separated-context bilinguals, the label 'English dominant' denotes individuals who are more balanced (see Figure S1). Negative values on the x-axis indicate lower frequency words. Shaded areas are standard errors of the means.