UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Prior Language Knowledge, the Language Environment, and Cognitive Resources Set the Stage for New Language Learning in Multilinguals

Permalink

https://escholarship.org/uc/item/05k1c30v

Author

Takahesu Tabori, Andrea Akemi

Publication Date

2022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, availalbe at https://creativecommons.org/licenses/by-nc-nd/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,

IRVINE

Prior Language Knowledge, the Language Environment, and Cognitive Resources Set the Stage for New Language Learning in Multilinguals

DISSERTATION

submitted in partial satisfaction of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in Language Science

by

Andrea A. Takahesu Tabori

Dissertation Committee:
Distinguished Professor Judith F. Kroll, Chair
Associate Professor Gregory Scontras
Associate Professor Julio Torres
Professor Mark Warschauer
Assistant Professor Xin Xie

2022

© 2022 Andrea Takahesu Tabori

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There have been three key figures who have mentored me at different stages of my career and continue to support me to this day: Judith Kroll (my graduate advisor at UCI), Jennie Pyers (my undergraduate advisor at Wellesley College), and Jill de Villiers (my PI as a lab manager at Smith College). You are living examples of what a woman can accomplish in this world with ambition and a great deal of curiosity. I also thank my other committee members Gregory Scontras, Julio Torres, Xin Xie, and Mark Warschauer, for your feedback on the project at different stages and for supporting me in my graduate studies more generally.

My journey in graduate school has been an unusual one involving constant relocation, but at the end of the day all of those experiences have helped me become more resilient and grow in different ways. I am thankful to all of the people that I met along the way at Penn State, the UCR, and at UCI.

I would like to thank my professors, colleagues and friends that I met at the Center for Language Science at Penn State: Annie Beatty-Martinez, Haoyun Zhang, Manuel Pulido, Michael Johns, Grant Berry, Melinda Fricke, Lindsey Ma, Paola Dussias, Matthew Carlson, and Mariana Nadeu. The research environment at the CLS was instrumental in shaping my interests and research program.

I would also like to the thank the present and past members of the Bilingualism Mind and Brain Lab: Christian Navarro-Torres, Kinsey Bice, Emily Mech, Justin Sarkis, Natsuki Atagi, Eve Higby, Megan Zirnstein, Marta Rivera Zurita, Jasmin Hernandez Santacruz, Nicole Fuentes Vargas, Guadalupe Mendoza, Alexandra Roman Irrizarry, Nicholas Sullier, Anne Therese Frederiksen, Luciana Brentano, Ingrid Finger, Ariel

Chan, Kellie Weast, and Sibylla Leon Guerrero. I also want to thank all of my fellow doctoral students and professors in the Department of Language Science: Richard Futtrell, Lisa Pearl, Bernard Tranel, Elizabeth Peña, Noa Attali and Jeremy Yeaton. I have learned a great deal about research, collaboration, and life from each of you.

I want to thank all 43 of my research assistants who did the hard work of collecting and coding data during my graduate studies: Yeojin Lee, Ciara Henschke, Dennis Wu, Jorge Juarez Lopez, Arlyn Ballesteros, Miriam Morkos, Belen Leon Perez, Melissa Ramos, Ashley Marin, Matthew Rios, Shayaan Syed, Vanessa Ledesma, Marissa Fernandez, Mayur Upparapalli, Adnan Syed, Zoya Mirza, Siyeon Park, Stephanie Lopez, Claudia Garcia, Edgar Ortuño, Maya Jedrzejczak, Will Zhang, Rachel Vu, Michelle Nguyen, Josue Mena, Emalie Mendoza, Deborah Kamliot, Abimael Hernandez Jimenez, Aaliyah Eslava-Deanda, Giancarlo Arzu, Geoffrey Cui, Kexin Zhang, Yuexin Chen, Yuhao Bi, Yushan Ye, Cesar Rosales, Khahnlihn Ngoc Hoang, Dayra Perez Bernal, Perry Shen Chen, Aadil Chishty, Sungju Kim, Anika Yu, and Shiyang Guo.

My work on this dissertation was funded by a National Science Foundation doctoral dissertation grant, National Science Foundation Partnerships in International Research grant, IRT-LLD fellowship by the US Department of Education, by a Neurolinguistics of Language doctoral dissertation fellowship from the William Orr Dingwall Foundation.

This research would not have been possible without the support of my family: Maria Fales, Robert Fales, Bob Fales and my grandmother Sofia Reyes Castillo (may she rest in peace). Thank you for having the patience to put up with me doing graduate school for almost seven years and moving around from one place to another. Thanks to my English bulldog, Benedict, for being himself and making me laugh all of the time. And to my husband of less than one month, Charles J. Guthy. You are the most supportive and understanding man in the world.

To my new family members, Charles G. Guthy, Patricia Guthy, Catherine McInnis, Dan McInnis, Fulton McInnis, Claire Crosby, and Chris Crosby, I thank you for your help and for always keeping me in your prayers. It has really been a team effort for me to finish graduate school, to get a postdoc in the greater Boston area, move across the country, to get married, and to do all of these things while writing up this dissertation. If it had not been for your intercession and help in various ways, I do not know where I would be.

Above all, I thank Christ Jesus. If I have accomplished anything in my life, it is because He has given me the Grace to do so.

VITA

Andrea A. Takahesu Tabori

Education

2015-2016	B.A. in Psychology Wellesley College
2015-2016	Ph.D. student in Psychology Pennsylvania State University
2016-2019	M.S. in Psychology University of California, Riverside
2019-present	Ph.D. in Language Science University of California, Irvine

Field of Study

Language processing and language learning

Publications

Takahesu Tabori, A., Wu, D., & Kroll, J. F. (under review). "Investigating the relationship between domain-general cognitive control and lexical production in Heritage bilingual speakers of Spanish." *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Kroll, J.F., Takahesu Tabori, A., & Navarro-Torres, C. (2021). Capturing the variation in language experience to understand language processing and learning. In M. Kail & F. Isel (Eds.), Language, plasticity and learning: Developmental issues. Special issue of *Language, Interaction, and Acquisition.*

Mech, E., Takahesu Tabori, A., Kroll, J. F., & De Bot, K. (2019). Psycholinguistics. In N. Schmitt and M. Rodgers (Eds.), *An Introduction to Applied Psycholinguistics, Third Edition.*

Takahesu Tabori, A. A., Mech, E. N., & Atagi, N. (2018). Exploiting language variation to better understand the cognitive consequences of bilingualism. *Frontiers in Psychology*, *9*, 1686.

Kroll, J.F., Takahesu Tabori, A., & Mech, E. (2017). Beyond typical learning: Variation in language experience as a lens to the developing mind. Commentary on Pierce, Genesee, Delcenserie, & Morgan. Linking early language experiences and language learning outcomes. *Applied Psycholinguistics*, 38(*6*), 1336-1340.

Kyuchukov, H., de Villiers, J. G., & Takahesu Tabori, A. (2017). Why Roma children need assessments in their native language of Romani. *Psychology of Language and Communication, 21*, 215-243.

Kyuchukov, H., de Villiers, J. G., & Takahesu Tabori, A. (2016). Assessing Roma children in their native language of Romani. In H. Kyuchukov (Ed.), *New Trends in the Psychology of Language* (31-53). Munich: LINCOM GmbH.

ABSTRACT OF THE DISSERTATION

Prior Language Knowledge, the Language Environment, and Cognitive Resources Set the Stage for New Language Learning in Multilinguals

by

Andrea A. Takahesu Tabori

Doctor of Philosophy in Language Science

University of California, Irvine, 2022

Distinguished Professor Judith F. Kroll, Chair

In this dissertation, I investigated how cognitive resources as well as formal, and informal language experience impact language learning in two studies. In the first study (Chapter 2), I examined the learning of Spanish grammatical gender by Chinese international students who were studying abroad in the US. The goal of that study was to uncover whether being immersed in their non-dominant language (English) to varying degrees and having ambient exposure to other languages in their environment would impact new language learning and modulate the relationship between cognitive control resources and learning. For the learning of grammatical gender, individual differences in the international students' degree of exposure to English predicted higher accuracy in learning. Moreover, international students' active use of an additional Chinese language modulated the relationship between cognitive control and reaction times for learning and generalization of grammatical gender in Spanish. These results suggest that both the

learners' current immersion context, their past language knowledge, and their cognitive resources come to shape new language learning in a third language.

In the second study (Chapter 3), I investigated whether prior knowledge of a tone language as an L1, individual differences in cognitive control and musical experience predicted the ability to perceive tones in a new language. Speakers of tonal native languages (Vietnamese or Bantu languages) and speakers of non-tonal native languages (English or Dutch) were asked to identify Mandarin tones. Critically, Vietnamese has a tone system that is similar to Mandarin because it uses the same two tonal cues as Mandarin (pitch height and pitch direction). On the other hand, Bantu languages are more dissimilar because they only use one of the tonal cues from Mandarin (pitch height). All participants had no prior knowledge of Mandarin, but all were L2 speakers of English, also a non-tonal language. L1 speakers of a tone system that was similar to that of Mandarin (Vietnamese) were more accurate in identifying the Mandarin tones relative to speakers of nontonal languages (Spanish or Dutch). However, L1 speakers of Bantu languages, which have a more dissimilar tone system to that of Mandarin, were less accurate in identifying Mandarin tones compared to speakers of nontonal languages. Higher levels of musical experience also predicted higher Mandarin tone identification, but only for Spanish and Vietnamese groups. Cognitive control abilities, as indexed by performance on the Simon task, did not predict tone identification. These findings are consistent with the claim that speech L2 perception is largely influenced by native language tuning to L1 phonetic cues and with previous research showing that experience tracking pitch in music confers benefits to tracking pitch in a linguistic context.

PREFACE

This dissertation project was initially proposed as a two-experiment study to investigate the learning of Spanish grammatical gender and phonotactic rules for two groups of bilinguals: Mandarin-English bilinguals living in a more linguistically diverse environment (Irvine, California) and Mandarin-English bilinguals living in a more linguistically homogenous location (State College, Pennsylvania).

The first experiment was intended to be an online behavioral study, while the second experiment was intended to be an electrophysiological study involving EEG data collection. For Experiment 1, data collection at both sites was to be carried out remotely. For Experiment 2, the plan was to collect data at our EEG lab at UCI and to travel to Penn State University to collect EEG data there. For Experiment 1, online data collection proceeded largely as planned although the targeted sample was broadened to the United States. This was done because at the time data collection began, COVID-19 cases were on the rise in the US and many of the Chinese international students were returning to China. This was especially the case at UCI, where an estimated eighty percent of the Chinese international students returned to China when the campus closed. Although we could have conducted the research with students in China and in the US for the same experiment, we decided against this because this would have changed the research question of the study. The US data for the online experiment are reported in Chapter 2.

The plan to travel to collect data for the EEG experiment at Penn State was first postponed and later abandoned when there was a surge of COVID cases on the Penn State campus. At that time, Penn State was conducting classes fully in person and had among the highest rates of COVID in the nation. The UCI campus was closed and we did not have access to our lab facilities (either EEG or behavioral equipment). Thus, the original plan to conduct an EEG experiment was also abandoned in favor of increasing the sample size of the online experiment.

Methodologically, one major change was made to the experimental design. I separated the Spanish grammatical gender learning from the phonetic learning component, which were originally embedded into a single learning task. Early pilot work revealed learning of Spanish grammatical gender, but not of the phonetic distinction of interest (between Spanish voiced and voiceless stops). This contrast appeared to be imperceptible to the learners. Because the phonetic distinction was not learnable in one experimental session and since we did not want to sacrifice the learning of the grammatical gender rule by changing the task parameters, we opted for examining phonetic learning of a more salient linguistic feature by other groups of bilinguals in a separate experiment. This gave rise to a new experiment that focused on phonetic learning of tone in Mandarin by Spanish-English, Vietnamese-English, Dutch-English, and Bantu-English bilinguals (Chapter 3).

CHAPTER 1:

Theoretical Motivation and Overview of Doctoral Dissertation

Children all over the world seem to learn the language(s) to which they are exposed, achieving developmental milestones at more or less the same age without specialized instruction from parents. This highly systematic and seemingly self-propelling characteristic of early language development has fascinated language researchers for decades. The mechanisms that underly the ability to acquire language have been and continue to be a topic of interest to this day (Bates & MacWhinney, 1987; Chomsky, 1957; Tomasello, 2005).

One of the most influential accounts of the constraints on language learning was proposed by Eric Lenneberg. Following in the footsteps of Chomsky (1957), Lenneberg (1967) put forth a series of arguments for the role of genetics in language learning. He took the precise timing of language development milestones to reveal a genetically-encoded biological program for learning language. This program seemed to be largely "resistant to the influence of environmental factors" (p. 60). Like Chomsky, he interpreted the fact that children's utterances were not mere copies of their parents' speech, but original utterances, to reflect the primary role of the genetic program and a very limited role of their individual language experiences on their language development.

The central claim in Lenneberg's theory was the observation that language learning was less successful when the learner was exposed to a second language later in life. He observed that most learners who started learning a second language after age twelve had accented speech and limitations in fully acquiring the grammar. It was this claim that has been tested in subsequent research to determine whether adults no longer have access to the same language learning mechanisms that are available during early childhood.

At the time that the Critical Period Hypothesis (CPH) was proposed, there was no direct way of comparing the neural mechanisms of children and adults. The neural mechanisms for language processing were not well-understood until techniques such as functional Magnetic Resonance Imaging (fMRI) and Event Related Potentials (ERPs) began to be used to investigate language processing in the 1990s.[1] The first of the studies that investigated neural processing in bilinguals emerged in the mid 1990s (Klein et al., 1995; Perani et al., 1996).

In the decades following Lenneberg (1967), before the widespread use of neuroimaging, behavioral studies testing the CPH focused on characterizing the extent to which the adult learners could acquire certain aspects of the L2 as a function of when they began learning the L2. These studies showed that later learning of the L2 was associated with poorer performance, especially in the domains of phonetics (e.g. Abrahamsson & Hyltenstam, 2009; Flege & Liu, 2001; Flege, Munro, & Mackay, 1995) and morphosyntax (e.g. Abrahamsson, 2012; Bialystok & Miller, 1999; deKeyser, 2000; Hartshortne, Tenenbaum, & Pinker, 2018; Johnson & Newport, 1989).

Although the findings on the consequences of the age of acquisition (AoA) appear to be consistent with the notion that adults no longer have access to the

[1] Prior to neuroimaging, neuroscientific knowledge was based on clinical case studies. That approach, however, was not used with nonclinical populations.

biologically-ingrained language learning program from early childhood, there are several issues with this interpretation. The first is that the learners compared in those studies differed in the amount of time that they had to learn the L2, with later-exposed learners having less time to learn than earlier-exposed learners (Steinhauer, 2014). The second critical issue is that the shape of the AoA function in many of the studies does not appear be consistent with a discrete and bounded critical period (for a discussion see Bialystok & Miller, 1999; Birdsong, 2006; Birdsong & VanHove, 2016). The third issue is that in addition to any maturational differences, there may also be other existing individual differences could have been confounded with the AoA effects. For these reasons, the evidence that adults no longer have access to the language learning mechanisms from early childhood is far from conclusive. In the next section, I review evidence from studies that have examined new language learning, showing that adults can often learn a second language with great facility.

Plasticity in Early Stages of Learning

Research that has used neural measures to track new language learning suggests that adults are able to learn new language patterns.[2] A seminal study by McLaughlin, Osterhout, and Kim (2004) tracked the neural signatures of classroom French learners and a control group of English monolinguals over the course of a nine months using Event Related Potentials (ERPs). French words and nonwords were presented auditorily to the learners in a lexical decision task. The learners were asked

[2] This is in contrast with studies that used neural measures to examine existing L2 knowledge, which sometimes show differential processing for bilinguals who acquired L2 at different points in life (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996; Wartenburger et al., 2003).

to decide whether the words were real or not real French words. They showed an increased negativity (N400) for nonwords revealing sensitivity to the words over nonwords after only 14 hours of classroom instruction. The learners' negativity increased over the course of two additional testing sessions. On the other hand, the English monolinguals did not show differential processing of the French words and nonwords. These results suggest that minimal instruction with a second language produced rapid neural changes. Other studies that have used neural measures to track new language learning in adults have also found emerging sensitivity to phonetic/phonological (Bice & Kroll, 2019; Cheng, Zhang, Fan, & Zhang, 2019; Wong & Perrachione, 2007), lexical (Bakker, Takashima, Van Hell, Janzen, & McQueen, 2015; Bice & Kroll, 2015), and grammatical (Friederici, Steinhauer, & Pfeifer, 2002; Morgan-Short, Sanz, Steinhauer, & Ullman, 2010; 2012) information.

Studies that have focused on the learning process itself tend to reveal adults' language learning ability more readily than studies that have focused on the product of learning ("ultimate attainment"). The discrepancy in learning outcomes between studies that examine new learning versus knowledge acquired from adults' natural learning environments has important implications. It suggests that instead of rigid biological constraints on learning, there may be environmental factors that limit how much adults learn from their natural learning environments. In the laboratory, we can examine learning conditions that can allow us to capture learning potential. Thus, a key motivation for focusing on new language learning in this doctoral dissertation is to capture adults' language learning potential.

Despite greater control over the learning conditions in a laboratory setting, a somewhat surprising finding in many of the studies that examine new language learning is that even though what the learners are being asked to learn is entirely new, there are striking individual differences in learning across participants. Thus, a complimentary approach is to examine the characteristics of learners that predict learning performance. Some studies have focused on the neural characteristics of the learners (Prat et al., 2016; Rodriguez et al., 2016; Wong, Perrachione, & Parrish, 2007), while others have focused on experiential factors that are theorized to share mechanisms with language processing such as musical experience (Bowles et al., 2016; Shook et al., 2013; Wong & Perrachione, 2007; Zhao & Kuhl, 2015). In behavioral studies, individual differences in cognitive resources have also been investigated in relation to learning (Batolotti & Marian, 2011; Granena, Jackson, & Yilmaz, 2016; Linck & Weiss, 2011; Martin & Ellis, 2012; Morgan-Short, Hamrick, & Ullman, 2022).

One criticism of this approach is that when regarded in isolation, correlations between learning performance and the various individual differences do not conclusively establish whether the individual difference of interest has a causal effect on learning outcomes. However, an important contribution of this approach is first to begin to identify the individual differences that are critical for learning different aspects of the new language. The role of these individual differences can then be more rigorously examined using experimental manipulations. Even so, there are some individual differences that are practically impossible to manipulate in the laboratory (the native language, bilingual status, musical experience, etc). While it is possible to train participants to some degree, it is only by comparing individuals who already differ in these regards that we have been able to discover the role of different experiences in language learning.

Another benefit of this approach is that by examining individual differences, we can then test hypotheses about how the relevant individual differences may be differentially affected by experimental manipulations. Although the experimental manipulations may have an average effect on a group of participants, how prior experience and the learning paradigm interact cannot be established without an examination of individual differences. Recent research suggests that prior experience plays a significant role in how well individual learners learn under different learning conditions (Morgan-Short et al., 2010; Perrachione et al., 2011; Sanz et al., 2016). In the present doctoral dissertation, I used a correlational approach to investigate how the learner's language experience, individual differences in cognitive resources, and other experiences predict new language learning. My doctoral dissertation is comprised of two different studies, one that examined morphosyntactic learning (Chapter 2) and one that examined phonetic learning (Chapter 3).

For logistical and theoretical reasons, I did not consider the same set individual differences in both studies. The specific factors that were examined in each study prioritized the individual differences that have been identified in the past as being the most relevant for that particular aspect of learning. In the study on morphosyntactic learning, I examined cognitive control and ambient language exposure within a group a Mandarin-English bilinguals. In the study on phonetic learning, I considered the role of the native language, musical experience, and cognitive control ability in perceiving tones in a tone language. The first study thus examined variation within a particular language group, while the second involved a set of group comparisons and individual differences analyses within those language groups.

Unlike previous research that has focused almost exclusively on monolingual adults learning a second language, the present dissertation focused on individuals who were already proficient bilinguals or multilinguals at the time of the study and were learning a new language (L3/Ln). Lifelong experience using two or more languages holds consequences for language processing and potentially for language learning. There are long-term adaptations due to bilingual language experience (Bialystok, 2017; Kroll, Bobb, & Hoshino, 2014). Focusing on language learning in multilinguals is an acknowledgment that in addition to the factors that shape language learning in monolinguals, prior bilingual language knowledge and the opportunities that bilinguals' language environment afford to use those languages may come to bear on new language learning. In the next sections, I briefly review the evidence that bilinguals adapt to both their language environments and to the specific languages that they know.

Long-term Adaptations to Bilingual Experience

Neuroscience has shed light on how knowing two languages proficiently leads to long-term neural changes. Proficient bilinguals represent their two languages in largely the same brain areas (Perani & Abutalebi, 2005; Honey, Thompson, Lerner, & Hasson, 2012; but see Xu, Baldaulf, Chang, Desimone, & Tan, 2017). One of the most important discoveries about bilingualism is that bilinguals activate their two languages without being consciously aware of doing so (Spivey & Marian, 2003; Thierry & Wu, 2007). Despite the dynamic activation and interactions across the two languages, proficient bilinguals are very skilled in producing and comprehending their two languages without confusion. There is behavioral (Blumenfeld & Marian, 2011; 2013; Linck et al., 2012) electrophysiological (Declerck et al., 2021; Jackson et al., 2001; Misra et al., 2012) and neuroimaging evidence (Guo et al., 2011; Kang et al., 2017) that bilinguals draw upon domain-general cognitive resources to regulate the activation of their two languages.

Lifelong experience using two languages and managing their coactivation alters brain function (DeLuca et al., 2019; Ferjan-Ramirez et al., 2017; Gold et al., 2013) and structure (De Luca et al., 2019; Garcia-Penton et al., 2014; Luk et al., 2011; Felton et al., 2017; Mechelli et al., 2004). These neural adaptations involve the recruitment of domain-general cognitive control resources for processing the two languages. It is important to note that monolinguals also recruit cognitive control when processing highly ambiguous structures (Hsu & Novick, 2016; Novick, Trueswell, & Thompson-Schill, 2010; Novick et al., 2014), but in bilinguals there is an additional requirement for regulating the activation of the two languages. In bilinguals, parallel language activation can be observed at the phonological (Chang, 2012), lexical (Jacobs, Fricke & Kroll, 2016), and syntactic (Dussias & Sagarra, 2007; Sanoudaki & Thierry, 2015) levels of processing.

Importantly, although all bilinguals have to juggle their two languages, not all bilinguals have the same communication requirements. Depending on where they live and on their language communities, bilinguals will use their two languages with the same or different speakers. Some bilinguals may use their dominant or native language more frequently, while others must use their nondominant language. There is increasing evidence that the language environment of the bilingual modulates which cognitive resources are recruited for language processing (Beatty-Martínez et al. 2020; Green & Abutalebi, 2013; Zhang, 2021).

Speaking two languages can be regarded as an exercise in managing uncertainty (Gullifer & Titone, 2021). Even though bilinguals consider their two languages for activation, they eventually must select one language. The recruitment of cognitive control is thought to play an integral part in language selection. However, bilinguals differ in the degree to which there is overlap in their use of the two languages with certain speakers, in certain contexts, and to talk about certain topics (Tiv et al., 2020). Thus, uncertainty arises in different ways across bilingual environments. For bilinguals who consistently encounter their languages in specific contexts and with specific interlocutors, it may be straightforward to predict which language to select. For bilinguals who use their two languages in overlapping contexts, the ability to select the relevant language depends on the immediate context and the language cues that arise during interactions. Previous research suggests that monitoring the environment for such language cues requires a proactive form of control whereas selecting the target language in a more predictable context requires reactive control (Beatty-Martínez et al., 2020; Gullifer et al., 2018).

One of the goals of this dissertation was to investigate whether bilingual adaptations to a particular interactional context not only shape cognitive control resources for processing the known languages but also influence the learning of an L3. This research question is addressed in Chapter 2, which examined the relationship between cognitive control resources and grammatical gender learning outcomes in L2-immersed bilinguals. I examined whether proactive control resources, which are recruited by L2-immersed bilinguals for processing their known languages (Beatty-Martínez et al., 2020; Zhang et al., 2021), were also predictive of L3 grammatical gender learning.

Adaptations to the native language

The native language appears to shape how well a new language is learned. In the domain of speech perception, early language experience tunes the individual to the features that were heard in the languages to which they were exposed during infancy (Werker & Tees, 1984). The speech categories from the native language(s) have been argued to form a kind of template through which subsequent speech sounds are perceived (Iverson et al., 2013; Kuhl et al., 2008). One of the factors that appears to be critical for whether new speech sounds can be perceived accurately is the perceived similarity between L1 and new speech categories. Speech sounds that are considered to be similar or equivalent by the learner (Flege & Bohn, 2020; Best & Tyler, 2007) are assimilated into the native language categories. Speech sounds that are judged to be dissimilar can form a new category.

Native language similarity also appears to play a role in learning the morphosyntax of a new language. For example, the ability to acquire grammatical gender appears to depend on whether the native language of the learner also has grammatical gender and also on whether it is the same type of grammatical gender system (masculine/feminine vs. masculine/feminine/neuter). A set of studies by Sabourin and colleagues found that both accuracy in learning grammatical gender in a new language and the processing of grammatical gender depend on the similarity between the gender systems in the native and second language (Sabourin, Stowe, & De Haan, 2016; Sabourin & Stowe, 2008). Learners whose native language has a grammatical gender system have better grammatical gender learning outcomes than those who do not have a grammatical gender system. For learners who have a grammatical gender system in their L1, having the same type of gender system in the L2 predicts better learning outcomes and similar processing in the L1 and L2.

There are accounts that posit that exposure to a native language results in the setting of parameters for universal grammar (Broselow & Finer, 1991), which might explain why having a similar system in the L1 predicts better learning of an L2 that has similar features (Phinney, 1987). However, there is no discrete “critical period” for syntactic development that has been shown to result in a sudden decline in L2 grammar learning ability (but see Hartshorne et al., 2018). The fact that adults who do not have the same grammatical gender system or a grammatical gender system in their native language can learn grammatical gender (Hopp, 2010) suggests that constraints on learning grammar are not as hard-wired as those for phonetic development. In other words, the parameters may be able to be “reset” once a sufficient level of experience with the grammatical feature is reached.

In Chapter 2, I examined the learning of grammatical gender by speakers of an L1 and L2 without grammatical gender (Mandarin and English). Grammatical gender is one of the most difficult aspects of morphosyntax to learn in a second language, yet it is learnable by adults. I investigated whether cognitive control resources and language exposure predict the ability to learn grammatical gender. If cognitive resources and language exposure differentially predict the ability to acquire grammatical gender in speakers without prior knowledge of a grammatical gender system, this would indicate that adult learners' ability to learn an unfamiliar linguistic feature is not set in stone by early language experience, but rather a function of the learner's cognitive abilities and language experience.

In Chapter 3, I compared the performance of different groups of bilinguals in learning tone in Mandarin as an L3. All the bilinguals knew English as an L2 but differed in their L1s. The ability to learn a new tone system in Mandarin as a function of the L1 and the similarity between the L1 and L3 was considered. The role of individual differences in musical experience and cognitive control resources was also considered. The goal of Experiment 2 was to examine the extent to which L1 features predict phonetic learning outcomes and how much explanatory power is gained by additionally considering individual differences that are primarily thought to reflect life experience more broadly.

Perceptual adaptations to the language environment

In addition to the demands of communicating in their known languages, individuals also differ in the degree of linguistic diversity afforded by their language environment. Some bilinguals live in places where languages that are unknown to them are widely used while others live in places where there is more limited language diversity. It is only recently that a number of studies have shown that people who have passive or ambient language exposure have some form of implicit knowledge stemming from their ambient language exposure of phonetic/phonotactic information (Au et al., 2002; Knightly et al., 2003) as well as of word-forms (Oh et al., 2020; Saffran et al., 1997). These studies suggest that ambient exposure to multiple language results in incidental learning of speech sounds and word forms.

There is also preliminary evidence that there may be a more general consequence of being exposed to language variation in the environment, particularly for learning new speech sounds. Adult learners who have larger social networks appear to also have better speech perception (Lev-Ari, 2018). This is thought to be because people with larger social networks hear more speaker variation in how speech sounds are realized and speaker variation has been found to be beneficial for learning new speech sounds. There is also evidence that greater ambient language diversity predicts better learning of phonology even when the ambient exposure was to different languages than what is being learned (Bice & Kroll, 2019).

One goal of the present study was to examine whether greater ambient language diversity not only benefits phonetic learning but also the learning of morphosyntactic. This question was examined in Chapter 2, which investigated the learning of Spanish grammatical gender in relation to the learners' social networks and exposure to different languages. In order to capture ambient language diversity, I opted for a two-pronged approach. I collected self-reported measures of ambient language exposure in which participants were asked to report how much they are exposed to different languages in their environment in general and also in their immediate social circle. I also developed a proxy measure of ambient language exposure from which previous exposure to those languages could be inferred. It is important to note that while there are available tools to collect naturalistic data such the Language ENvironment Analysis (Oetting et al., 2009) and Electronically Activated Recorder (Robbins et al., 2018), these tools require significant resources for data collection and transcription. Most importantly, these audio recording devices are typically given to the participants in the laboratory, where they have to be trained on how to wear them, use them, and when to return them. Because the present studies took place entirely online and sampled participants all around the US (for Experiment 1), physical audio recording devices were not an option.

How particular linguistic features engage cognitive resources

There is longitudinal evidence that cognitive control resources are important for language learning, especially from developmental research (Gooch et al., 2016). For bilingual children too, there is also evidence that cognitive control resources are important for language learning (Santillan & Khurana, 2018). However, the vast majority of this evidence comes from studies that look at language very wholistically. Not much is known about which aspects of language learning engage these cognitive resources. A goal of my doctoral research was to examine the relationship between learning of particular linguistic features (grammatical gender in Experiment 1 and tone in Experiment 2) and cognitive control resources.

In Experiment 1, I investigated the learning of grammatical gender in Spanish for Mandarin-English international students living in the US. These bilinguals are hypothesized to be relying on proactive control for regulating language activation. The goal was to examine whether proactive and reactive control, which are the relevant components for regulating language activation in the known languages, will also be recruited for L3 learning of a novel linguistic feature. In order to do this, I used the AX-Continuous Performance task, a continuous measure of proactivity/reactivity as the measure of cognitive control for Experiment 1. I hypothesized that learners would rely on proactive control to learn the Spanish grammatical gender pattern that was presented explicitly and without any exceptions. The logic is that proactivity involves tracking frequent patterns and forming a prediction to anticipate exemplars that conform to the pattern. When there is regularity in the input, learners are hypothesized to rely on proactive control to anticipate the pattern. However, if the expected pattern is violated, learners' might use reactive control to adjust their expectations. Learners who are more proactive are hypothesized to be more rule-oriented whereas learners who are more reactive are hypothesized to be prone to learning with a more exemplar-based focus. Because the learning input only included examples with transparent gender marking, it was hypothesized that proactivity would be more advantageous for learning the grammatical gender rule.

In Experiment 2, I examined the identification of Mandarin tones in bilinguals who knew different native languages that were either tonal (Vietnamese, Bantu) or non-tonal (Dutch, Spanish). The focus was on whether reactive control would play a role in perceiving the new Mandarin tones. Previous research suggests that reactive control efficiency predicts better speech perception (Lev-Ari & Peperkamp, 2013; 2014). This may be because individuals who have better reactive control are able to flexibly adjust to item-specific variation. As previously mentioned, proactivity may be associated with generating strong predictions and requiring more evidence before adjusting to the new information. Thus, reactivity may support the kind rapid adaptation that is involved in speech perception. The cognitive control measure selected for Experiment 2 was the Simon task, a measure of reactive control efficiency.

To review, this doctoral dissertation consists of two studies. The first study investigated the learning of grammatical gender in Spanish by Mandarin-English bilinguals. The goal of this study was to examine learning ability as a function of ambient language exposure and of cognitive control resources which are relevant for processing the known languages. Critically, the first experiment held constant the languages known by the participants. Because neither of their known languages have grammatical gender, this group was particularly likely to draw upon other forms of language experience (ambient language) and cognitive resources for learning grammatical gender in Spanish as an L3.

The second study focused on the learning of Mandarin tones by multilinguals who all knew English as an L2 but had different L1s: Spanish, Vietnamese, Dutch, and Bantu languages. The native languages were either tonal (Vietnamese, Bantu languages) or nontonal languages (Dutch, Spanish). The goal of this study was to examine the extent to which the native language, its similarity to the target language (Mandarin) and individual differences in musical experience and cognitive resources explain the ability to learn Mandarin tones. If the native language and its features fundamentally place a limit on the ability to perceive new features, then we might expect a primary role of the native language and limited evidence for other individual differences such as cognitive control and musical experience. However, the extent to which cognitive control and musical experience explain variation in learning Mandarin tone would provide evidence that constraints set by early language experience are also modulated by other forms of experience.

CHAPTER 2:

Investigating the Relationship Between Individual Differences in the Language Environment and Cognitive Resources for New Morphosyntactic Learning

Although traditional accounts of language learning have assumed the existence of hard constraints beyond a critical period early in life (Lenneberg, 1967), advances in cognitive neuroscience reveal that second language (L2) exposure in adulthood induces rapid neural changes (Hofstetter, Friedmann, & Assaf, 2017; Mclaughlin, Osterhout, & Kim, 2004). Measures of neural activity reveal that there are substantial individual differences in language processing that are not always evident in behavior (Bice & Kroll, 2015; 2019; Wong & Perrachione, 2007). A challenge is to understand these observations of neuroplasticity in the face of the persistent difficulties reported for acquiring certain aspects of the L2, especially morphosyntax (Clahsen & Felser, 2006; Hartshorne, Tenenbaum, & Pinker, 2018; Johnson & Newport, 1991; deKeyser, 2000) and phonetics-phonology (Flege, 2007; Flege, Yeni-Komshian, & Liu, 1999; Iverson et al., 2003).

A prevalent approach in the study of late L2 acquisition has been to examine how L2 learning reflects transfer from L1 and potential interference between the two languages. Several accounts of L2 learning attribute L2 learning difficulties to the way in which the language system is tuned to the L1 (Best, McRoberts, & Goodell, 2001; Iverson et al., 2003; Kuhl et al., 2008; Sabourin, Stowe, de Haan, 2006). Although this approach reveals many specific difficulties for L2 learners based on their L1, there are also individual differences in late L2 learning outcomes among L2 learners with the same L1 (Wong & Perrachione, 2007; Wong et al., 2007). Much less research has investigated the environmental factors and cognitive skills that play a role in new language learning.

The current study focused on investigating the sources of individual variation in language learning outcomes within the same group of bilinguals. I investigated whether individual differences in cognitive control resources and language exposure predict third language (L3) learning in bilinguals. Chinese international students studying abroad in the US were asked to learn grammatical gender in Spanish, a language of which they have no formal knowledge. The Chinese international students had varying degrees of exposure to English, Mandarin, and other languages while living in the US. Of particular interest was the degree to which the language diversity of their current language environment in the US and individual differences in cognitive control skills would contribute to learning Spanish.

In the next sections, I will review the evidence that the bilingual's two languages interact in ways that require cognitive control, but that different components of control are shaped by the context in which the two languages are used. A number of recent studies have tested this idea for language processing, but very little research has delved into the implications that these differences may hold for new language learning. This is the primary aim of this study.

Cognitive Control in Bilingual Language Processing

The relationship between cognitive control and language processing has been investigated for both monolinguals and bilinguals. For monolingual speakers, there is overlap between the substrates that underlie conflict resolution and ambiguity resolution during language processing (Hsu, Jaeggi, & Novick, 2017; Novick, Trueswell, & Schill, 2010). When reading a garden path sentence such as “the horse raced past the barn fell,” there is ambiguity stemming from the word *raced* having two interpretations: one as a finite verb and one as a passive participle. Because the most frequent use of the word *raced* activates the finite verb interpretation, readers are often led down the garden-path and forced to reinterpret once they reach the end of the sentence. Like monolinguals processing ambiguous sentences with multiple interpretations, bilinguals too face ambiguity. However, it is much more pervasive because bilinguals always co-activate their languages in parallel.

Neuroimaging research reveals that the networks of the brain that support L1 and L2 processing in proficient bilinguals are the largely the same (Perani & Abutalebi, 2005; Honey, Thompson, Lerner, & Hasson, 2012; but see Xu, Baldaulf, Chang, Desimone, & Tan, 2017). Consequently, bilinguals’ two languages are always jointly activated and competing for selection (Spivey & Marian, 2003; Thierry & Wu, 2007) even when they are only consciously aware of using one language. This parallel language activation is pervasive and can be observed at the phonological (Chang, 2012), lexical (Jacobs, Fricke & Kroll, 2016), and syntactic (Dussias & Sagarra, 2007; Sanoudaki & Thierry, 2015) levels of processing. Language co-activation fundamentally changes how bilinguals process each of their languages because whether features across languages converge (Hoshino & Kroll, 2008; Van Assche, Duyck Hartsuiker, & Diependaele, 2009) or diverge, can result in either facilitation or interference (Durlik, Szewcyk, Muszynski, & Wodniecka, 2016; Jared & Szucs, 2002). Because there are differences in how bilinguals process their languages relative to monolinguals,

monolinguals are not the best reference for understanding how bilinguals process language. Instead, understanding how different types of bilinguals differ from one another may be more likely to reveal the specific ways in which cognitive resources are recruited to enable language processing.

Two general alternatives have been proposed as ways for bilinguals to select the target language: selective attention to the target language and inhibition of the non-target language. According to the selective attention view, bilinguals might activate both languages but be able to selectively access the target language by exploiting environmental language cues (e.g. Costa, Miozzo, & Caramazza, 1999). The second view is that bilinguals consider both target and non-target languages for selection and they both compete for selection until inhibition is applied to the non-target language (Green, 1998). There is extensive behavioral (Meuter & Allport, 1999; Linck, Sunderman, & Kroll, 2009; Kleinman & Gollan, 2018), neuroimaging (Abutalebi & Green, 2008), and electrophysiological (Misra et al., 2012) evidence supporting the view that bilinguals select the target language of speech by using cognitive control to inhibit the non-target language (Abutalebi & Green, 2008).

One approach in the study of cognitive control has been to examine how bilinguals recruit cognitive control by shifting between proactive and reactive control mechanisms. According to the Dual Mechanism Framework of cognitive control (Braver, 2012), proactive control can be thought of as an anticipatory adaptation of behavior oriented towards accomplishing a future goal. In contrast, reactive control can be understood as a rapid adjustment in behavior when unexpected environmental changes occur. While there is extensive research investigating cognitive control performance in bilinguals (Bialystok, Craik & Viswanathan, 2004; Martin-Rhee & Bialystok, 2008; Morales, Gómez-Ariza, Bajo, 2013; Morales et al., 2015), very little is known about the role cognitive control plays in new language learning (but see Bartolotti et al., 2011).

The focus of the current study was to examine whether cognitive control resources support L3 learning in bilinguals. Using the Dual Mechanism Framework of cognitive control, I investigated whether individual differences in proactivity/reactivity were predictive of new morphosyntactic learning. The critical prediction was that the effect of cognitive control on new learning would be modulated by bilinguals' language environment. Below, I review the evidence that not all bilinguals engage the same cognitive control resources. One critical factor that determines how they recruit cognitive control is the interactional context imposed by their language environment.

Cognitive Control and the Language Environment

A feature of the language environment that has been shown to modulate cognitive control engagement during bilingual language processing is the immersion context. While most bilinguals have a dominant language, they differ in whether they are immersed in their dominant language or their non-dominant language. Previous research suggests that bilinguals who are immersed in their non-dominant language (e.g. study-abroad students) are inhibited in their dominant language relative to bilinguals who are immersed in their dominant language or L1 (Baus, Costa, & Carreiras, 2013; Jacobs et al., 2016; Linck et al., 2009). Critically, although immersed bilinguals have reduced access to their dominant language in the environment, they still continue co-activate their L1 when speaking their L2 (Jacobs et al., 2016). Because L2-

immersed bilinguals are in an environment in which their dominant language is not widely used, they are able to proactively inhibit their entire dominant language whereas the non-immersed bilinguals need to maintain access to their dominant language and might use inhibition reactively as specific language demands arise.

Two recent studies illustrate the role of the interactional context on the recruitment of domain-general cognitive resources during language processing. Beatty-Martínez et al. (2020) compared lexical production in a simple picture naming task for three groups of native Spanish speakers, all of whom were highly proficient in English as the second language. One group lived in Spain, in a location in which Spanish and English were used separately. For the Spain bilinguals, Spanish was the language of the home and community and English was the language used at work and school. A second group lived in Puerto Rico, where Spanish and English are used almost interchangeably and where speakers code switch frequently with each other. The third group lived in the US, where they were studying at an American university in an English dominant context. They code switched with each other but never knew with whom they might be able to speak Spanish. An independent assessment of cognitive control showed that the group immersed in English, the L2, relied on proactive control to control language processing, whereas the group in Spain, using the two languages separately, relied on reactive control. The results suggested that having to monitor the environment to determine with whom you can speak each language is a critical factor that may determine how cognitive resources are allocated to enable proficiency in both of a bilingual's languages.

A study with native Mandarin speakers who were proficient in English as the L2 provided converging support for the conclusion of the Beatty-Martínez et al (2019) study. Zhang, Diaz, Guo, and Kroll (2021) compared the cognitive control ability of Mandarin-English bilinguals living in China from an earlier experiment (Zhang et al., 2015) and of Mandarin-English bilinguals studying abroad in the United States after a short-term cognitive control training study. As shown in previous studies with Mandarin-English bilinguals living in the United States (Zirnstein et al., 2018; Bogulski et al, 2019), these bilinguals maintained their Mandarin dominance despite the predominant influence of English in their environment. Results showed that only the Mandarin-English bilinguals in living in China, but not the bilinguals studying abroad in an English dominant context, benefitted from the cognitive control training. The Mandarin-English bilinguals who were immersed in English in the US had superior proactive control at the start of the experiment, suggesting that the immersion environment itself had the consequences of training cognitive control. Critically, when two groups of bilinguals were compared on performance in a language switching paradigm, the immersed bilinguals showed significantly greater inhibition of Mandarin, the L1. Zhang et al. (2021) proposed that bilinguals who were immersed in their non-dominant language actively engaged proactive control to retain their native language, like the immersed group in Beatty-Martínez et al. (2020).

Evidence for the role of inhibitory control in L2 learning

One of the most important consequences of dual language activation is the need to regulate L1 interference is during L2 learning. Co-activation of the non-target language has not only been observed in proficient bilinguals, but also in relatively less proficient adult L2 learners (Bice & Kroll, 2015; Jacobs, Fricke, & Kroll, 2016), suggesting that the process of learning a second language itself requires cognitive control. Because skilled bilinguals are accustomed to learning concepts with multiple lexical mappings, they may become experts at inhibiting cross-language interference during new language learning. A statistical learning study with adults found that bilinguals learned words with multiple mappings better than monolinguals (Poepsel & Weiss, 2016). These results suggest that bilinguals are indeed more efficient at suppressing ambiguity from multiple lexical representations during language learning. Similarly, studies with infants suggest that cognitive control may be critical for new language learning. Infants with dual language exposure have been found to outperform aged-matched infants with monolingual language exposure when learning a new artificial language (Kovacs & Mehler, 2009).

Although the majority of these studies have been monolingual and bilingual comparisons, there is longitudinal evidence for the role of cognitive control in L2 learning. A recent large-scale longitudinal study by Santillan and Khurana (2018) examined inhibitory control grow in preschoolers attending the Head Start program. The preschoolers differed in whether they were native speakers of English or second language learners of English. The researchers reported that relative to the English monolingual children, the transition from being a monolingual in a non-English language to becoming bilingual was associated with an accelerated increase in inhibitory control growth. Similarly, a study by Hartanto, Toh, and Yang (2019) found that the negative effects of low SES on executive functions were reduced in bilinguals compared to monolinguals. These studies implicate cognitive control in bilingual language learning.

There is relatively little research that has investigated language learning in bilingual adults. Several behavioral studies that investigated language learning in bilinguals found that bilinguals were better word learners than monolinguals (Cenoz & Valencia, 1994; Kaushanskaya & Marian, 2009b; Keshavarz & Astaneh, 2004; Sanz, 2000; Van Hell & Mann, 1997). However, very few studies have examined the role of cognitive control on language learning. Notably, Bartolotti and Marian (2012) examined Morse Code learning outcomes in bilingual speakers of different languages. Since Morse code does not overlap in form with natural languages, language competition from participants' known languages was avoided. The researchers first introduced a Morse sequence and followed it by a second sequence that conflicted with the first. Results showed that individual differences in cognitive control only predicted learning outcomes for the second, high conflict sequence. The results suggest that bilinguals may benefit from inhibitory control when learning an L2 in a high conflict scenarios. A more recent study by Bogulski, Bice, and Kroll (2019) suggests that it is not cognitive control resources per se but bilinguals' expertise using cognitive control to regulate L1 interference that may be important for new language learning. This study compared novel Dutch word learning in different groups of bilinguals, with the novel Dutch words being learned either via L1 or L2. Among the bilingual groups tested, only the bilinguals who learned via the L1 were advantaged relative to monolinguals, suggesting that the regulation of L1 interference during L2 learning may be the aspect of cognitive control that supports language learning. Taken together, preliminary evidence suggests that cognitive control resources support new language learning in bilinguals by aiding in the regulation of L1 interference.

Ambient Language Exposure and Language Learning

Recent psycholinguistic research suggests that ambient exposure to other languages may facilitate new language learning even in monolinguals. Monolingual speakers who have ambient exposure to multiple languages in their environment have been found to be advantaged in learning a phonological rules in a new language compared to monolinguals with less diverse language exposure (Bice & Kroll, 2019). Another illustration of the influence of ambient language diversity on new language learning is the case of childhood overhearers. Childhood overhearers are adults who as children grew up listening to a second language in their environment, although they did not go on to acquire that language. The childhood overhearers that have been studied are English speakers who were exposed to Spanish as a heritage language in their home although they did not go on to become bilingual in English and Spanish, and remained functional English monolinguals. These overhearers of Spanish have been found to be advantaged in learning the phonology/phonetics of Spanish, but not its morphosyntax (Au et al., 2002; Knightly et al., 2003). This evidence suggests that even passive exposure to other languages may be critical for tuning the phonetic/phonological system of the L2 learner, creating a perceptual openness to non-native speech sounds (Petitto et al., 2012).

There is also a growing body of work suggesting that the size of individuals' social network affects speech perception. Lev-Ari has conducted a number of experiments in which she examines the relationship between network size and speech perception in monolinguals. Her research finds that monolinguals who have larger social networks are better able to perceive speech sounds in noise (Lev-Ari, 2018). This may be because speakers with larger networks are exposed to more variability in their input, which may help them to separate the relevant phonetic information from other irrelevant variation during speech perception. To my knowledge, the question of whether there is a relationship between ambient language exposure and the learning of morphosyntax, has not been investigated in past research.

In the current study, I examined whether individual differences in language exposure and cognitive control predict new language learning in adult learners. The learners were Chinese international students studying abroad in the US. They completed the AX-CPT, a measure of cognitive control that measures proactive and reactive control continuously. They also completed the Spanish Noun Phrase learning task, which involved learning Spanish noun phrases (e.g *ese gato/that cat*) and being tested on their knowledge of gender congruency between the article and noun. In addition, participants completed two self-report measures aimed at characterizing their exposure to languages in their environment and to the languages used by people in their social circle.

Predictions

The language environment may play multiple roles in language learning. First, it may modify how cognitive resources are recruited for processing the known languages (Mandarin, English), which may also hold consequences for new learning. If so, we might expect the international students who use their dominant language more in the US so also draw more proactive control resources for maintaining their L1 active in an L2-dominant environment (Beatty-Martínez et al., 2020; Zhang et al., 2021). This recruitment of proactive control resources may also be involved in explicit learning of the grammatical gender rule in Spanish. Proactive control is hypothesized to enable the learners to exploit contextual cues (e.g. word final morpheme and definite article co-occurrence) to bias their behavior in line with the regularities provided in the learning task. The second way that the learning environment may support new learning is in providing more varied exposure to languages, which may predispose the learner to be more sensitive to the novel language patterns that will be learned.

Method

Participants

Seventy-five speakers of Mandarin and English who were born and raised in China but currently living in the US participated in the current study. They were recruited through a flyer (Appendix A) that was posted on the UCI campus, distributed through various online platforms including Facebook, the Chinese social media platform WeChat, Reddit, and through listservs. Interested potential participants first completed a short-screening survey (Appendix B). The screening survey was designed to distinguish between Heritage speakers of Mandarin born and raised in the US (who would be disqualified from the study) and Chinese international students. The flyer was written in Mandarin and specified that the researchers were looking for Chinese international students to participate.

In order to qualify for participation, the participants had to be international students in any school in the US or recent graduates still living in the US, to have been born in China, to be currently residing in the US at the time of the screening, and to not have taken a Spanish course previously.[3] If any languages that have a grammatical gender system were listed in the languages that were known to the participant in the screening survey, the experimenter followed up with the potential participant to determine if the participant had some level of proficiency in that language. Only participants who had achieved a beginning level of proficiency but had discontinued taking coursework in a language with a grammatical gender system could be included in the study. Interested participants who reported having taken a course in a language with a grammatical gender system were asked to explain their current knowledge of that language. The experimenter only invited these participants to join the study if they indicated that they had no longer had functional knowledge of those languages.[4] Eight participants were disqualified due to one or more of the criteria. Five participants were excluded because they had taken more than one year of coursework in a language with grammatical gender, one participant was excluded due to currently taking coursework in a language with grammatical gender, and another due to learning English not as an L2 but an L1.

Of the 75 international students, 70 reported Mandarin as a native language and 5 reported Cantonese as a native language. Thirty-one participants reported knowledge of at least one language other than Mandarin, Cantonese, and English. These additional languages included Japanese (n = 12), French (n = 7), German (n = 1), Taiwanese (n =

[3] One participant indicated having studied Spanish for a week using Duolingo, but that was not considered a course for the purpose of this study.
[4] Participants were specifically asked if they remembered how to say any phrases in the grammatically gendered language that they had studied. They could only participate if they remembered how to say single words (e.g. hello, goodbye, thank you). Once they were invited to participate in the study, they were also asked to report their proficiency in the language they had studied. The role of prior coursework in a language with grammatical was not analyzed in the current study.

3), Korean (n = 6), Spanish (n= 1), Hebrew (n = 1), and Thai (n = 1). Participants' self-reported exposure to Mandarin, English, and other languages is reported in Table 2-1.

Table 2-1. Participant demographic and language history characteristics.

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Mandarin Fluency	58	62.46	10.349	42	56	68	89
English Fluency	57	35.58	6.612	22	31	40	51
Age (Years)	77	22.417	2.929	18.209	20.326	24.282	31.86
% Mandarin Exposure	70	49.486	18.177	4	40	60	85
% English Exposure	70	48.214	18.397	15	36.25	58.75	95
% Other Exposure	71	2.352	5.715	0	0	1	40
English AoA	70	6.429	2.996	0	5	7.75	18

Notes. Verbal fluency scores are the sum of the total number of exemplars produced in each language across the four semantic categories. The percentage of exposure to all languages adds up to one hundred percent. Age and Age of Acquisition (AoA) are in years.

Tasks

Language History Questionnaire

The language history questionnaire included a subset of questions from the Language Experience and History Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) to assess age of acquisition, proficiency, and demographics. In addition, the questionnaire included questions regarding the reason for the participants' living in the US, their place of origin in China, their current place of residence in the US, and questions regarding formal and ambient language exposure both in China and in the US (Appendix C). The questionnaire was administered through Qualtrics and took ten to fifteen minutes to complete.

AX-Continuous Performance Task (AX-CPT)

Materials. A version of the AX-CPT without distractors was used (Gullifer, Chai, Whitford, Pivneva, Baum, Klein, & Titone, 2018). The AX-CPT consisted of 20 practice trials and 200 experimental trials. Of the experimental trials, 140 (70%) were AX trials. The remaining three trial types, AY, BX, and BY had 20 trials each (10%). The non-A cues and non-X probes were chosen to not resemble the letters A and X and included the letters: C, G, J, F, O, P, S, T, U. The same cue and probe letter pairs were never presented simultaneously (Appendix D). Although multiple versions of the AX-CPT have been used in previous research, I chose this specific version because excluding the three distractor letters reduced the time spent on experimental trials. This time savings enabled me to double the number of trials. Increasing the number of trials relative to what is collected in the laboratory was important because the online delivery of the task was expected to introduce additional measurement error.

Procedure. This task is a continuous measure of proactive and reactive control. In this task, participants saw sequences of two letters presented on the screen in rapid succession. The task involved pressing buttons in response to the combination of the first (cue) and last letters (probe). Participants were instructed to respond YES to all trials that have AX cue-probe sequences and to press NO to any other cue-probe combinations. Seventy percent of the trials were AX trials. Thirty percent were either AY, BY, or BX trials, with each occurring 10% of the time. Critical trials were AY and BX trials. AY trials required participants to shift from proactive to reactive control to respond successfully. This shift occurs due to AY trials leading to the expectation that an AX trial might occur, but then requiring an adjustment upon seeing a non-X probe when the response has to be made. BX trials required the maintenance of proactive control and the disengagement of reactive control to respond successfully. This shift occurs due to the B probe leading to an expectation that the response will be NO but the X cue reactivating the AX response YES. Error rates were used to compute the Behavioral Shift Index (BSI), a continuous metric of proactive and reactive control.[5] The BSI can be computed from reaction times and from error rates (Braver et al., 2010). In the present study, I computed the BSI from error rates because the online format of task delivery was anticipated to introduce more noise for reaction times than for accuracy.

Spanish Noun Phrase Learning[6]

Materials. Thirty-two Spanish nouns of imageable objects and animals were selected as experimental stimuli. Of these, 16 were used for studying and testing knowledge of the studied items and 16 were used for examining generalization. Images for the nouns were primarily taken from the database SoyVisual (https://www.soyvisual.org/) and from fair image searches on the internet.

The selected nouns had a transparent gender marking based on the final phoneme of the word; all of these nouns ending in [a] were feminine and all of the nouns ending in [o] were masculine. A native Spanish speaker was recorded saying the noun phrases for the 32 different nouns. The noun phrases were of inanimate objects,

[5] Formula: AY – BX/[AY + BX] (Braver et al., 2010).

[6] Previous research has found that the effects of ambient language exposure are not always revealed by testing studied items but by generalization to novel items (Bice & Kroll, 2020). Because it was not clear which aspects of learning performance (learning of studied items, generalization to novel items, learning trajectory) would be most likely to reveal effects of ambient language exposure and of cognitive control skills, I created the Noun Phrase Learning task in a way that would allow me to examine these three aspects of learning.

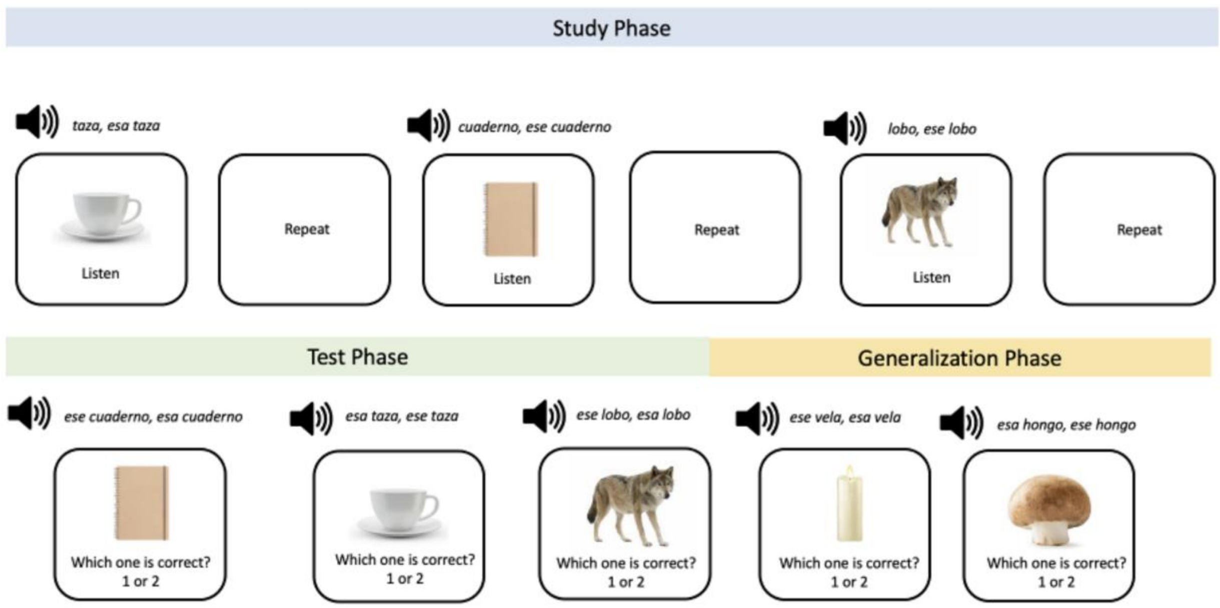
with the exception of five of animals.[7] A native Spanish speaker was recorded while producing the 32 noun phrases in the format: [Noun], that [Noun]. The noun phrases were produced with the gender congruent demonstrative article (ese$_{\text{MASC}}$ gato$_{\text{MASC}}$) and with the gender incongruent demonstrative article (esa$_{\text{FEM}}$ gato$_{\text{MASC}}$). The gender congruent recordings were paired with the pictures for study trials. For test trials, both gender congruent and gender incongruent noun phrases were used (Appendix E).

Procedure. Participants were told at the beginning of the experiment that they would learn noun phrases in Spanish. Two examples of noun phrases with the definite article were provided: one masculine (*ese micrófono/that microphone*) and one feminine (*esa libreta/that notepad*) along with its corresponding image and translation. Participants were told that in Spanish there are masculine and feminine nouns and that the primary way to tell whether a noun is masculine or feminine is whether the noun ends in an [a] or [o]. They were told that based on the final morpheme, a gender congruent article was to be paired with each noun (“Masculine nouns ending in [o] are paired with *ese* and feminine nouns ending in [a] are paired with *esa”)*. They were told that in the learning task, they would encounter noun phrases and study them. The goal was for them to be able to remember which noun goes with which article because they would be tested on that knowledge. They were then informed that they would be doing multiple blocks of studying and testing. The participants had 40 minutes to complete the task on the experimental platform FindingFive (FindingFive Team, 2019). The task involved three study blocks and three test blocks, interleaved. During the study and test blocks, participants studied and were tested on the same sixteen noun phrases. There

[7] Most of the animals selected (cat, duck, wolf, bee, turtle) could not be identified as biologically male or female based on their picture with the exception of the cow.

was one final block of generalization with 16 unstudied items. An overview of the Noun Phrase learning task is shown in Figure 2-1.

Figure 2-1. Noun Phrase Learning task overview.



Language Recognition

Materials. Recordings of native speakers of different languages were obtained from volunteers and from publicly available videos of interviews and shows found on the internet. The recordings were of speakers of the following languages: Spanish, Portuguese, Russian, Tagalog, Hebrew, Arabic, French, Catalan, Malay, Thai, Finnish, Burmese, Hindi, Indonesian, Persian, Vietnamese, Romanian, Cambodian, Tibetan, and Mingrelian. The recordings ranged from 2 to 5 seconds in length (Appendix F). Two measures of language recognition were computed: Spanish recognition d’ and Other language recognition d’. Both scores were signal detection measures of sensitivity (d’).

The d’s were computed by determining the number of hits (trials on which participants correctly answered YES) and false alarms (trials on which participants incorrectly answered YES when the correct answer was NO). The z-transformed difference between Hits, and False Alarms was the d’.[8]

Procedure. Participants were presented with a question (“Is this Japanese?”, “Is this Spanish?”) and simultaneously heard an audio snippet of a person speaking in a language unknown to the participant. They were then asked to answer the question to the best of their ability with a button press (YES/NO). Since they were not familiar with the languages presented, they were told to “make their best guess” as to the language that was spoken in each audio snippet. Participants took ten to twelve minutes to complete the task on FindingFive. They were told to only listen to the audio snippet once unless they did not hear the audio snippet the first time it was presented.[9]

Verbal Fluency

Materials. Two lists of semantic categories were used, one per language. For Mandarin the categories used were: *vegetables, animals, body parts, and family members*. For English, the categories used were: *fruits, colors, school supplies, and musical instruments*. As a language cue, flags of China and of the US were used to indicate that Mandarin or English was to be spoken. Participants’ total number of correct

[8] This measure was intended to be used as a proxy measure for ambient language exposure. However, as we will later see, it turned out not to reflect ambient language exposure. Instead, we will be using self-rated exposure as the measure of ambient language exposure in analyses.
[9] This was done to reduce the amount of time it took to complete the task while allowing the participants to relisten to the audio if there was some noise in their environment.

responses across the four categories for each language were used as a measure of fluency in each language.

Procedure. Participants were instructed to produce as many nouns as possible belonging to a semantic category in the language indicated by the flag and to avoid repeating themselves. The category name was displayed in English for the English trials (along with the US flag) and in Mandarin for the Mandarin trials (along with the flag of China). Participants were instructed to begin speaking as soon as they heard a “ding” sound and saw the word START appear on the screen. They were told that it would be time to stop when they saw the word STOP and heard a second “ding” sound after thirty seconds had elapsed. Audio recordings were coded offline.

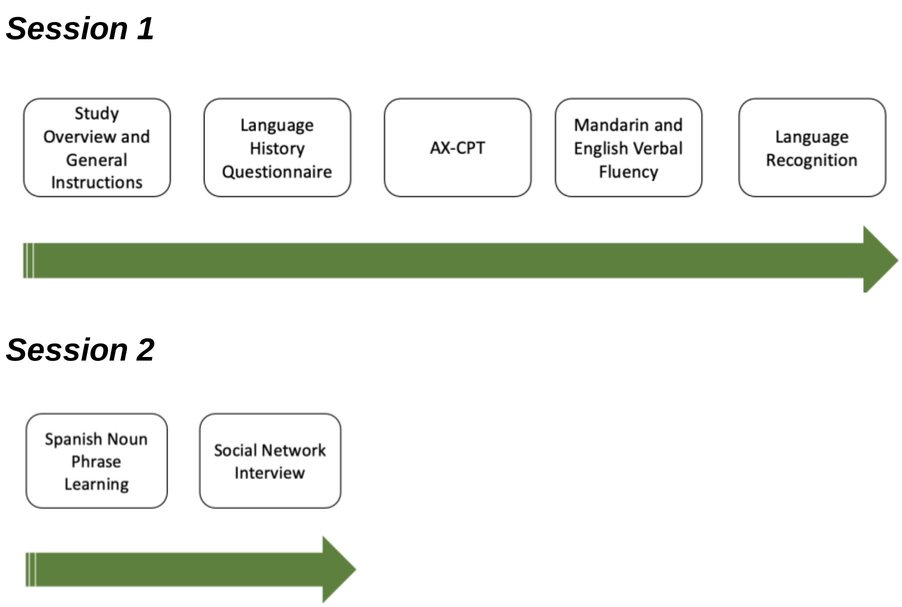
Social Network Survey[10]

Procedure. Participants were asked to list every person with whom they communicate every week for at least 5 minutes using an abbreviation (regardless of whether the communication was via in-person interaction, phone call, video call, texting/messaging, etc). The five minute criterion is based on the social network survey protocol used by Lev-Ari et al. (2018). The participant was then asked to indicate to the best of their knowledge the languages spoken by each person in their social network, their country of origin, the amount of time in hours that the participant communicated with the individuals and how much of their communication occurred in Mandarin, English, or any other language known to the participant. For each participant, the

[10] The social network survey was added to the experimental battery after the experiment was underway as a way of adding another measure of ambient language exposure. Of the sample, 41 participants completed the social network survey.

number of speakers with whom each participant communicated in English, Mandarin, or other languages was computed by adding the number of speakers that were spoken to in English/Mandarin/Other languages 20% of the time or more.

Figure 2-2. Overview of experimental procedure.



General Procedure

Participants completed two experimental sessions virtually on Zoom with an experimenter. The experimenter provided the participant with general instructions (Appendix C) at the beginning of the session and specific instructions for each task were contained within each experimental task. The experimenter gave the participant access to each task as they were completed by the participant. Participants were compensated at the end of each session. At the end of their two sessions, participants were debriefed by email (Appendix D). An overview of each experimental session is shown in Figure 2-2.

Analysis

The first goal of the analysis was to examine the convergent validity of the proxy measure for ambient language exposure: the language recognition task. In Part One, I examined the correlations between self-reported ambient language exposure on the Language History Questionnaire (LHQ) and of social network size on the Social Network Survey (SNS) to scores on the language recognition task. In Parts Two and Three, I examined whether ambient language exposure and cognitive resources predicted two learning outcomes: the learning of studied items and the trajectory of learning over the course of three study blocks. In Part Four, I examined the extent to which ambient language exposure and cognitive resources predicted the extension of knowledge from learned noun phrases to unstudied ones.

Results and Discussion

Part One: Convergent Validity of the Language Recognition Measure

The language recognition measure was developed as a way to capture participants' wholistic knowledge or familiarity with different languages of which they had no formal knowledge. It was hypothesized that the ability to correctly identify a language by an audio snippet spoken by a native speaker would reflect their prior exposure to those languages. In the following analyses, I examined: 1) whether language recognition scores correlated with self-reported measures that reflect ambient language exposure to languages other than Mandarin and 2) whether the self-reported variables that reflect ambient language exposure correlated with one another.

This initial analysis enabled me to evaluate whether Spanish recognition and Other language recognition scores were valid measures of cumulative language exposure and to determine if, alternatively, the self-report measures could be used as indicators of cumulative language exposure. In this analysis, I correlated Spanish d' and Other language d' scores using a set of variables extracted from the LHQ and the SNS.[11]

Table 2-2. LHQ and SNS variables considered in analysis of language exposure.

Task	Variable Name	Description
LHQ	EnglishExposure	Percentage of the time participants are currently exposed to English in the US (scale 0-100)
LHQ	OtherChineseExposure	Percentage of the time participants are exposed to a Chinese language other than Mandarin in the US (scale 0-100)
LHQ	OtherNonChineseExposure	Percentage of the time participants are exposed to languages other than English or any Chinese languages in the US[12] (scale 0-100)
LHQ	NonMandarinExposureChina	How often participants were exposed to languages other than Mandarin when in China (scale 0-3)
SNS	ChineseNetworkSize	The number of people in their social network to whom they speak in Mandarin or any Chinese language in the US
SNS	NonMandarinNetworkSize	The number of people in their social network to whom they speak in Mandarin or any other Chinese language in the US
SNS	NonChineseorEnglishNetworkSize	The number of people in their social network that speak languages other than English/Mandarin/other Chinese languages in the US[13]

[11] Note that the SNS was administered to a subset of the participants from the larger study same (n = 41). Hence, separate regressions were conducted for the LHQ and SNS measures of ambient language exposure. The correlations between LHQ and SNS variables for the smaller sample are shown at the end of this section.
[12] Throughout the paper, I use the term "Chinese languages" to refer to languages that are natively acquired in China and "non-Chinese languages" to refer to languages that are typically not natively acquired in China.
[13] This variable is based on the languages that the participants' reported that people in their social networks knew. This variable thus reflects not the number of languages used by the participants, but potential ambient language exposure from linguistic diversity of the people in their social network.

Spanish Recognition and LHQ Measures

A linear regression was fitted in which Spanish recognition d’ scores were predicted from the four LHQ variables listed above (Table 2-2). None of the LHQ measures significantly predicted Spanish d’ scores (Table 2-3).

Table 2-3. Model summary of linear regression predicting Spanish recognition d’ from LHQ language exposure variables.

Predictors	*Estimates*	*CI*	*P*
(Intercept)	0.00	-0.23 – 0.24	0.974
USEnglishExposure	0.06	-0.19 – 0.31	0.649
USOtherNonChineseExposure	-0.04	-0.32 – 0.23	0.749
USOtherChineseExposure	0.14	-0.14 – 0.41	0.325
ChinaNonMandarinExposure	0.01	-0.24 – 0.26	0.943

Notes. Model formula: Spanish recognition d’ ~ USEnglishExposure + USOtherNonChineseExposure + USOtherChineseExposure + ChinaNonMandarinExposure

Spanish Recognition and SNS Measures

A linear regression was fitted in which the Spanish d’ scores were predicted from the three SNS variables (Table 2-2). The number of total speakers of any Chinese language was predictive of Spanish recognition scores, with a smaller Chinese network size predicting marginally higher Spanish recognition scores (Table 2-4).

Table 2-4. Model summary of linear regression predicting Spanish recognition d’ from SNS variables.

Predictors	*Estimates*	*CI*	*P*
(Intercept)	0.80	-0.23 – 1.82	0.125
ChineseNetworkSize	-0.19	-0.38 – 0.00	0.053
NonMandarinNetworkSize	0.03	-0.12 – 0.17	0.726
NonChineseorEnglishNetworkSize	0.06	-0.19 – 0.31	0.633
ChinaNonMandarinExposure	0.13	-0.22 – 0.49	0.459

Notes. Model formula: Spanish recognition d’ ~ ChineseNetworkSize + NonMandarinNetworkSize + NonChineseorEnglishNetworkSize + ChinaNonMandarinExposure

Other Language Recognition and LHQ Measures

A linear regression was fitted in which Other language recognition d’ scores were predicted from the four language exposure variables from the LHQ (Table 2-2). None of the SNS measures significantly predicted Spanish d’ scores (Table 2-5).

Table 2-5. Model summary of linear regression predicting Other recognition d’ from LHQ language exposure variables.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.01	-0.23 – 0.24	0.951
USEnglishExposure	-0.08	-0.33 – 0.17	0.517
USOtherNonChineseExposure	0.04	-0.24 – 0.31	0.796
USOtherChineseExposure	-0.02	-0.29 – 0.25	0.874
ChinaNonMandarinExposure	0.01	-0.24 – 0.26	0.936

Notes. Model formula: Other recognition d’ ~ USEnglishExposure + USOtherNonChineseExposure + USOtherChineseExposure + ChinaNonMandarinExposure

Other Language Recognition and SNS Measures

A linear regression was fitted in which the Other language recognition d' scores were predicted from the three SNS variables (Table 2-2). None of the SNS measures significantly predicted Other language d' scores (Table 2-6).

Table 2-6. Model summary of linear regression predicting Other d' scores from SNS variables.

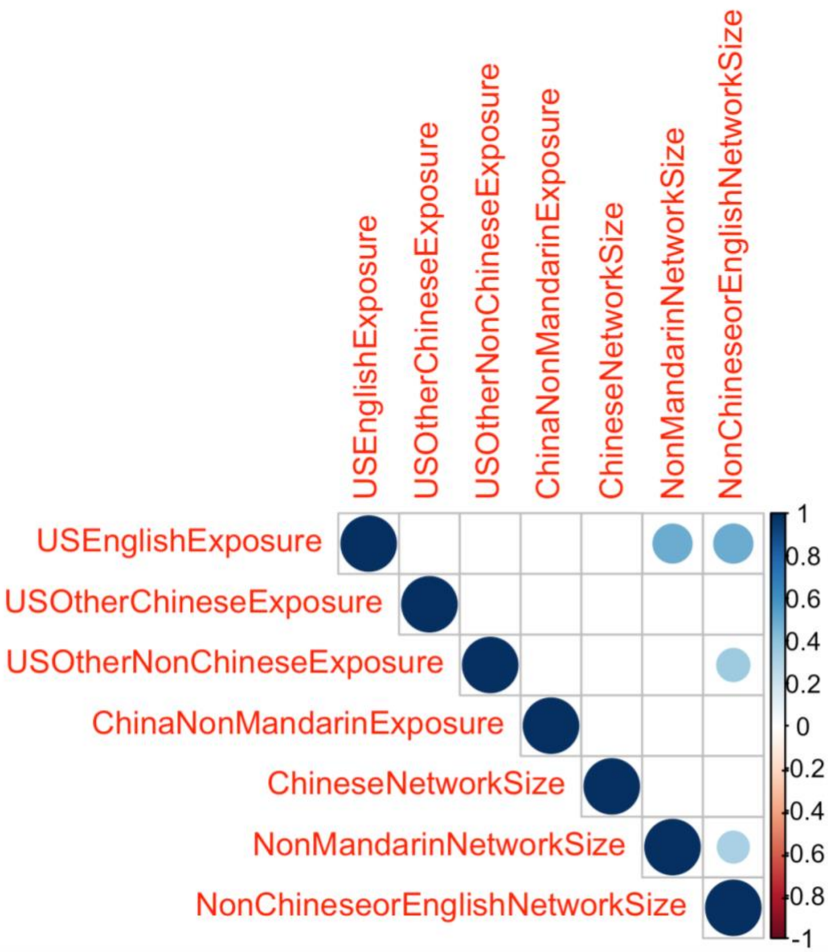
Predictors	*Estimates*	*CI*	*P*
(Intercept)	0.06	-0.31 – 0.43	0.739
ChineseNetworkSize	0.13	-0.25 – 0.50	0.499
NonMandarinNetworkSize	-0.04	-0.44 – 0.36	0.843
NonChineseorEnglishNetworkSize	-0.08	-0.47 – 0.31	0.686
ChinaNonMandarinExposure	-0.00	-0.38 – 0.38	0.994

Notes. Model formula: Other recognition d' ~ ChineseNetworkSize + NonMandarinNetworkSize + NonChineseorEnglishNetworkSize + ChinaNonMandarinExposure

LHQ and SNS measures

The LHQ and SNS measures of ambient language exposure generally did not predict Spanish or Other language recognition scores. The only exception was that a smaller Chinese network size predicted better Spanish recognition. Although this relationship suggests that the Spanish recognition score reflects the diversity of the international students' social network, this relationship was only marginally significant. Thus, the results seem to indicate that language recognition may reflect a form of explicit language knowledge, which may not be particularly related to the kind of implicit language knowledge that is acquired from passive language exposure. Therefore, as an alternative way of measuring ambient language exposure, I considered using the self-report measures from the LHQ. I evaluated the convergent validity of the self-reported measures from the LHQ by examining the relationships between the LHQ and SNS measures.

Figure 2-3. Correlations between LHQ language exposure variables and SNS variables.



Notes. $N = 41$. The correlation coefficient, Pearson's r, is represented by the size and color of the circle. Correlation coefficients (indicated by the circles) are only shown for significant correlations.

The correlation matrix above (Figure 2-3) shows several significant correlations between LHQ and SNS measures that reveal convergent validity. The positive correlation between US English exposure (LHQ) and Non-Mandarin network size (SNS) indicates that international students who reported greater English exposure in the US also routinely interacted with more people who did not speak Mandarin. There was also a significant correlation between US English exposure (LHQ) and Non-Chinese or English network size. This correlation reflects that international students who were exposed to English more frequently knew more people who spoke languages other than Mandarin or other Chinese languages. Thus, the LHQ measure of US English exposure appears to reflect multiple aspects of linguistic diversity in participants' immediate social circle. On one hand, it reflects the number of English speakers and on the other it reflects the number of speakers of English who also spoke other languages.

Section Discussion

In this section, we conducted a set of analyses to determine whether the language recognition measures, Spanish recognition (d') and Other language recognition (d'), seem to broadly reflect language knowledge that has emerged from ambient language exposure. If so, this would justify including the language recognition measures as predictors of Spanish learning in subsequent analyses. As we have seen, the analyses revealed that the language recognition measure correlated with self-rated language use to a very limited extent. One likely reason is that language recognition is a skill that develops not only as a function of language exposure but also of other factors such as world knowledge of languages and geography.

In order evaluate the role of ambient language exposure in learning Spanish grammatical gender, I considered using the self-report measures from the LHQ. To establish their convergent validity, I correlated LHQ and SNS self-report measures. The pattern of correlations confirmed that self-reported language exposure in the LHQ was associated with several social network characteristics. Therefore, to examine the contribution of different sources of language exposure in the following analyses, I used the individual self-reported measures of language exposure from the LHQ because these were collected for all of participants.

Part Two: Spanish Grammatical Gender Learning

In the Noun Phrase learning task, participants alternated between studying a set of noun phrases and being tested on their knowledge of the gender congruency of the definite article and the studied nouns. They then completed a final block in which they were tested on their knowledge of the gender congruency of unstudied items. The research question was whether individual differences in ambient language diversity and in cognitive control (proactivity/reactivity) predict the learning of grammatical gender. I hypothesized that higher levels of ambient language diversity and of proactivity would predict better learning.

The descriptive statistics for the experimental tasks that were examined in the inferential statistics are summarized in the figures below. In the learning task, participants became more accurate across the three blocks and their reaction times became shorter. Average accuracy for the final block of learning and generalization were above chance (50%), although there was variability in learning scores (Figure 2-4).

For the AX-CPT, BSI scores ranged from -1 to 1, with a distribution skewed towards higher BSI scores (Figure 2-5). The distribution indicates that the majority of the BSI scores tended to be in the range between balanced proactivity/reactivity (score of 0) and high proactivity (1). This pattern is consistent with previous studies showing that young adults are generally proactive (Braver et al., 2001; Paxton et al., 2008; Bugg, 2014). The distribution of self-rated exposure to different language categories examined in the LHQ reveals that the international students were largely exposed to English in the US, with some participants reporting exposure to other Non-Mandarin Chinese languages and non-Chinese languages (Figure 2-6A). In terms of the international students' prior exposure to non-Mandarin languages in China, about half of the sample reported no exposure to languages other than Mandarin (Figure 2-6B). The remainder of the sample were evenly distributed across the remaining exposure categories (daily/weekly, monthly, and yearly).

Analysis. I conducted the analyses in the programming environment RStudio (RStudio Team, 2020) using the programming language R (version 4.1.2; R Core Team, 2022). Analyses were conducted using the *lmer* function. The analyses reported here were linear regressions, with both the dependent and independent variables being averages. This approach was chosen over mixed models because the models did not converge well enough to consider the predictors of interest when modeling item-level performance. The only exception were the learning trajectory analyses, which were mixed models used to capture growth curves. As a significance threshold, I used the p-values generated via Satterthwaite method by the *lmer* package. I conducted analyses predicting average accuracy and reaction times in block 3 of testing, which was the final block that tested knowledge of the studied noun phrases.

Figure 2-4. Accuracy and reaction times on the Noun Phrase learning task.

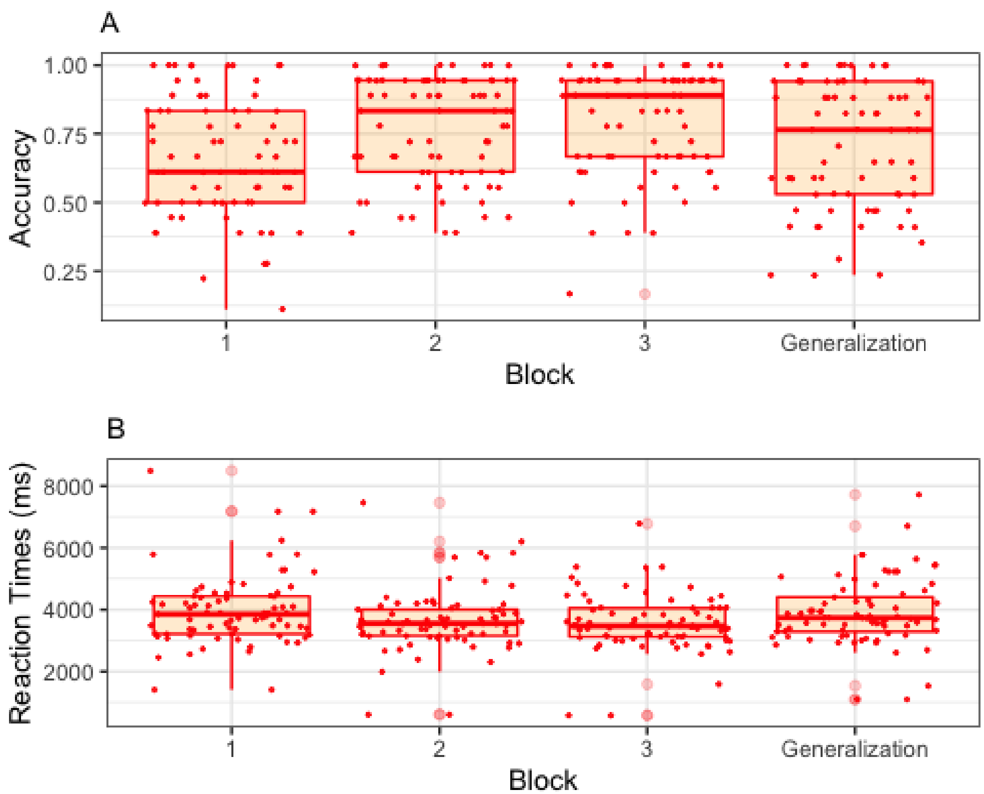


Figure 2-5. Distribution of BSI scores from AX-CPT.

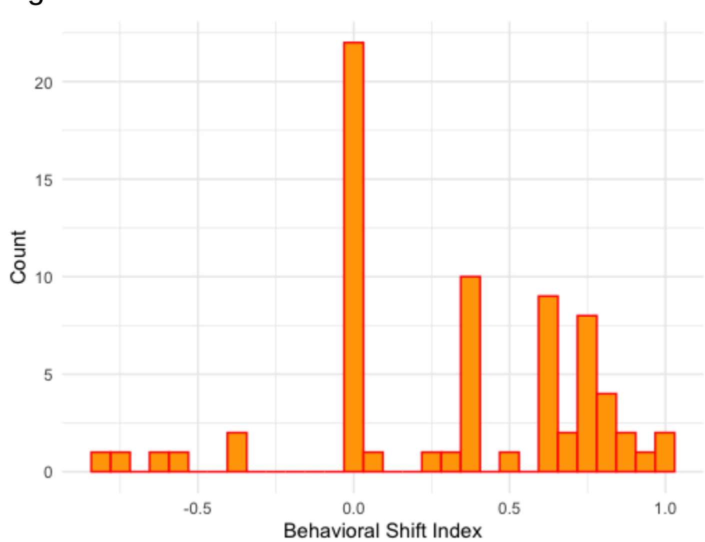
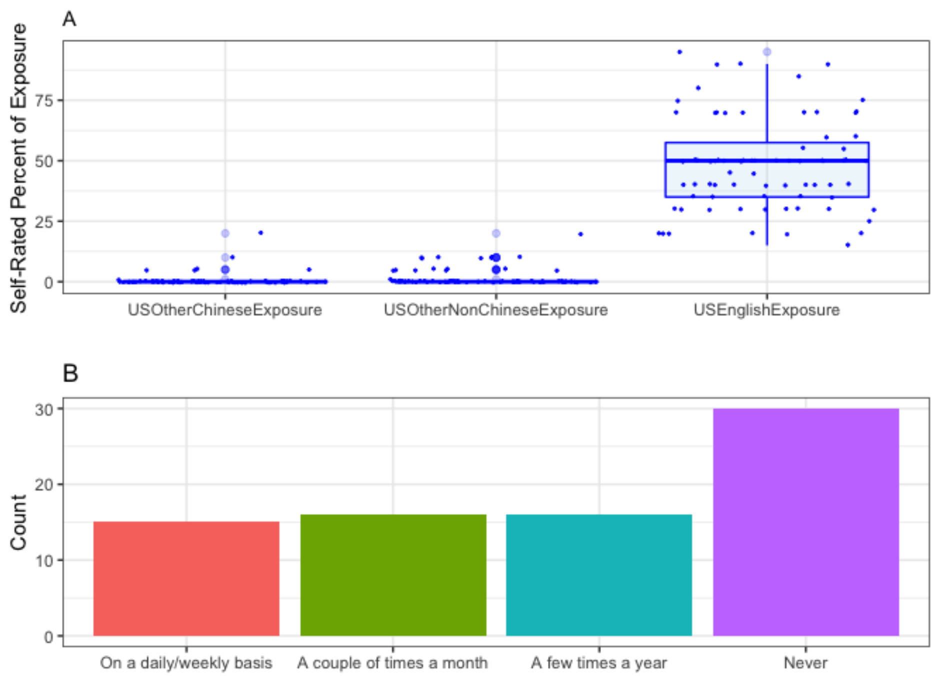


Figure 2-6. (A) Distribution of self-rated exposure percentages to other Chinese languages, other non-Chinese languages, English in the US and (B) Frequency of exposure to non-Mandarin languages in China.



Notes. Exposure to languages in the US and in China are plotted separately due to the scale for the US questions being out of 100, while the question about exposure in China is on a scale from 0 (Never) to 3 (On a daily/weekly basis).

Accuracy and reaction times were for the gender congruency decision (which one is correct? *el gato*/*la gato*). Accuracy and all continuous predictors were z-scored and reaction times were logged for analysis. All figures for models are plotted from model predicted values. For all analyses, I used a model comparison approach to avoid overfitting the data. I incrementally tested whether these five predictors significantly improved model fit: US English Exposure, US Other Chinese Exposure, US Other Non-Chinese Exposure, China Non-Mandarin Exposure and the BSI from the AX-CPT. The

BSI has been used in multiple studies as an individual difference measure of cognitive control (Beatty-Martínez et al., 2020; Zhang et al., 2015; 2021). A higher BSI value indicates greater proactivity, while a lower value indicates higher reactivity. If predictors did not significantly improve model fit over the intercept-only model, they were left out of the final model. The final model is reported in the text.[14]

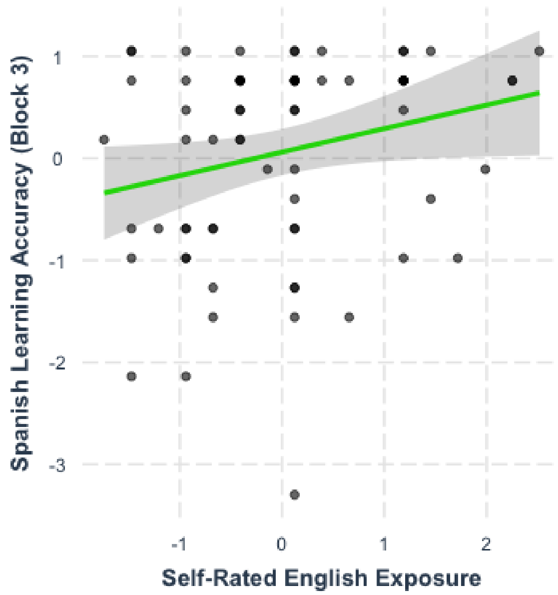
Accuracy

A linear regression was fitted with average learning accuracy in test block three as the dependent variable. The only addition to the intercept-only model that significantly improved model fit was of US English Exposure, $\chi 2 = 2.25$ $p = .02$. The additions of US Other Chinese Exposure ($\chi 2 = .53$, $p = .46$), US Other Non-Chinese Exposure ($\chi 2 = 1.83$, $p = .18$), China Non-Mandarin Exposure ($\chi 2 = .45$, $p = .50$), and BSI ($\chi 2 = .10$, $p = .74$) did not improve model fit. The final model included Non-Mandarin Exposure as a predictor.

The formula for the final model was: Average block 3 accuracy ~ US English exposure. In the final model, there was a significant effect of English exposure on accuracy, with higher self-rated English exposure predicting higher learning accuracy for studied items in block 3, $\beta = .24$, $t = 2.28$, $p = .02$ (Figure 2-7). International students who reported higher exposure to English in the US were also more accurate at making the Spanish grammatical gender decision for the studied noun phrases.

[14] Tables are included for all models except those than only have one significant predictor. In that case, the model will be summarized in the text.

Figure 2-7. Relationship between learning accuracy in test block 3 and self-rated English exposure in the LHQ.



Notes. Learning accuracy and self-rated English exposure values were z-scored.

Reaction Time

A linear regression was fitted with average reaction times in test block three as the dependent variable. The addition of US Other Chinese Exposure ($\chi 2$ = 2.25 p = .02) and BSI ($\chi 2$ = 3.65 p = .06) to the null model marginally improved model fit. A model with these two predictors together was significantly better fitting than the null model, F = 3.96 p = .02. A model in which the two predictors interacted was also significantly better than the one without the interaction, $\chi 2$ = 6.84, p = .01). The additions of US English Exposure ($\chi 2$ = 1.83, p = .18), US Other Non-Chinese Exposure ($\chi 2$ = 1.83, p = .18), and China Non-Mandarin Exposure ($\chi 2$ = .45, p = .50) did not improve model fit. The final model included US Other Chinese Exposure, BSI, and their interaction (Table 2-7).

Table 2-7. Model summary of linear regression predicting learning reaction times in test block 3 from US Other Chinese Exposure and BSI

Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.12	8.05 – 8.19	<0.001
USOtherChineseExposure	-0.16	-0.28 – -0.04	0.008
BSI	-0.12	-0.19 – -0.04	0.004
USOtherChineseExposure *BSI	-0.24	-0.42 – -0.06	0.011

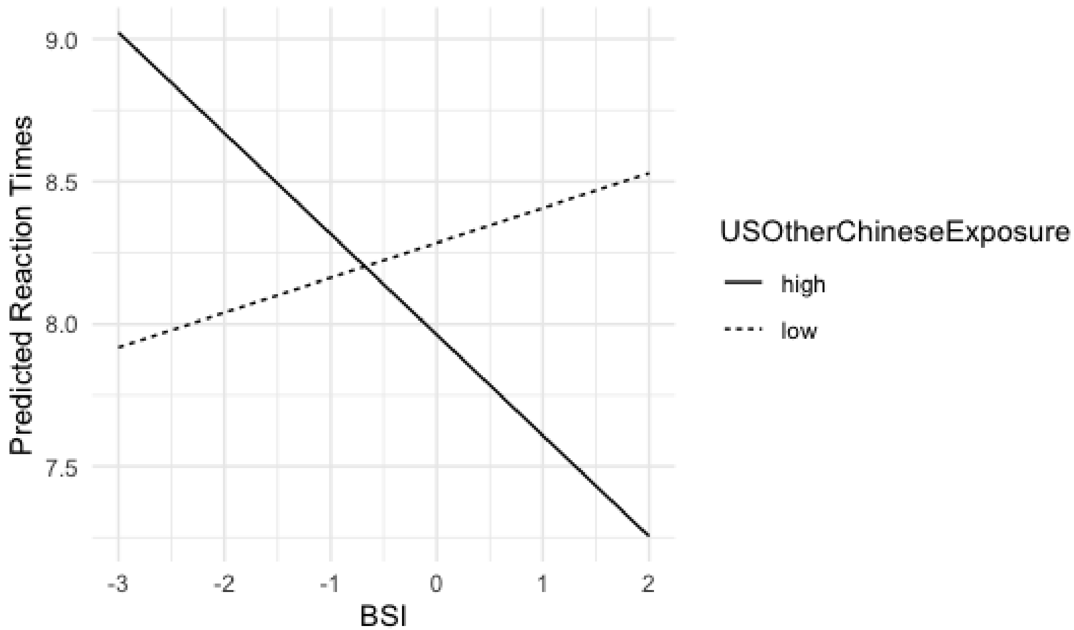
Notes. Model formula: Average block 3 accuracy ~ US Other Chinese Exposure*BSI.

There was a main effect of US Other Chinese Exposure which indicated that international students who reported being exposed to a Chinese language other than Mandarin (e.g Cantonese)[15] more often in the US were faster in making the Spanish grammatical gender decision correctly. There was also a significant effect of BSI, with a higher score (greater proactivity) predicting a faster response. There was a significant interaction between US Other Chinese Exposure and BSI. To probe this interaction, I examined the effect of BSI at high and low levels of Other Chinese exposure.[16] Simple effects revealed that at low levels of Other Chinese exposure, there was no effect of BSI, $\beta = .12$, $t = .12$, $p = .13$. However, at high levels of Other Chinese exposure, a higher BSI (greater proactivity) predicted faster responses for the Spanish gender decision, $\beta = -.35$, $t = -3.12$ $p = .002$. The interaction is illustrated below (Figure 2-8). This interaction suggests that for bilinguals who were actively using an additional

[15] In our sample, most of the Other Chinese exposure reported was of Cantonese, but it also included data from four other participants who reported exposure to Shanghainese and Taiwanese in the US.
[16] High and low levels of other Chinese exposure were operationalized as 1 SD below and above the mean. For all future interactions reported, 1 SD below and above the mean will be used as high and low values.

Chinese language (on top of Mandarin) in the US, the recruitment of proactive control resources seemed to benefit the speed of answering the gender congruency questions.

Figure 2-8. Interaction between US Other Chinese Exposure and BSI.



Notes. Predicted reaction times are the model predicted log reaction times.

Section Discussion

In this section, we conducted a set of analyses to examine how individual differences in exposure to different languages and cognitive control predict the learning of grammatical gender in Spanish. In these analyses we specifically examined accuracy and reaction times in the final test block for studied items. The accuracy analyses revealed that increased exposure to English in the US was predictive of higher learning accuracy in Spanish at the end of the learning task. This result seems to indicate that despite all of the international students being immersed in a broader environment in which English was the primary language, international students who were more exposed to English in their daily lives were also able to make the Spanish gender congruency decision faster. As we saw in Part 1, higher self-rated exposure to English in the LHQ was associated with greater social network diversity, particularly with having more social network members who were not speakers of Chinese languages and with having more social network members who spoke languages other than English. These results therefore suggest that participants who used English more frequently were also more likely to encounter different languages in their immediate social circle, which in turn may have contributed to their sensitivity to the novel grammatical gender pattern in Spanish. Importantly, English exposure did not interact with performance on the cognitive control measure (BSI), consistent with a direct contribution between language exposure to the learning of gender in Spanish.

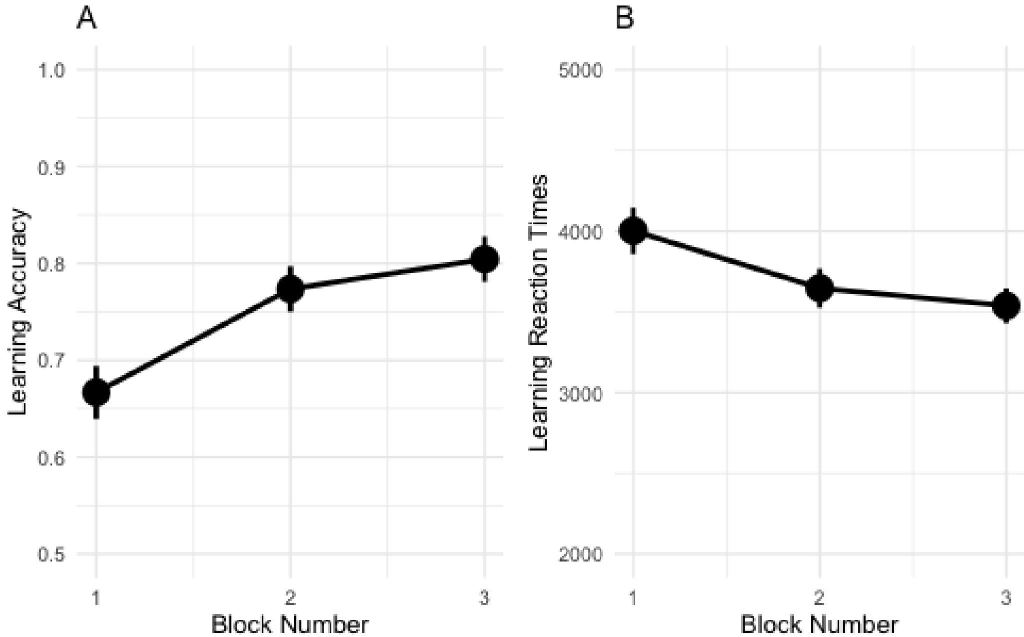
In the reaction time analyses, there was a significant interaction between degree of exposure to Other non-Mandarin Chinese languages and the BSI score. For international students who were more exposed to a non-Mandarin Chinese language in the US (e.g. Cantonese), a higher BSI (greater proactivity) predicted faster performance in the Spanish learning task. For international students who had a low degree of exposure to a non-Mandarin Chinese language in the US, there was no effect of BSI on Spanish scores. This indicates that international students that spoke an additional Chinese language other than Mandarin, greater proactivity was associated with a faster response. For international students who did not know an additional Chinese language other than Mandarin, proactivity was unrelated to reaction times for responding to the final Spanish test. This result is the first indication that individual differences in the language environment of the international students modulated the relationship between cognitive control and Spanish learning. Although this result is consistent with the general prediction that the immersion environment affects the relationship between cognitive control recruitment and learning, it was an unexpected finding.

Why might using Cantonese or a non-Mandarin Chinese language be associated with reliance on proactive control for learning? The bilinguals who reported active use of an additional Chinese language were L1 speakers of Cantonese, L2 speakers of Mandarin, and L3 speakers of English while those who reported low use of another Chinese language were L1 speakers of Mandarin and L2 speakers of English. While Mandarin use tends to be highly prevalent in Chinese international student circles, the number of speakers of other Chinese languages is much smaller. Thus, native speakers of non-Mandarin Chinese languages (e.g. Cantonese) are in a sense more fully L2-immersed because they do not encounter as many people with whom they can speak their L1 as regularly. They are immersed in English, the predominant language in the larger environment, and to a lesser extent in Mandarin, the predominant language in Chinese international student circles. The finding that for the students with high exposure to Chinese languages other than Mandarin proactivity predicted learning outcomes echoes the finding from the language processing literature that bilinguals immersed in their nondominant language recruit proactive control to keep their suppressed dominant language active (Beatty-Martínez et al., 2020; Zhang et al., 2021). This finding therefore suggests that there is a parallel between how cognitive control is recruited for language processing and new language learning. The component of control that is recruited by bilinguals to process their known languages in their particular interactional context is also the same component of cognitive control that predicted better learning.

Part Three: Spanish Grammatical Gender Learning Trajectory

In this analysis, accuracy and reaction time data from the three test blocks were modeled using a growth curve analysis. The accuracy and reaction time trajectories for these two variables across the three blocks can be seen in Figure 2-9. The learning trajectories were modeled using second order orthogonal polynomial contrasts (linear and quadratic) in a mixed model using the *lmer* function. To the model with the polynomial terms, I added the same four LHQ language exposure predictors and the BSI as fixed effects. These predictors were allowed to interact with the linear and quadratic terms. Their contribution to model fit was evaluated and served as the basis for inclusion/exclusion from the final model. If there were convergence issues with the predictors when allowed to interact, we instead included them as non-interacting terms and evaluated improvement in model fit that way. The maximal random effects structure including the polynomial terms and any significant individual difference predictors was also attempted. In cases where there was a convergence failure, the random effects structure was first simplified by removing the correlation between terms. If convergence failure continued, the predictor causing the convergence failure was excluded.

Figure 2-9. (A) Spanish learning accuracy and (B) reaction time trajectories for three test blocks.



Accuracy

A regression was fitted in which the dependent variable was the average learning accuracy for each block. An initial model revealed significant model fit improvement after adding the linear and quadratic time terms into the model, $F = 43.44$, $p < .001$. Following the model comparison procedure outlined above, the only additional fixed effect that improved model fit was US English exposure when allowed to interact with the polynomial terms, $F = 9.38$, $p = .02$. The final model also included by-participant random slopes for the time terms (Table 2-8).

There was a main effect of US English exposure on accuracy overall, with international students that had higher English exposure also having higher learning scores. There was no interaction between English exposure and the linear or quadratic terms, suggesting that the rate of learning did not depend on English exposure.

Table 2-8. Model summary of polynomial regression predicting learning accuracy trajectories from Block and English exposure

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.04	-0.15 – 0.24	0.660
Linear	3.70	2.62 – 4.78	<0.001
Quadratic	-1.17	-2.25 – -0.09	0.034
USEnglishExposure	0.25	0.05 – 0.45	0.014
Linear* USEnglishExposure	0.04	-1.01 – 1.10	0.935
Quadratic* USEnglishExposure	-0.64	-1.69 – 0.42	0.234

Notes. Model formula: Average Accuracy (for a given block) ~ LinearTerm + QuadraticTerm + USEnglish Exposure + LinearTerm*USEnglishExposure + QuadraticTerm*US English Exposure +(Block|Participant)

Reaction Times

A regression was fitted in which the dependent variable was the average of the log reaction times for each block. An initial model revealed significant model fit improvement after adding the linear and quadratic terms into the model, $F = 34.91$, $p < .001$. The only additional fixed effect that improved model fit was US Other non-Chinese exposure when allowed to interact with the polynomial terms, $F = 57.89$, $p < .001$. The final model is presented below (Table 2-9).

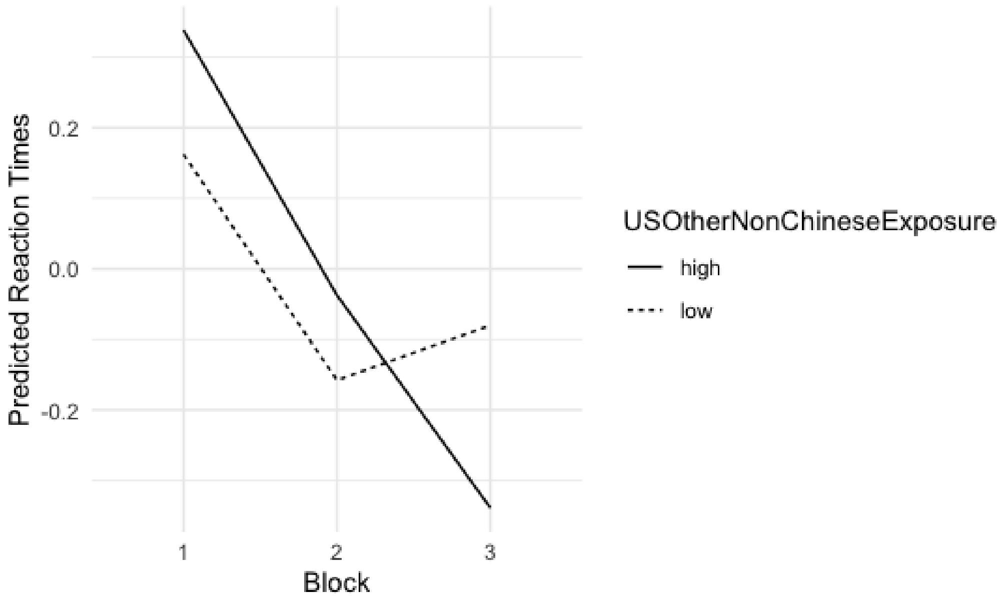
Results show that there was a significant linear decline in reaction times over the course of three blocks of testing. Moreover, there was a significant interaction between the linear time term and US Other non-Chinese exposure. To probe this interaction, I examined the effect of the linear time term at high and low levels of Other non-Chinese exposure. For international students who reported high levels of exposure to languages that were not native to China and to English, the linear decrease in reaction times over the course of three blocks was more pronounced (β = -3.86, t = -5.90, p < .001) than for international students who were minimally exposed to other languages, β = -1.38, t = -2.18, p = .03. This interaction is illustrated in Figure 2-10.

Table 2-9. Model summary of polynomial regression predicting reaction times trajectories from Block and US Other Non-Chinese exposure

Predictors	*Estimates*	*CI*	*p*
(Intercept)	-0.02	-0.25 – 0.21	0.871
Linear	-2.62	-3.45 – -1.79	<0.001
Quadratic	0.78	-0.05 – 1.61	0.067
USOtherNonChineseExposure	0.01	-0.27 – 0.28	0.964
Linear*USOtherNonChineseExposure	-1.24	-2.20 – -0.28	0.011
Quadratic*USOtherNonChineseExposure	-0.53	-1.49 – 0.43	0.274

Notes. Model formula: Average reaction times (for a given block) ~ LinearTerm + QuadraticTerm + USOtherNonChineseExposure + LinearTerm* USOtherNonChineseExposure + QuadraticTerm*USOtherNonChineseExposure +(Block|Participant)

Figure 2-10. Interaction between US Other Non-Chinese Exposure and Linear time term for Block.



Notes. Predicted reaction times are the model predicted log reaction times.

Section Discussion

In this section, I conducted a set of analyses that examined whether individual differences in exposure to different languages and cognitive control predicted the trajectory of learning of grammatical gender in Spanish. I specifically examined the trajectory of accuracy and reaction times over the course of three test blocks. The accuracy analyses revealed that higher exposure to English was predictive of higher learning accuracy overall, although exposure to English did not predict different learning trajectories. These results are consistent with prior analyses from Part 2 which showed that higher levels of English exposure predicted better learning accuracy in block 3.

The reaction time analyses showed that the reaction time trajectory over the course of three test blocks depended on the level of exposure to languages that were not English or any Chinese languages. International students who reported higher exposure to new languages showed a consistent and more accelerated decline in reaction times over the course of the three test blocks whereas international students who had low levels of exposure to unknown languages showed a decline that began to reverse in the final test block. The reversal of reaction times in the final test block suggests that the international students in a less diverse language environment were beginning to fatigue during the final test block, and required more time to correctly respond the gender congruency question.

Taken together, the results of the trajectory analyses indicate that increased English exposure and increased exposure to Other Non-English and Non-Chinese languages predicted better learning outcomes, although only exposure to Other Non-English and Chinese languages predicted a different learning trajectory. There appeared to be no contribution of cognitive control resources for the trajectory of learning.

In the next section, I conducted a set of analyses that examined the learners' ability to extend their knowledge of the patterns learned from studied noun phrases to new unstudied noun phrases. Performance on test block 4, which tested this extension of knowledge to unstudied items, will be referred to as generalization. Generalization was hypothesized to reveal more general knowledge of grammatical gender, as it does not test knowledge of grammatical gender using the same items that were included in the study phases. In addition to considering the self-reported measures of language exposure and the BSI score that were included in prior analyses, learning accuracy from block 3 was also included as a predictor.

Part Four: Spanish Grammatical Gender Generalization

Accuracy

A regression was fitted in which the dependent variable was the average generalization accuracy for the final test block. The addition of learning accuracy in block 3 ($F = 38.04$ $p < .001$) and US Other Chinese Exposure ($F = 16.04$ $p < .001$) to the intercept-only model significantly improved model fit. The interaction between learning accuracy and Other Chinese Exposure significantly improved model fit over the model without the interaction, $F = 4.75$ $p = .01$. The final model included learning accuracy and Other Chinese Exposure as predictors (Table 2-10).

Table 2-10. Model summary of linear regression predicting generalization accuracy from Block 3 learning accuracy and US other Chinese exposure.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	-0.15	-0.37 – 0.06	0.149
NP LearningACC3	0.82	0.58 – 1.06	<0.001
USOtherChineseExposure	-0.31	-0.70 – 0.09	0.126
NP LearningACC3* USOtherChineseExposure	0.85	0.21 – 1.49	0.010

Notes. NP Learning ACC3 = Noun Phrase Learning Accuracy in Block 3; Model formula: Average generalization accuracy ~ NP Learning ACC3 + US Other Chinese Exposure + NP Learning ACC3*US Other Chinese Exposure

There was a main effect of learning accuracy, with international students who were more accurate in the final test block for studied items also having higher generalization scores. There was a significant interaction between US Other Chinese language exposure and learning score in block 3. To probe this interaction, I examined the role of US Other Chinese exposure at high and low levels of learning accuracy. Simple effects revealed that for less accurate learners, higher use of another non-Mandarin Chinese language predicted lower generalization accuracy, $\beta = -1.15$, $t = -2.52$, $p = .01$. For highly accurate learners, the effect of another non-Mandarin Chinese language exposure predicted marginally higher generalization scores, $\beta = .54$, $t = 1.99$, $p = .05$.

Reaction Time

The addition of US Other Chinese Exposure ($F = 2.25$ $p = .02$) and BSI ($F = 3.65$ $p = .06$) to the null model marginally improved model fit. A model with these two predictors together was significantly better fitting than the null model, $F = 3.96$ $p = .02$. A model in which the two predictors interacted was also significantly better than the one without the interaction, $F = 6.84$, $p = .01$. The additions of Learning accuracy in block 3 ($F = .21$, $p = .64$), US English Exposure ($F = 1.83$, $p = .18$), US Other Non-Chinese Exposure ($F = 1.83$, $p = .18$), and China Non-Mandarin Exposure ($F = .45$, $p = .50$) did not improve model fit. The final model included US Other Chinese Exposure, BSI, and their interaction (Table 2-11).

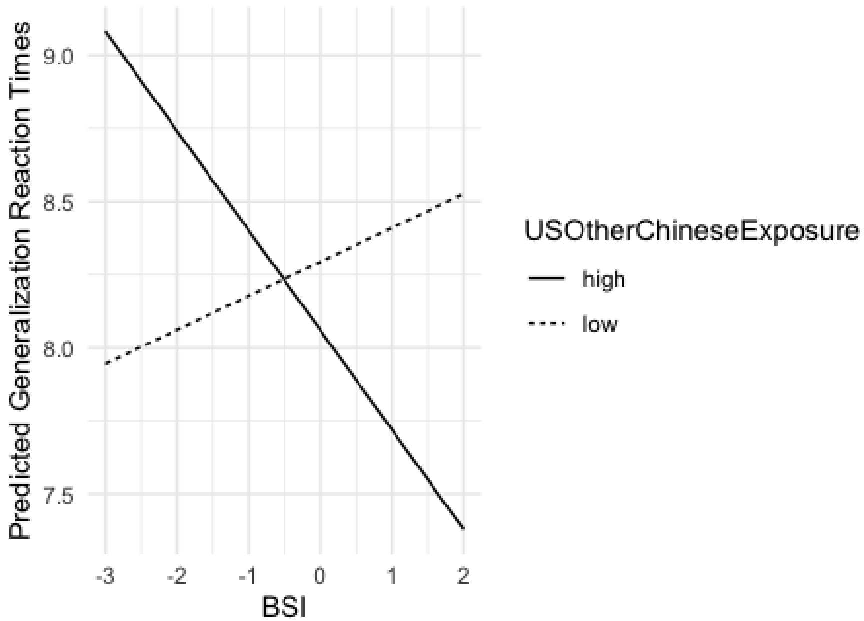
There was a main effect of BSI, such that participants who were more proactive also responded faster to the gender congruency question for unstudied items. There was also a main effect of US Other Chinese exposure, such that learners who had more exposure to non-Mandarin Chinese languages also responded faster. The interaction between BSI and Other Chinese exposure was also significant. This interaction was broken down by examining the effect of BSI at high and low levels of Other Chinese exposure. At high levels of Other Chinese exposure, increased BSI (more proactivity) predicted faster generalization reaction times, $\beta = -.34$, $t = 5.24$, $p < .001$. At low levels of Other Chinese exposure, increased BSI (more proactivity) predicted an increase in reaction times, $\beta = .11$, $t = 2.48$, $p = .01$. The interaction is illustrated in Figure 2-11.

Table 2-11. Model summary of linear regression predicting generalization reaction times from US other Chinese exposure and BSI.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	8.18	8.14 – 8.22	**<0.001**
BSI	-0.11	-0.16 – -0.07	**<0.001**
USOtherChineseExposure	-0.12	-0.18 – -0.05	**0.001**
BSI*USOtherChineseExposure	-0.23	-0.33 – -0.13	**<0.001**

Notes. Model formula: Average generalization reaction time ~ BSI + USOtherChineseExposure + BSI*USOtherChineseExposure

Figure 2-11. Interaction between US Other Chinese exposure and BSI.



Section Discussion

Results for the accuracy analysis showed that learning accuracy in the final test block for studied items was the best predictor of generalization performance. This suggests that the efficiency of the learning process is largely captured by how well learners acquired grammatical gender knowledge of the specific noun phrases. At the same time, there was evidence that prior language knowledge also played a role in generalization ability.

International students who had high learning accuracy seemed to benefit from exposure to an additional Chinese language. For those who had low learning accuracy, however, greater exposure to an additional Chinese language predicted lower generalization accuracy. This result could suggest that greater linguistic diversity helps those who are good at learning to extend their knowledge to new noun phrases. For struggling learners, active exposure to a third language seems to predict worse generalization of grammatical in Spanish. It is unclear why higher use of another Chinese language predicts *lower* generalization accuracy in struggling learners.

The reaction time analysis showed that reaction times were predicted by both cognitive control resources and exposure to another Chinese language. For learners who had high levels of exposure to another Chinese language, higher proactivity was predictive of better generalization scores. For learners who had low levels of exposure to another Chinese language, higher reactivity was predictive of better generalization scores. These results echo the results from the learning reaction time analysis, which also showed that the learners who were more exposed to another Chinese language in their environment responded faster if they were more proactive control whereas bilinguals who had less exposure to another Chinese language did not benefit from proactivity. As previously explained, this interaction between exposure to another Chinese language and the BSI may reflect differential recruitment of cognitive resources due to a different type of language immersion. The international students who were exposed another Chinese language were in a sense more deeply immersed in their nondominant languages (English and Mandarin) than the Mandarin-English bilinguals in the sample, who had easier access to their L1 (Mandarin) in their international student Chinese community. The key result was that for the international students who were exposed another Chinese language, having better proactive control was predictive of faster generalization. This result suggests that proactive control resources may have been involved in learning of grammatical gender for L1 speakers of another Chinese language. If so, this pattern would be consistent with the international students who knew another Chinese language recruiting proactive control for learning, which is the component of cognitive control that has been implicated in language regulation for L2-immersed bilinguals. On the other hand, for the learners who were not exposed to another Chinese language, having more reactive control was predictive of faster generalization. This finding suggests that for participants who only knew Mandarin and English, reactive control resources may have been recruited for generalizing Spanish grammatical gender. Although an unexpected result, this pattern is consistent with the recruitment of reactive control for L1-immersed bilinguals. The fact that the Mandarin and English speakers recruited reactive control for learning might reflect that they were not as deeply immersed in their L2 as the international students who were exposed another Chinese language. The former had more opportunities to speak their L1 (Mandarin) than the latter (Cantonese). Thus, even though both groups were L2-immersed by virtue of studying abroad in the US, these results suggest that the different interactional contexts for L1 speakers of Mandarin versus L1 speakers of another Chinese language in the US shaped how they recruited cognitive control for learning.

General Discussion

The present study investigated the learning of Spanish grammatical gender by Chinese international students. The goal of the study was to uncover whether individual differences in ambient language exposure to both known and unknown languages and cognitive control preferences were predictive of Spanish learning outcomes. There were two general hypotheses. The first was that the relationship between of cognitive control and language learning would differ as a function of the language environment. This prediction follows from the research on language processing, which shows that bilinguals draw upon different cognitive resources depending on their language environment and its demands (Green & Abutalebi, 2013). The second hypothesis was that ambient language exposure may directly contribute to new language learning ability.

Did the relationship between of cognitive control and Spanish learning differ as a function of ambient language exposure?

There were multiple findings that illustrated different relationships between cognitive control and Spanish learning as a function of language exposure. For both learning and generalization reaction times, frequency of exposure to a Chinese language other than Mandarin modulated the relationship between cognitive control and learning outcomes. International students who were regularly exposed to another (non-Mandarin) Chinese language seemed to benefit from proactive control while those who were minimally exposed to a non-Mandarin Chinese language seemed to benefit from reactive control for answering learning and generalization questions more quickly.

Since the larger sample of international students contained a small number of native speakers of other Chinese languages (n = 5), such interactions should be interpreted with caution. However, if this pattern is robust, it would suggest that the nature of the immersion of the international students leads to differential cognitive control recruitment for language processing, which may in turn hold consequences for how cognitive control is engaged in new language learning. To test this idea more rigorously, more research will be needed that addresses the comparison between immersed bilinguals whose native languages differ in prevalence in their immersion environment.

Did ambient language exposure contribute to new Spanish learning?

There was evidence that multiple sources of language exposure predicted Spanish learning outcomes. First, higher exposure to English was associated with more accurate learning at the final test block for studied items. Second, higher exposure to languages other than English, Mandarin or other Chinese languages was also associated with a more consistent decrease in reaction times throughout the three test blocks. These findings suggest that even though the international students were all immersed in English in their broader language environment, the degree to which their own social circle includes people with whom English is used regularly and who know other languages seems to contribute to their language learning potential.

An important implication of the present study is that some aspects of the ambient language experience seem to emerge from how learners' prior language knowledge shapes their interaction with their current language environment and some seem to be primarily driven by the diversity available in their current language environment. Without having characterized some of the learners as trilinguals by assessing their language experience in China, these speakers would have appeared to have a more diverse language environment in the US. Thus, by measuring only current ambient language exposure without appropriately characterizing formal language knowledge, the source of language experience that contributes to language learning can be misidentified. It is possible that this may have happened in studies where participants who were thought to be naïve overhearers might in fact have had some existing receptive language skills of the overheard language (Au et al., 2002; Knightly et al., 2003).

While the self-report measures of language exposure used in this experiment were not objective measures of the language environment, they have had their own advantages. Estimating language diversity by using demographic information for a given geographic area is not well-suited to capturing individual differences because in order to be accurate, one must assume that the person lives primarily that region and has limited contact with people that are outside of that region. Due to technological advances in communication, transportation, and media, it is very unlikely that demographic estimates capture the language environment of all individuals living in a given region accurately. Another advantage relative to other objective measures (e.g. LENA or the EAR) is that self-report measures may capture language exposure that is infrequent or distant from the experimental session. Although recordings of the language environment may be more objective in terms of quantifying language input, they only capture the very immediate language environment of the participant. Thus, by using self-report measures, accuracy in the amount of input may have been lost but sensitivity to cumulative language experience gained.

What kind of language knowledge can be gained from hearing people speak foreign languages?

Recent studies suggest that passive language exposure can lead to implicit knowledge of phonotactics (Oh et al., 2020), phonetic contrasts (Knightly et al., 2003) of the languages to which one has been exposed. However, the findings of this study suggest there is a more general contribution of varied language exposure than learning particular features or particular words. In this study, different sources of ambient language exposure predicted the learning of grammatical gender in Spanish. Increased exposure to English in the international students' environment, their nondominant language predicted better learning of Spanish.

The present study suggests that individuals who are studying abroad and are willing to become more deeply immersed in their non-dominant language are better language learners of a new language. This could be because by intentionally exposing themselves to their nondominant language (English), they learn to adapt to the demands of their language environment, which involves suppressing their dominant language. The ability to suppress the dominant language may be critical for learning to attend to a different set of linguistic features and for learning new information (Bogulski et al., 2019). Therefore, the contribution of exposure to the non-dominant language for new language learning would seem to originate from the shared demands of language processing and language learning.

At the same time, the degree of English exposure the international students reported having in the US was also associated with having more diverse social networks. This finding suggests that the international students' social circle and the linguistic diversity it afforded, also contributed to the learning. Hearing different languages spoken by their friends and family and also hearing different varieties of English may have contributed to the learners' sensitivity to new language patterns. Previous research suggests that there are individuals who are able to take advantage of highly variability when learning new phonetic distinctions while others are not

(Perrachione et al., 2011). The individuals who are able to learn from highly variable environments have been found to have neuroanatomical predispositions that support more detailed auditory processing (Wong et al., 2008). Although such perceptual predispositions could be innate, it is also possible that a more sensitive auditory system is the result of experience-induced plasticity. Just as hearing more speaker variation in one's language environment may improve speech perception (Lev-Ari, 2018), living in a more diverse language environment may also induce such perceptual changes in the learner, making them more attuned to new language patterns. Alternatively, we must consider the possibility that people who have more fine-grained perceptual abilities may be more likely to pursue more linguistically diverse environments, and to be better language learners by recruiting certain cognitive resources. Although these possibilities are not mutually exclusive, more research will be needed to investigate precisely how language experience affects the language learner.

CHAPTER 3:

Investigating the Impact of Native Tone Language Knowledge, Musical Experience, and Cognitive Control for New Tone Learning.

Early exposure to two languages has been found to have enduring consequences for language processing and language learning. Infants come into the world able to perceive the speech sounds of different languages, even languages to which they have not been exposed. In a seminal study, Werker and Tees (1984) found that this language-universal perception declines by the end of the first year. At that point, children have tuned to the language(s) to which they were exposed and appear to have lost sensitivity to languages to which they had not been exposed. This early tuning to the native language (L1) has been conceptualized as a perceptual “magnet” because all new speech sounds come to be perceived through existing speech categories (Kuhl, 1993; Kuhl et al., 2008). According to theories of second language speech perception, new speech sounds may be perceived as L1 sounds or assimilated into existing L1 categories if they are similar, while sounds that are dissimilar from those found in the L1 may be perceived as new categories (Flege, 1987; Tyler & Best, 2007).

There is extensive evidence that early perceptual tuning to the L1 is enduring and that there may be a privileged role of early language experience in paving the way for future language learning. The enduring consequences of early language exposure are perhaps most evident in individuals who had an initial language exposure that was permanently discontinued. Pierce et al. (2014) examined the brain activity of three groups of adolescents living in Canada while listening to Mandarin sentences: Mandarin-French bilinguals who were actively using both languages, international adoptees who as infants had been born in China and learned French when they were adopted into Canadian families, and French monolinguals who had been born and raised in France. On average, the adoptees lived in China until about age three and discontinued Mandarin exposure permanently since immigration to Canada, becoming functionally monolingual in French.

Participants were asked to listen to auditory stimuli that were three syllable sequences of pseudowords that were either hummed or produced normally. All sentences contained tonal information and participants were asked to press a button to indicate if the final syllable of each sentence was the same or different. The results showed that the Mandarin-English bilinguals and the adoptees both activated predominantly the left Superior Temporal Gyrus (STG) and to a lesser extent the anterior STG in the right hemisphere.[17] In contrast, a control group of French monolinguals only activated the right STG, which is consistent with processing Mandarin tones as nonlinguistic auditory stimuli. The results of the study suggest that despite the discontinued use of their native language (Mandarin), the adoptees' early experience with Mandarin resulted native processing of tone as a linguistic feature.

Another population that reveals the continued influence of early L1 tuning for language learning in adulthood are childhood overhearers. Childhood overhearers are adults who as children were exposed to a language other than their native language but never learned to speak that language. Au et al. (2002) asked whether having overheard a second language in childhood might confer befits to learning a second language as an adult. They examined three aspects of language learning: the pronunciation of voiceless

[17] The STG is an area that contains the auditory cortex and Wernicke's area.

stops, the lenition of stops, and the production of number and gender agreement. This study compared functionally English monolinguals who had overheard Spanish as children, English monolinguals who had not overheard Spanish, and Spanish-English bilinguals' performance in these aspects of Spanish.

Since Spanish has a shorter Voice Onset Time (VOT)[18], if the overhearers experienced any benefits from their early overhearing exposure, they were expected produce Spanish VOTs shorter than those of English monolinguals and comparable to those of Spanish-English bilinguals. Moreover, since Spanish voiced stops are lenited in intervocalic contexts, if the overhearers experienced a benefit due to their early exposure, they would also produce a higher percentage of lenited voiced stops in word medial position relative to word initial position. The results showed that the childhood overhearers produced VOTs shorter than those of English monolinguals and more similar to those of proficient Spanish-English bilinguals. They also lenited voiced stops in intervocalic position at a similar rate to the proficient Spanish-English bilinguals. However, in a grammaticality judgement task that assessed sensitivity to grammatical gender and number, the overhearers were less accurate than the Spanish-English bilinguals and performed similarly to the English monolinguals. These results suggest that if early language has enduring consequences, it is most likely for the perception of speech sounds, not for morphosyntax. A follow-up study with more controlled materials reached the same conclusion (Knightly et al., 2003).

The evidence reviewed is consistent with a version of a critical period of speech perception early in life in which the L1 paves the way for second language (L2) speech

[18] This is an acoustic measure of the time between the burst of a stop and the onset of voicing.

perception. As we will see in the next sections, in some cases, L1 tuning can be facilitatory (Chang & Mishler, 2012; Bohn & Best, 2012) while in others it can translate to insensitivity to L2 speech categories (Iverson et al., 2003; Yamada, 1995).

L1 interference in L2 perception

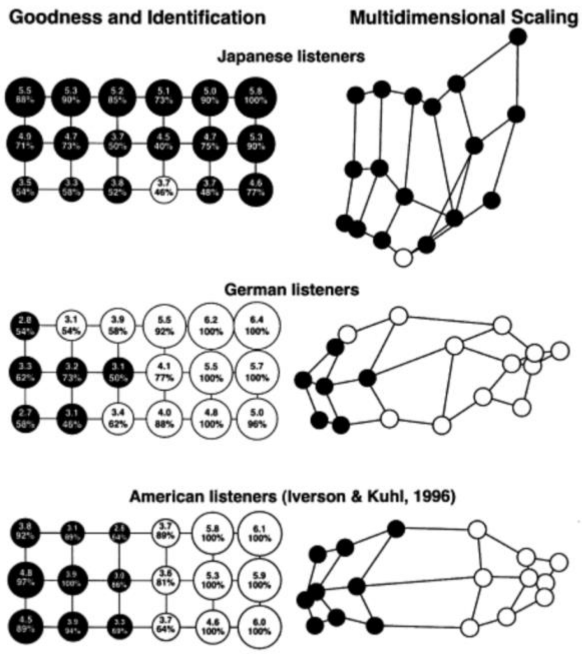
Perhaps the most striking example of the influence of the early language experience for the ability to perceive non-native contrasts is the acquisition of the English /r/ and /l/ by native speakers of Japanese. In Japanese, there is a single approximant category, with phonetic realizations that are generally alveolar taps but also include lateral approximants. Whether an alveolar tap or a lateral approximant is produced does not affect meaning in Japanese. In English, approximants and lateral approximants are contrastive (the words *fir* and *fill* have different meanings). Studies that have investigated the perception of the English contrast between /r/ and /l/ by native Japanese speakers find evidence native Japanese speakers assimilate the two English approximants into their one native Japanese approximant category and are unable to distinguish between the two English approximants categorically (Goto, 1971; Iverson et al., 2013; Yamada & Tohkura, 1992).

Iverson et al. (2013) compared perceptual discrimination of the English /r/ and /l/ contrast by adult speakers of different native languages (English, German, and Japanese). Critically, German and English speakers have categories for /r/ and /l/ in their L1s, while Japanese speakers do not. Thirty equidistant exemplars on an F2 (second format frequency) and F3 (third formant frequency) continuum (between /r/ and /l/) were used to obtain goodness and identification ratings from each of the groups.

Using multidimensional scaling, the perceptual spaces of the three groups of speakers were simulated. Results showed that the perceptual spaces of the three groups of speakers were structured differently as a function of their native language. The findings indicated that for German and English speakers, there is a shrinkage of perceptual similarity of exemplars within the /r/ and /l/ categories and increased perceptual dissimilarity between /r/ and /l/ sounds (Figure 3-1). For the Japanese speakers, however, the exemplars were mostly perceived as being /r/. Critically, their perceptual space did not show convergence between goodness and identification and the shape of their perceptual space, suggesting that they did not have two different speech sound categories for /r/ and /l/.

The reason for Japanese speakers' inability to discriminate between the English /r/ and /l/ seems to lie in their reliance on uninformative acoustic cues. The primary and most salient acoustic cue for distinguishing between these approximants in English is the F3, with the /r/ having a lower F3 than the /l/. Japanese speakers seem to rely on the F2, which is irrelevant for discriminating between approximants in English. These findings suggest that adults' ability to perceive new speech sounds largely depends on their L1 and whether the cues for perceiving speech sounds are the same or different in the L1 and L2.

Figure 3-1. Perceptual spaces for speakers of different L1s (from Iverson et al., 2013).



Notes. The color of each circle represents the most commonly perceived sound for each exemplar by each language group. Black circles represent tokens classified as /r/ and white circles represent those classified as /l/.

L1 facilitation in L2 perception

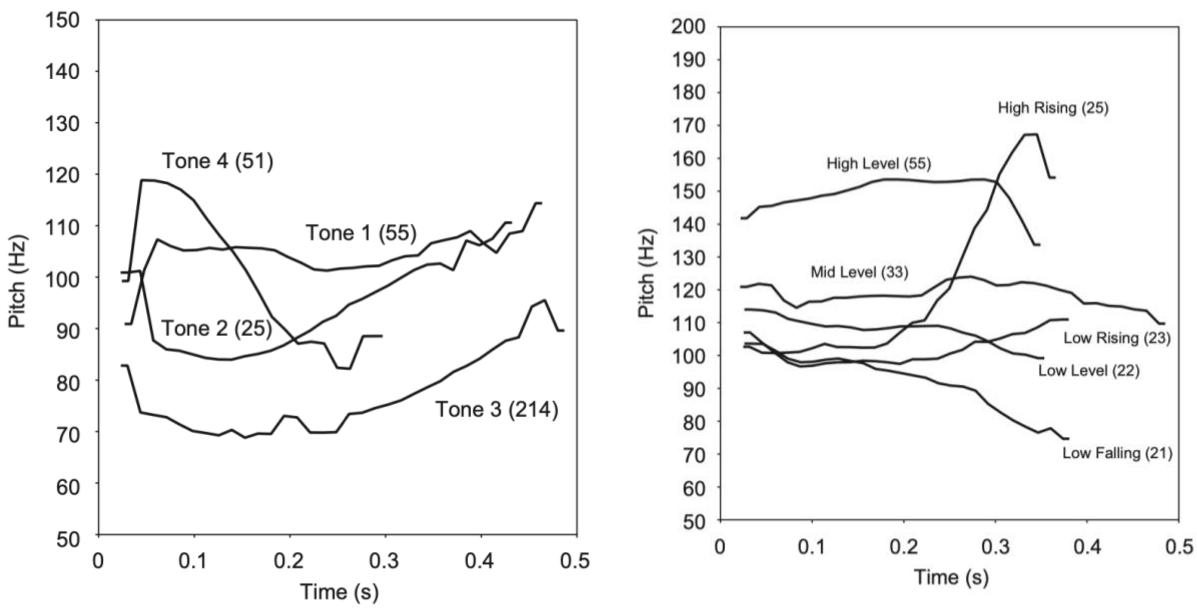
L1 tuning is not always detrimental to L2 speech perception. A study by Chang and Mishler (2012) examined the identification of English unreleased stops by native Korean speakers and native English speakers. In American English, final voiceless stops are sometimes realized as unreleased and sometimes as released. For example, the sound [t] in the word *cat* may be pronounced as released (with a prominent [t] sound) or as unreleased (without a prominent [t] sound). In Korean, however, all final voiceless stops are obligatorily unreleased. The authors hypothesized that by virtue Korean requiring an unreleased stop word-finally, the native Korean speakers would be more sensitive to identifying word final released/unreleased stops even in a new language relative to native speakers of that language (American English) who did not have such native language requirements. Results showed that American English speakers were less well able to identify released/unreleased stops in English, their native language, than Korean native speakers who did not speak English. Taken together, these findings indicate that L1 tuning does not inherently constrain L2 speech perception, but rather it sets up the learner to better acquire L2 features that are present or similar in the L1.

The facilitatory role of L1 knowledge on L2 learning has also been documented for speakers of a tone language learning a new tone language. There are several studies that have examined how knowledge of a tonal native language influences the perception and production of tones in a new tone language (Francis, Ciocca, Ma, & Fenn, 2008; Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh, et al., 2003; Hallé, Chang, & Best, 2004; Krishnan, Xu, Gandour, & Cariani, 2005; Wang, Behne, Jongman, & Sereno, 2004; Wong, 2002; Xu, Gandour, & Francis, 2006). All of the aforementioned studies find that knowing a tone language is associated with better tone perception overall than not knowing a tone language. This is the case when examining performance across all tones. However, there is also evidence that knowledge of a tone language can selectively benefit or interfere with the perception of tones in the L2 depending on the relevant cues for tone in the L1 and L2.

The role of cues

A study by Francis et al. (2008) examined the perception of Cantonese tones by native Mandarin and English speakers. The study involved training the Mandarin and English speakers over the course of ten days, with a pre and post-test of tone discrimination. Cantonese and Mandarin both have similar tone systems in that they both involve tones that vary in pitch direction (e.g upward, downward, level) and in relative pitch height (Figure 3-2). The results of the study found that there was no difference in the perception of Cantonese tones by Mandarin and English speakers overall.

Figure 3-2. Pitch trajectory over time for Mandarin tones (left) and Cantonese tones (right). Image taken from Francis et al. (2008).



A more fine-grained analysis showed the Mandarin and English speakers perceived two of the Cantonese tones (high rising and high level) very accurately at the beginning of the study. Accuracy for the low falling and low level tones was lower prior to training but improved after training for both groups. The finding that both groups of speakers learned two tones equally well suggests that it is not the case that learning new tones requires having an existing set of tone categories in the L1 on which to map the new L2 tones. For the low rising tone, however, only the English speakers improved with training. The investigators suggested Mandarin speakers' difficulty with the low rising tone could be due to them being more sensitive to pitch direction (up/down) than pitch height. Mandarin speakers' difficulty seemed to arise from the Cantonese low rising tone having a very subtle rising contour compared to the rising tones in Mandarin.

This result thus suggests that it is not knowledge of a tone system per se that aids in learning a new tone language but rather having an L1 bias to use the same cue in the L1 and in the L2 (e.g tracking pitch direction/height in a given syllable). Whenever there is a difference in the cue(s) used in the two languages (such as F2/F3 in Japanese and English for /r/ and /l/ or pitch height/pitch direction in Cantonese and Mandarin for the low rising tone), interference may result. If the two languages align in how the phonetic cues are used, there may be a facilitatory effect from knowing how to use the cue in the L1 and transferring that knowledge to the L2.

All studies that have examined tone learning except for one (Wang et al., 2004) have investigated the impact of tonal knowledge via a language that uses tone lexically. In languages that use tone lexically, the same syllable combined with different tones results in different meanings. Although there are differences in the weighing of cues between different languages that use tone lexically (Mandarin vs Cantonese), these differences may be relatively small because these languages use contour and register tones which involve both pitch height and direction. On the other hand, there is another set of tonal languages that do not use contour tones: African languages in the Bantu family. These languages use register (level) tones that differ in pitch height and have a stable pitch direction. If learning to use a different linguistic cue is what leads to interference in learning new tones, one hypothesis that can be made is that speakers of Bantu languages may have difficulty perceiving Mandarin tones because they do not have experience tracking pitch direction at the syllable level.

To my knowledge, there is no prior research that has investigated how knowing a tonal language that only has register (level) tones, may impact the learning of a language such as Mandarin, that has contour tones. In the current study, I investigated this issue. Another important contribution of the current study will be to consider the role of individual differences in learning tones in a new language. Previous research suggests that the ability to learn new phonetic contrasts depends not only on existing language knowledge but also on other kinds of perceptually enriching experiences such as musical experience and on cognitive skills. In the next sections, I review the evidence that these individual differences may be important to consider in order to more accurately assess the contribution of L1 experience to L2/Ln speech perception.

Individual differences in Cognitive Control

Cognitive control is a term that refers to the ability to adjust behavior adaptively depending on current goals. Cognitive control involves a set of neural mechanisms in the prefrontal cortex (Miller, 2000). Behavioral measures of cognitive control typically involve asking participants to make a response while inhibiting a prepotent response. For example, in the Stroop task (Stroop, 1935), participants are asked to read color words that are printed either in the same color as the name of the word (**RED**) or in a different color (**RED**). Reaction times for color incongruent words are typically longer than color congruent words. This congruency effect, known as the Stroop effect, is reflective of efficiency in adjusting behavior to match task demands in the face of interference.

There is evidence that such behavioral cognitive control measures predict speech perception (Lev-Ari & Peperkamp, 2013; 2014). In Lev-Ari and Peperkamp (2014), French native speakers completed an auditory lexical decision task and two cognitive control measures: the Simon and Stroop tasks. In the lexical decision task, participants listened to French words that had a voiced stop. Some of the words were targets (e.g *codé*) that had a lexical neighbor containing a voiceless stop in the same position (*coté*) while others were control words that did not have a voiceless neighbor (*brodé*). In French, voiced stops are prevoiced (negative VOTs) and voiceless stops have short positive VOTs. All words were edited to have shorter VOTs (more prevoicing) or longer VOTs (less prevoicing). The shorter and longer VOT versions of each word were presented in separate blocks, with order counterbalanced. The logic was that if the French speakers activated the voiceless lexical neighbors, which have shorter VOTs, they were more likely to be slowed down by the longer VOT manipulation and to benefit from the shorter VOT manipulation for targets. The reaction time difference between the short and longer VOT conditions was used as an individual difference measure of sensitivity to pre-voicing, with a shorter difference indicating greater sensitivity to prevoicing. Results showed that smaller conflict scores both on the Simon and Stroop tasks predicted greater sensitivity to prevoicing for targets but not controls when the version of the words with the longer VOTs was presented first. This finding suggests that when hearing a realization of the target words in which the voiced stop approximated the VOT values of the voiceless neighbor, French speakers who had better cognitive control abilities were better able to take advantage of the prevoicing cue more efficiently to resolve lexical competition. When the version of the targets with the shorter VOTs was presented first, smaller conflict scores in the Stroop and Simon tasks predicted greater sensitivity to prevoicing for both targets and controls. However, when the shorter VOT condition was heard first (and participants had previously heard the more canonical targets), lexical competition was reduced. Even so, cognitive control abilities were predictive of better speech perception in this less ambiguous context. Taken together, these findings suggests that cognitive control ability may support the allocation of attentional resources to exploit informative acoustic cues when there are varying degrees of contextual ambiguity. Similarly, research on word processing in bilinguals suggests that the ability to overcome L1 phonological activation during L2 listening (and vice versa) depends on cognitive control ability; bilinguals with more efficient cognitive control are able to mitigate non-target language activation during language comprehension in the other language (Blumenfeld & Marian, 2011; 2013).

Perhaps the most direct evidence of cognitive control involvement in speech perception comes from a magnetoencephalography (MEG) study conducted by Ferjan Ramirez et al. (2017). This study investigated speech perception in eleven-month-old infants who were monolingually-exposed to English or bilingually exposed to English and Spanish. The infants listened to a speech sound that was present in both English and Spanish (standard) as well two sounds that were unique to each language (deviants).[19] The standard sound was presented 80% of the time and the deviants 10% of the time. The Mismatch Response for the two languages (the difference between standard and deviants for each language) revealed that the bilingually-exposed infants were sensitive to English and Spanish while the monolingually-exposed infants were only sensitive to English sounds. Importantly, bilingually-exposed infants' (but not monolingually-exposed infants') brain activation extended to prefrontal and orbitofrontal areas, areas associated with cognitive control. This study reveals that speech perception in two familiar languages involves the recruitment of cognitive control mechanisms even at the earliest stages of language development. This study also suggests that bilingual experience provides an opportunity to examine the engagement of cognitive resources for speech perception.

There is also evidence that adults learning a new language engage cognitive control resources when perceiving L2 speech sounds. A neuroimaging study by Wong et al. (2007) examined Mandarin tone perception by native English speakers who did not know Mandarin. The study found that learners who were more successful in identifying Mandarin tones activated bilateral superior and middle temporal regions and the Inferior Temporal Gyrus, areas associated with speech and word processing. Critically, learners who were less successful activated the right medial frontal, anterior cingulate, and middle frontal areas, which is a broader pattern of recruitment that includes cognitive control resources. These findings suggest that cognitive control resources may be important for learning to track pitch in a different way than it is

---

[19] The standard was a voiceless alveolar unaspirated stop (perceived as /da/ in English and /ta/ in Spanish). The Spanish deviant was a prevoiced alveolar stop /da/ and the English deviant was a voiceless aspirated alveolar stop / t^ha/.

typically used in the L1,[20] especially for learners who have less perceptual sensitivity to the novel distinction. Thus, even speakers of nontonal languages may recruit cognitive control for perceiving tones in a new language, especially if the tonal distinction is less perceptible to them. These findings raise the question of whether speakers of different languages, which vary in whether they are tonal or nontonal and, if tonal, in the form of their tone system, would differ in their degree of cognitive control recruitment for tone perception. One prediction is that speakers of nontonal languages, who have less expertise tracking pitch at the syllable level, may be more likely to recruit cognitive control resources for perceiving tones in a new tone language. In addition, if the specific cues for tracking tone differ in the native and target language, then we might also expect greater cognitive control recruitment than for speakers of languages that have a more dissimilar tone system to the target language.

Individual Differences in Musical Experience

There is also evidence that individual differences in musical experience play a role in tone learning. A series of studies by Perrachione and Wong have investigated Mandarin tone discrimination and lexical tone learning in native English speakers without prior Mandarin experience (Perrachione & Wong, 2007; Wong, Perrachione, & Parrish, 2007). These studies show that there are substantial individual differences in learning tones in a new language. Some of those differences appear to be driven by anatomical brain differences (Wong et al., 2007) while others may be at least partly induced by

[20] Variation in pitch is used in all languages in different ways. Although pitch is not used contrastively at the syllable level in English, it is used to express a range of meanings such as talker uncertainty (Ward & Hirschberg, 1985).

experiences such as musical training. Perrachione and Wong (2007) investigated the learning of Mandarin tones and new vocabulary using the newly learned tones by native English speakers. Tone discrimination involved listening to a tone and deciding whether that tone was rising, falling or level. The lexical tone learning task involved learning the meanings of novel English syllables paired with a particular tone. These studies have found that individuals who reported being musicians had better pitch discrimination ability and that pitch discrimination ability in turn predicted word learning outcomes.

The finding that having musical experience predicted better tone discrimination indicates that having played music may have improved perceptual sensitivity by training the tracking of pitch height and direction.[21] Thus, even nonlinguistic experiences may contribute to learning tone in a new language. A study by Bowles, Chang, and Karuzis (2016) examined English speakers lexical tone learning ability in Mandarin. In a lexical tone learning task, the English speakers were asked to learn the meanings of novel words that have the same syllable but differ in tone. The study considered the role of individual differences in pitch recognition, language learning aptitude, musical experience, and cognitive ability for learning the novel words with the Mandarin tones. By far the strongest predictor of lexical tone learning was tone recognition. The years of musical experience reported by the participants also predicted word learning to a lesser extent. Although the contribution of years of musical experience was relatively small, a Principal Component Analysis revealed that pitch ability and musical experience loaded onto one component. These data support the notion that pitch ability and musical experience are aspects of the same auditory processing system (Perrachione,

[21] Such experience-induced effects likely also reflect the tendency of people to pursue activities that build on their biological predispositions.

Fedorenko, Vinke, Gibson, & Diley, 2013). Taken together, these results suggest that musical experience may contribute to sensitivity to track pitch, which is important both for tone perception and for learning to use tones in word learning.

Present Study

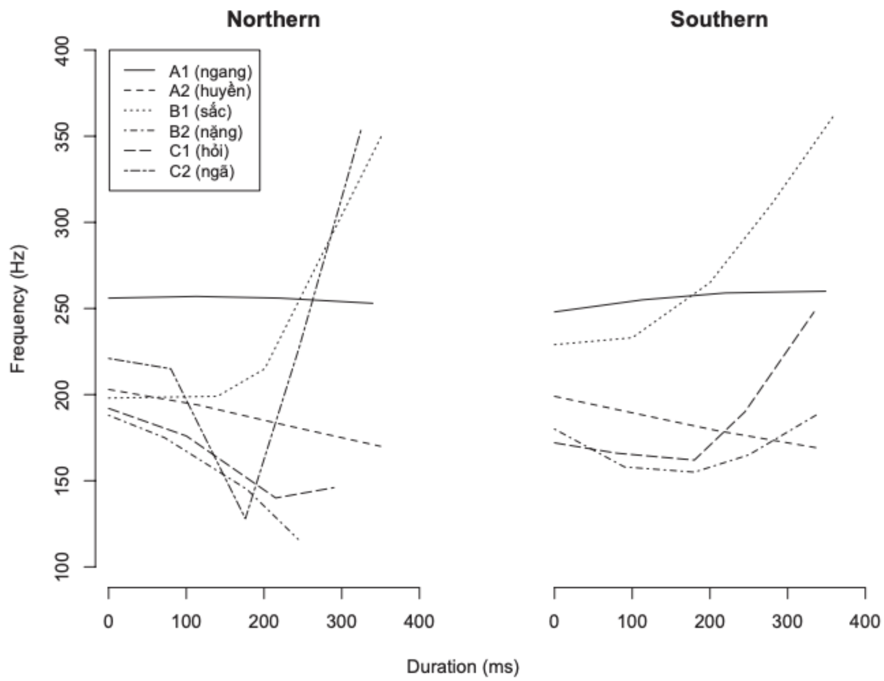
In the present study, I investigated the impact of L1 tone knowledge and of the type of L1 tone system for learning tones in a new language (Mandarin). Participants were all multilinguals who spoke different L1s but all knew English as an L2. None of the participants had prior knowledge of Mandarin. I was interested in two comparisons. The first was between the L1 speakers of tonal languages (Vietnamese, Bantu languages) versus L1s speakers of nontonal languages (Spanish, Dutch). The second comparison was between native speakers of an L1 that uses pitch height and pitch direction as cues for tone (Vietnamese) and native speakers of an L1 that only uses pitch height as a tonal cue (Bantu languages). I was also interested in the degree to which individual differences in cognitive control ability and musical experience account for variance in tone perception. I first review the critical features of the two tonal languages considered in the present study (Vietnamese, Bantu languages) and their relationship to the target language (Mandarin). I then lay out a set of predictions about group and individual differences.

Vietnamese

Vietnamese is a language from the Austroasiatic language family. It has two main dialects: Northern Vietnamese (spoken in Hanoi) and Southern Vietnamese (spoken in

Ho Chi Mihn). Both dialects have tone inventories consisting of register (level) and contour tones. The differences between the dialects lie in the number of tones and the cues used to distinguish between them (Brunelle, 2009; Kirby, 2010). In Northern Vietnamese, there are six tones. Of the six Northern tones, one is a level tone (*năng*), one is a rising tone (*sắc*), one is falling (huyền), and the last two (*ngã, hói*) have an initial fall and rise again at the end. The last two Northern Vietnamese tones that fall and rise (*ngã, hói*) are distinguished by a combination of the pitch direction, pitch height, glottalization and creakiness. The *Ngã* tone falls, then comes to a stop with glottalization and ends on a dramatic rise. The *Hói* tone falls until there is a creaky voice quality and then it rises (but less sharply than the *ngã* tone). In Southern Vietnamese, there are only five tones. The level (*năng*), rising (*sắc*), and falling (huyền) tones are the same as in Northern Vietnamese. The only difference is that in Southern Vietnamese the *ngã* and *hói* tones have merged into a single tone that falls and rises without creakiness. The tone inventories of Northern and Southern Vietnamese are shown in Figure 3-3.

Figure 3-3. Tone inventories of Northern and Southern Vietnamese. Image from Kirby (2010).



Like Mandarin, Vietnamese (whether it is the Northern or Southern variety) has both register and contour tones. Both Mandarin and Vietnamese use tone lexically; every syllable has a tone assigned to it and the use of different tones changes word meaning. For example, the syllable *ba* paired with the six different tones produces six different meanings (the flat tone is indicated by no diacritic):

ba ‘three’

bà ‘grandmother’

bá ‘to embrace’

bạ ‘to strengthen'

bả 'bait'

bã 'residue'

Thus, Vietnamese and Mandarin have very similar tone systems, with a combination of register and level tones. Both languages use tone lexically to convey word meaning. Although Vietnamese has more tones than Mandarin and more tonal acoustic cues, from the perspective of a Vietnamese speaker learning Mandarin, the types of tones of Mandarin should be familiar because the rising, falling, and level contours are present in both languages.

Bantu Languages

Proto-Bantu is a language family of approximately 500 languages spoken in Africa (Downing, 2011). The vast majority of Bantu languages are tonal,[22] with a two-tone system.[23] The high tone is regarded as the active tone, while the low tone is regarded as the default tone. All of the Bantu languages that are tonal have register (level) tones, with the tones differing from one another in pitch height. This is in contrast to languages like Mandarin and Vietnamese, in which there are both register (level) tones and contour tones (e.g rising, falling).

Bantu languages are agglutinative. Words are formed by adding prefixes to a root morpheme in order to convey different information. Word meaning is largely determined by the combination of morphemes. In Bantu languages, morphemes

[22] The exception is Swahili which was once tonal but lost tonality after contact with Arabic.
[23] The exceptions to the two-register tone system are languages in which there are three of more register tones.

themselves have tones associated with them, but once they are joined into a word, there are several contextual factors that determine where tones will surface. These factors include word boundaries, syllable position (Cassimjee, 1998; Kenstowicz & Kisserberth, 1990; Kisseberth, 1992; Myers & Carleton, 1996), and syntactic boundaries (Kisseberth, 1994).

Most relevant for the present investigation of tone perception, Bantu languages do not use tone lexically (i.e., speakers of Bantu languages do not exploit tone as a source of semantic information at the syllable level). There are very few minimal pairs in Bantu languages. Tone placement is the result of various phonological processes (e.g Clements & Goldsmith, 1984; Goldsmith, 1984) that delete adjacent tones and dictate where the high tones will surface. This means that Bantu languages resist having certain tone sequences and, in some cases, eliminate tone sequences altogether. Some analyses of Bantu languages suggest that tone is used in a more accentual manner (Downing, 1996; McCawley, 1970; 1978). Another aspect of Bantu tone that is different from both Mandarin and Vietnamese is that tone in this language family interacts with syntactic information, causing tones to spread in specific ways (Kisseberth 1994).

To summarize Bantu tone differs from Mandarin tone in the following ways:

(1) Bantu languages do not have contour tones, only register (level) tones. This means that it is pitch height that is the relevant cue for tone in Bantu languages, whereas Mandarin uses both pitch height and pitch direction as cues.

(2) In Bantu languages, tone placement is highly variable depending on a number of contextual factors.

(3) In Bantu languages, tone is generally not used to distinguish between the meanings of words.

Predictions: Tone Discrimination

Group Level. If it is a general perceptual expertise in attending to pitch that supports the perception of tones in a new language, then we might expect: (1) L1 tonal language speakers (Vietnamese and Bantu) to perceive Mandarin tones more accurately than nontonal language L1 speakers (Spanish, Dutch). However, if it is more specific experience in attending to target language relevant cues in the L1 that facilitates L2 learning of another tone language, we might expect Vietnamese speakers to learn to distinguish Mandarin tones more easily than the other three groups who have no knowledge of a tonal language that has contour tones (Spanish, Dutch, Bantu). If there is both a contribution of tone knowledge and an effect of the pitch cues used in the native language for learning new tones, then we might expect a graded effect of tone knowledge such that: (1) Vietnamese and Bantu speakers will distinguish Mandarin tones better than the nontonal language speakers (Spanish, Dutch) and (2) Bantu language speakers will perceive Mandarin tones less accurately than Vietnamese speakers.

Individual differences. Musical experience may be another experience (in addition to having learned a tone language as an L1) that aids in the ability to track pitch. Thus, higher levels of musical experience should predict better tone identification in Mandarin. If musical experience and tone perception are part of a larger auditory perceptual ability (Bowles et al., 2016), then we might expect that participants who have prior tone knowledge will benefit less from musical experience than participants without prior tone knowledge. This prediction assumes that there is a limit to pitch discrimination acuity and that learners draw upon their native language experience first, since it is generally learned before musical training starts. Once the native language is learned, musical experience is presumed to further develop pitch discrimination acuity. The degree of benefit from musical experience may thus depend on how much the native language has already developed pitch acuity. Learners who had prior tone experience are expected to need less musical experience to achieve the same level of pitch acuity as nontonal language speakers.

If cognitive control recruitment depends on how salient the tonal cues are to the learners (Wong et al., 2007), then we might expect that for the learners who do not have experience using pitch height and pitch direction as tonal cues (Spanish, Dutch) cognitive control will predict the accuracy in Mandarin tone perception. Because the Bantu speakers have experience attending to at least one of the relevant cues, we might also expect individual differences in their cognitive control ability to predict Mandarin tone perception but to a lesser extent than for the nontonal L1 groups. Since Vietnamese speakers have experience using pitch height and pitch contour as tonal cues, these cues are expected to be highly salient to them in Mandarin, and they may not recruit cognitive control for perceiving Mandarin tones. Therefore, there may not be a relationship between Vietnamese speakers' cognitive control skills and their accuracy for tone perception in Mandarin.

Method

Participants

28 Vietnamese-English bilinguals, 25 Spanish-English bilinguals, 20 Bantu-English bilinguals, and 22 Dutch-English bilinguals participated in this study. The same experimental materials were used for all groups with instructions provided in English. The Vietnamese and Spanish speaker groups were both heritage speakers of Vietnamese and Spanish living in the US. They were recruited using flyers, social media, by word of mouth, and to a lesser extent using the platform Prolific Academic. Because the Vietnamese and Spanish groups were heritage speakers living in the US, they were expected to be more English-dominant than the other two groups who were living in countries where their L1s were spoken. The so-called "Bantu-English bilingual" group consisted of native speakers of multiple Bantu languages currently living in either South Africa or Botswana. The tone languages represented in the Bantu group included: Zulu (n = 17), Tswana (n = 13), Sotho (n = 9), Xhosa (n = 6), Tsonga (n = 3), Venda (n = 3), Ndebele (n = 2), and Swazi (n = 2).[24] All of these languages have two register tones (high and low). On average, participants in the Bantu group knew three tone languages (M = 3.10, SD = .85). Other than English, participants' nontonal languages reported included German (n = 2), Hindi (n =1), Afrikaans (n = 8), and Swahili (n = 1). To participate in the study, the speakers of Bantu languages had to know at least one tone language.[25] The Dutch group was also of interest because Dutch is not a tonal language and because English is widely used in the Netherlands. The

[24] All participants spoke more than one tone language, therefore they are counted multiple times.
[25] This was generally the profile for all participants recruited. The exception were some participants who reported knowledge of Afrikaans and Swahili (not tonal languages). These potential participants could only be included if they spoke at least one Bantu language with tone.

Bantu and Dutch-English groups were recruited on Prolific (www.prolific.co). The participants' demographic and language characteristics are summarized in Table 3-1.

Table 3-1. Summary of participant characteristics.

	Groups			
Variable	Vietnamese	Spanish	Dutch	Bantu
	M(SD)	*M(SD)*	*M(SD)*	*M(SD)*
Age	23.85(4.99)	22.48(3.67)	26.90(5.01)	26.80(4.73)
Education	4.11(1.52)	4.19(1.23)	5.28(1.38)	5.00(1.55)
L1 Exp	31.77(18.04)	34.19(11.54)	67.14(14.62)	32.10(15.73)
English Exp	65.37(17.84)	64.42(10.61)	27.32(12.07)	48.80(18.44)
Other Exp	2.85(4.97)	1.38(3.06)	5.52(6.01)	27.1(12.23)
Musical Experience	10.53(9.01)	4.56(4.16)	7.76(8.45)	4.70(6.07)

Notes. The L1s were Vietnamese, Spanish, Dutch, or the native Bantu language depending on the language group. Exp = Exposure.

General Procedure

Participants completed the Language History and Musical Experience survey on Qualtrics. They then completed the Simon and pitch identification tasks. Participants recruited on Prolific completed all tasks on FindingFive, whereas participants recruited via other methods completed the Simon task on tatool and the other tasks on FindingFive. Specific materials and procedures for the tasks are provided below.

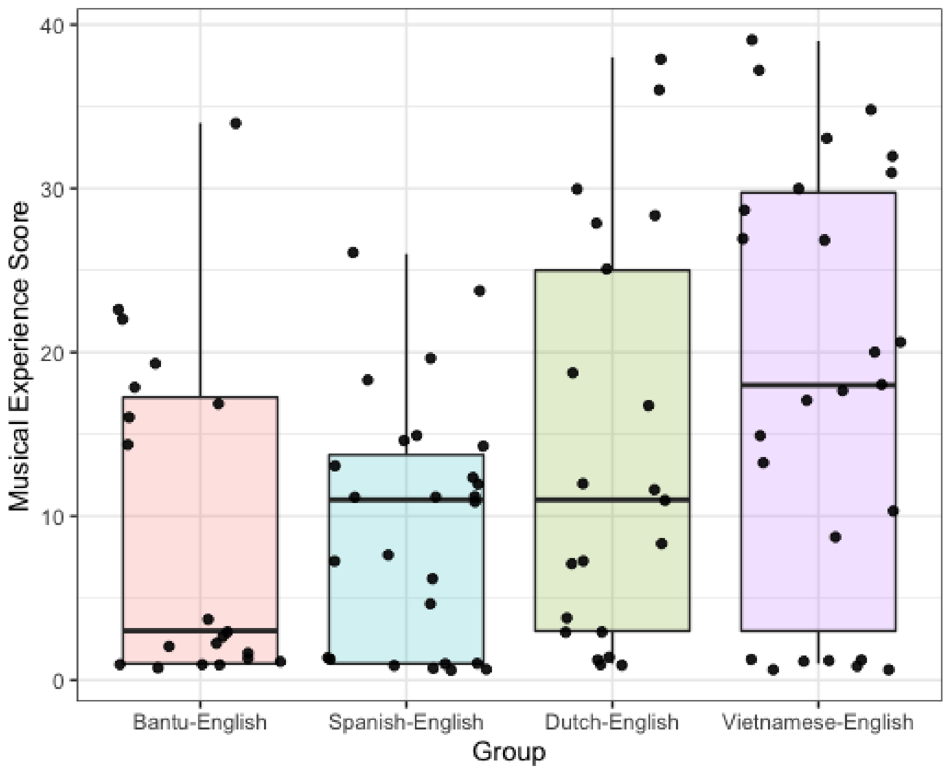
Measures

Language history and musical experience questionnaire

Participants were asked general demographic questions such as their age, highest education level, and about any history of hearing problems. Regarding language experience, they were asked to report the languages that they knew, their proficiency, age of acquisition and exposure to each language. In addition, they were also asked to report if they had previously played instruments, how often they practiced, how many years of instruction they had and whether they had recently practiced these instruments. They were also asked if they had ever taken classes in music theory, or sung in a group chorus or played in an orchestra. The full questionnaire is provided in Appendix G. From the questions regarding musical experience, I created a composite score of musical experience.[26] This score was the sum of the number of years participants had played an instrument and one additional point was added for additional musical experiences they reported (having taken music lessons, having recently practiced at least one musical instrument in the past thirty days, composing music, singing in a chorus/playing in an orchestra, teaching another person how to play an instrument, etc). The distribution of musical experience scores for each group is shown in Figure 3-4. Although all groups varied in musical experience, the Dutch and Vietnamese had more variation than the Spanish and Bantu groups. Another key difference is that the Bantu group is the only group with a bimodal distribution of musical experience scores.

Figure 3-4. Distribution of musical experience scores for each group.

[26] Previous research that has investigated musical experience in relation to tone learning has operationalized musical experience in a variety of ways. Some studies have used musician status as a categorical variable (Wong & Perrachione, 2007). Other studies have used years of musical experience as a continuous predictor (Bowles et al., 2016). In the current study, a composite score that heavily weighted years of musical experience was used but additional language experiences were also counted in order to capture individual differences more sensitively.



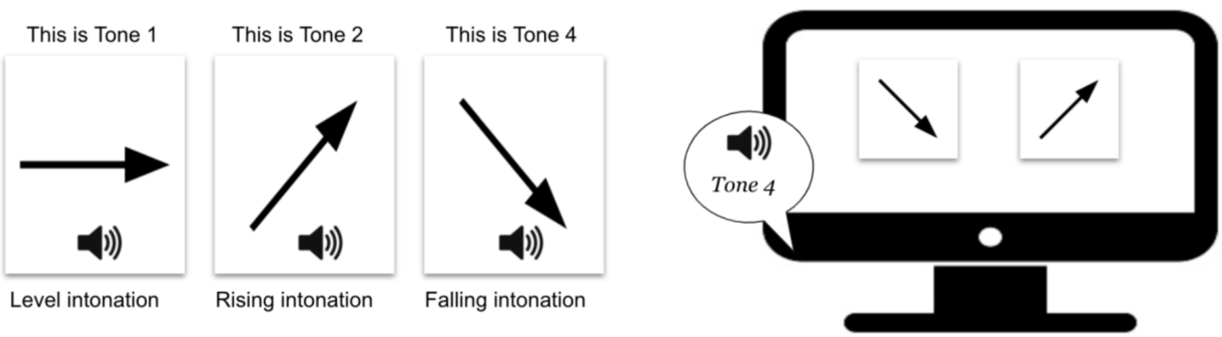
Pitch identification (Wong & Perrachione, 2007)

Materials. Two male and two female native speakers of Mandarin were recorded saying the vowels [a, e, i, o, y] with Mandarin tone 1 (level). The recordings were then resynthesized to superimpose the three of the Mandarin tones on them: tone 1 (level), tone 2 (rising) or tone 4 (falling).[27] The same stimuli from the original Wong and Perrachione (2007) study were used (Appendix H).

[27] Tone three was not investigated because the original study by Wong and Perrachione (2007) described the third tone as being particularly difficult to learn for native speakers of English and also as confusable by both native and nonnative speakers.

Procedure. Participants were first shown a set of three images that corresponded to tones one, two, and four (Figure 3-5). In a practice phase, they listened to a syllable produced with the different tones in order to learn to associate the shape of the tonal contour with the direction of the arrow. In the test phase, they were asked to listen to a speaker produce one of five syllables [a, e, i, o, y] with either tone 1 (level), tone 2 (rising) or tone 4 (falling) in Mandarin and to identify whether the tone on the syllable was level, rising, or falling. Participants had unlimited time to respond with a button press.

Figure 3-5. Overview of pitch identification task.

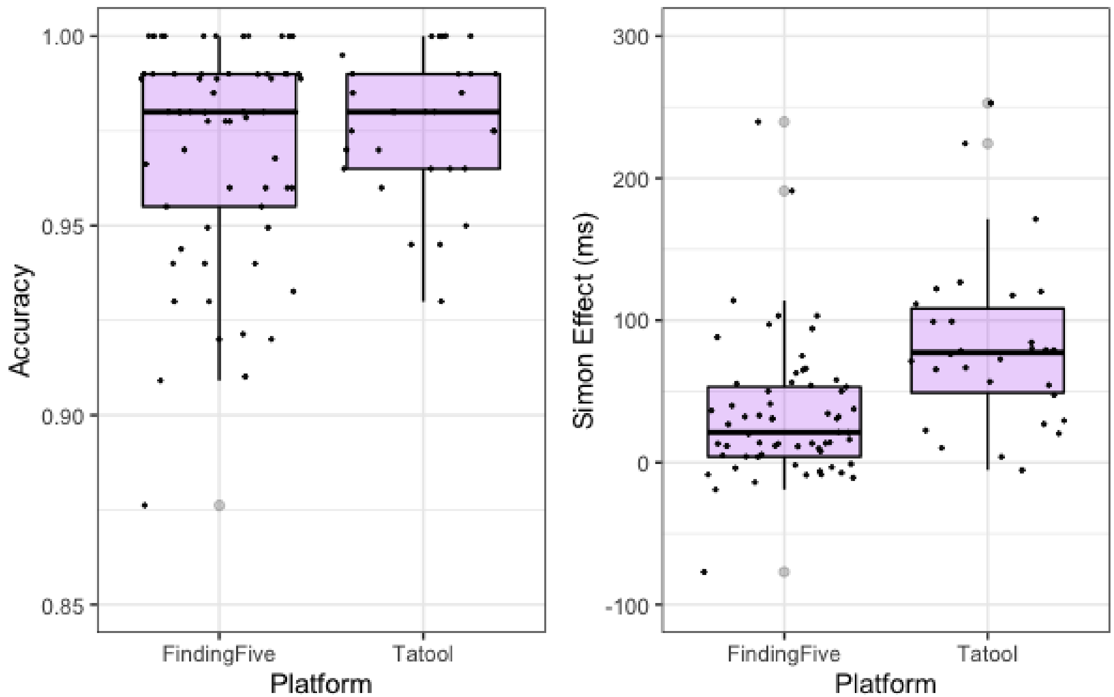


Notes. On the left are the images of the tones that participants were initially trained to associate with the images of arrows that show pitch trajectory (level, rising, falling). On the right is a sample test trial, in which participants heard a vowel with tone 4 (falling) and were asked to select the corresponding tone using the image.

Simon task (Von Bastian, Locher, & Ruflin, 2013)

The version of the Simon task that was used in this study had red and green circles as shapes and included 50% congruent and 50% incongruent trials. Participants were asked to press the left button when presented with a green circle and the right button when presented with a red circle regardless of the location of the circle. There were no central trials. The majority of the Spanish-English and Vietnamese-English bilinguals' data for the Simon task were collected on the website tatool (https://www.tatool-web.com/). However, for other groups, for whom data were collected solely via Prolific, it was necessary to combine all experiments into a single platform: FindingFive. Thus, the FindingFive version of the Simon task programmed to match the design and timing of the tatool version. Both versions had 100 trials. Reaction times for incorrect responses were excluded. Reaction times above or below 2.5 standard deviations above or below the mean for each participant were also excluded from analysis. Participants tested on Tatool (*n* = 32) and FindingFive (*n* = 66) did not significantly differ in their average accuracy (*t* = .91, *p* = .36). However, participants who completed the task on tatool (*t* = 3.20, *p* = .001) had a significantly larger Simon Effect (Figure 3-6). Given that there were differences in the magnitude of the Simon effect by platform, I z-scored the Simon Effect separately for each group in order to conduct the analyses with all participants. This preserves the relative ranking of each participant, although the raw values differed for participants with the same z-score that did the experiment in the two platforms.

Figure 3-6. Overall accuracy and magnitude of the Simon Effect by experimental platform.



Analysis

Data were analyzed in the programming environment RStudio (RStudio Team, 2020) using the programming language R (version 4.1.2; R Core Team, 2022). The analysis focuses on the factors that predicted Mandarin tone identification. This analysis was conducted using a mixed logistic regression using the package *lme4* (version 1.1-27.1; Bates, Mächler, Bolker, & Walker, 2015). Binary accuracy outcomes for tone identification (correct, incorrect) were fitted using the function *glmer*. We considered the following predictors for inclusion in the model: group (Spanish, Vietnamese, Dutch, Bantu), Musical Experience Score, and the Simon Effect. The maximal random effects structure was attempted. Whenever convergence issues arose, correlations were first removed and if convergence issues persisted problematic terms were removed. Continuous predictors were z-scored (Musical Experience Score and the Simon Effect). The final model is summarized in a table. Only statistics not summarized in the tables are included in the text. Figures are provided to illustrate interactions.

Results

A logistic regression was fitted with trial-level pitch identification accuracy as the dependent variable. Adding Group as a predictor significantly improved model fit to the intercept-only model, $\chi 2$= 36.07, p <.001. Adding musical experience score to the model with Group also improved more fit, $\chi 2$= 17.85, p= .001. The addition of the Simon Effect did not improve model fit, $\chi 2$= 7.35, p= .49. The final model included by-item random slopes for group. The final model is summarized in Table 3-2.

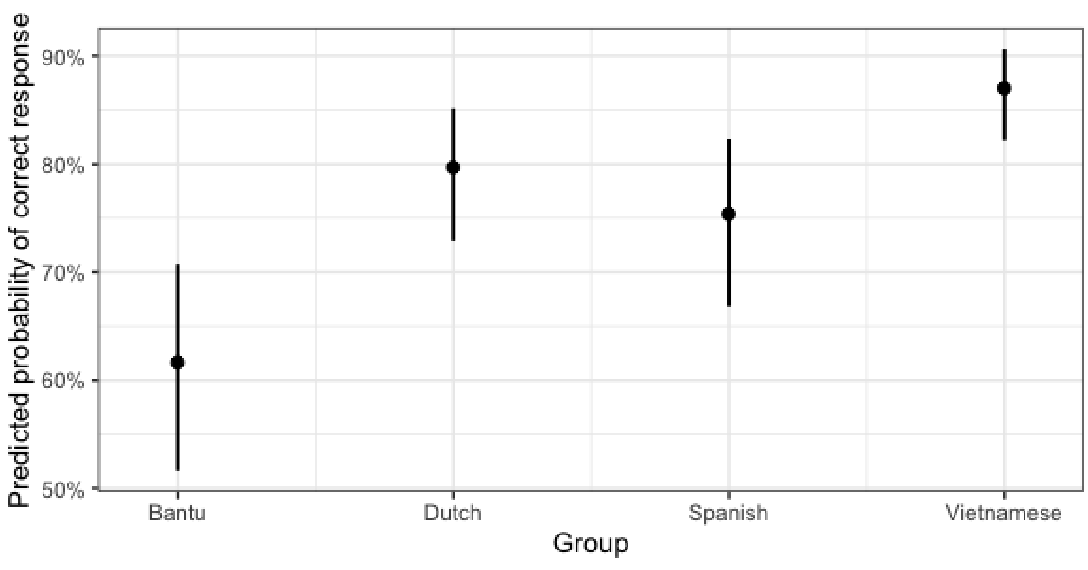
Pitch identification accuracy by group is shown in Figure 3-7. Relative to the Spanish-English group, the Dutch-English bilinguals did not differ in their likelihood of accurately identifying Mandarin tones. Thus, the two non-tonal language groups performed equally well on tone identification. The Vietnamese-English group, which had the most similar tone system to Mandarin, had a significantly higher likelihood of accurately identifying Mandarin tones than the Spanish-English bilinguals (z = -2.74, p =.005), the Dutch-English bilinguals (z = -5.39, p <.001), and the Bantu-English bilinguals (z = -2.74, p =.005). This result is consistent with a benefit of L1 tone knowledge when the tone system in the L1 (Vietnamese) uses both of the relevant cues as the target language.

Table 3-2. Summary of pitch identification model.

Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	3.28	2.13 – 5.06	**<0.001**
Group [Bantu]	0.48	0.27 – 0.86	**0.014**
Group [Dutch-English]	1.22	0.69 – 2.13	0.493
Group [Vietnamese-English]	2.21	1.25 – 3.88	**0.006**
MusicalExperience	2.47	1.33 – 4.60	**0.004**
Group [Bantu] *MusicalExperience	0.42	0.19 – 0.93	**0.032**
Group [Dutch-English] *MusicalExperienceScore	0.50	0.25 – 1.02	0.057
Group[Vietnamese-English]*MusicalExperience	0.64	0.32 – 1.27	0.205

Notes. The reference level for Group was Spanish-English bilinguals, but additional comparisons are reported in the text.

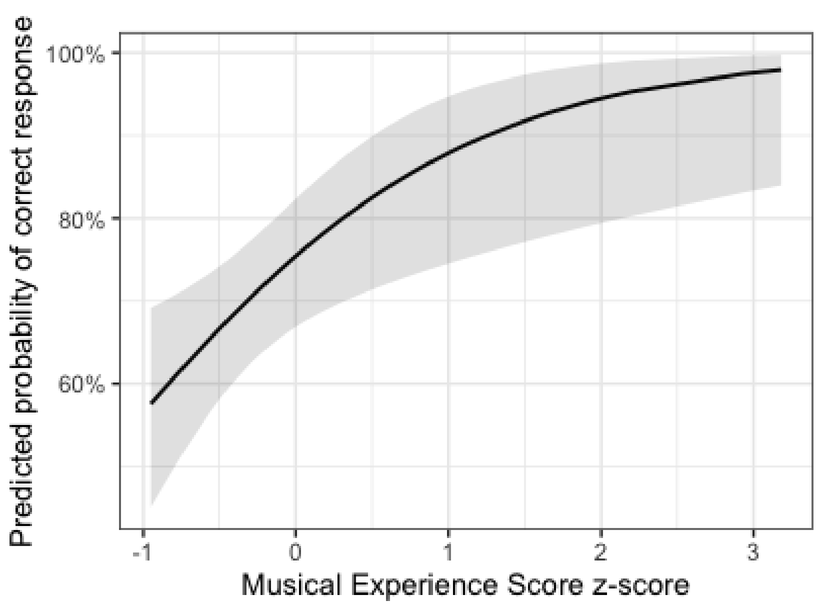
Figure 3-7. Model predicted probabilities of accurate Mandarin identification by group.



The Bantu-English bilinguals were predicted to have a significantly lower likelihood of accurately identifying Mandarin tones than the nontonal groups ($z_{Spanish}$ = 2.46, p =.01; z_{Dutch} = 3.30, p <.001). This result suggests that the benefits of L1 tone knowledge do not extend to cases when only one of the relevant tonal cues from the target language is used in the L1. The finding that Bantu speakers were less likely to accurately identify Mandarin tones than the nontonal groups suggests that the Bantu speakers experienced interference from their L1 only exploiting one of the two tonal cues relevant in Mandarin.

There was also a main effect of musical experience, with participants who had a higher musical experience score being predicted to have an increased likelihood of accurate tone identification than people who had less musical experience (Figure 3-8).

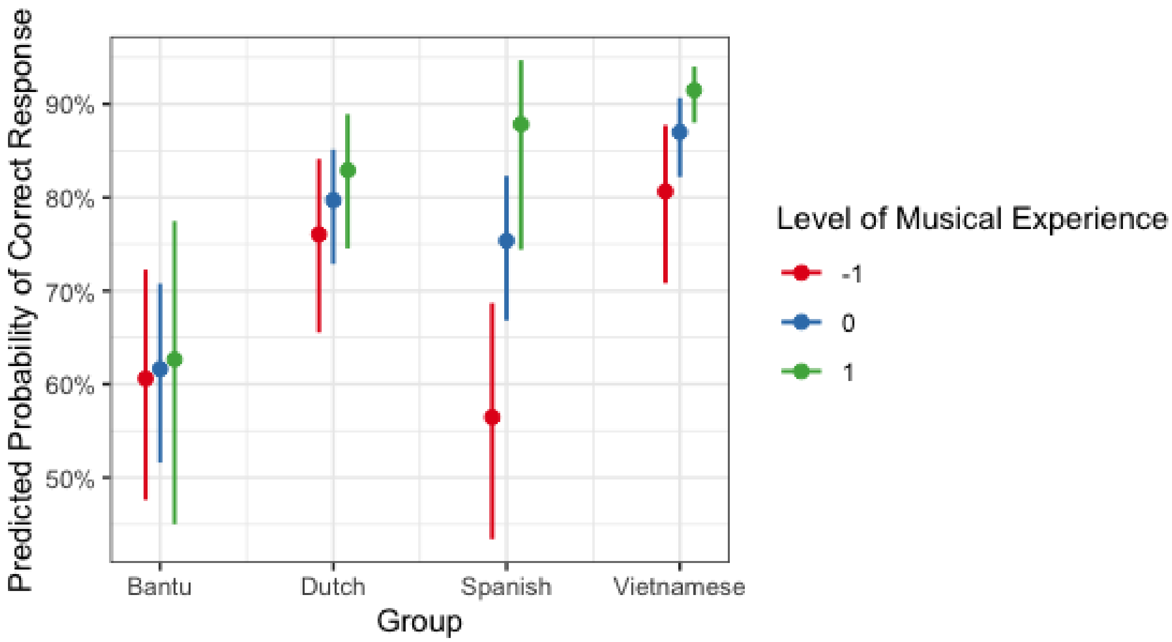
Figure 3-8. Relationship between musical experience and Mandarin tone identification.



The effect of musical experience was qualified by an interaction with group. The effect of musical experience was only significant for two of the groups, Spanish and Vietnamese speakers. This interaction is illustrated below in Figure 3-9. For Spanish and Vietnamese groups, higher levels of musical experience predicted a higher likelihood of correct Mandarin tone identification. However, musical experience did not predict differential Mandarin tone identification for Bantu and Dutch groups. This result is inconsistent with the prediction that musical experience would be more predictive of Mandarin tone perception for groups with nontonal experience, since Bantu speakers had prior tone experience. The fact speakers of Vietnamese and Spanish both benefitted from prior musical experience but Bantu and Dutch speakers suggests that the effect of musical experience may depend on sampling differences. I return to the sampling issue in the General Discussion. Finally, there was no relationship between the Simon Effect and tone identification.[28]

[28] One concern is that the FindingFive version could have truncated variation in the Simon Effect, making it more difficult to detect a relationship between cognitive control score and tone discrimination. Thus, we fitted the same model but only with the participants who completed the Simon task on tatool. Again, there was no relationship between Simon scores and tone identification, $z = -.66$ $p = .39$.

Figure 3-9. Predicted probability of accurate Mandarin tone identification by level of musical experience and by group.



Notes. For Level of Musical Experience, -1 means 1 SD below each group's mean level of musical experience, 0 is each group's mean level of musical experience, and 1 is 1 SD above each group's mean level of musical experience.

Discussion

In this study, I examined how prior knowledge of a tone language affects the perception of tones in a new tone language. Four groups of participants who did not have prior knowledge of Mandarin were asked to identify Mandarin tones: Vietnamese-English, Spanish-English, Dutch-English, and Bantu-English bilinguals. Two groups spoke a tone language as an L1 (Vietnamese or a Bantu language) and two were native speakers of a nontonal L1 (Dutch or Spanish). Participants were asked to listen to speakers of

Mandarin producing the Mandarin tones and to identify each tone as level, rising, or falling.

I tested three hypotheses about how tonal knowledge in the L1 affects the ability to identify tones in a new language. The first hypothesis was that general knowledge of a tone language may aid in learning a new tone language. For this hypothesis to be supported, L1 speakers of tone languages (Vietnamese and Bantu) should have better Mandarin tone identification than the non-tonal L1 speakers (Dutch and Spanish). The second hypothesis was that whether the tonal L1 required attending to the same or different cues for tone (pitch height, pitch direction) would differentially predict tone perception in the new tone language. For this hypothesis to be supported, Vietnamese speakers (who know an L1 that uses pitch direction and pitch height like Mandarin) should distinguish Mandarin tones better than the other three groups who do not know a tone language that uses pitch height and direction as cues (Spanish, Dutch, Bantu). The third hypothesis was that if both general knowledge of a tone language and familiarity with the specific pitch cues in the target language contribute to tone discrimination in a new language, then one might expect a graded effect of tone knowledge with Vietnamese speakers distinguishing Mandarin tones better than the nontonal language speakers (Spanish, Dutch) and with Bantu language speakers distinguishing Mandarin tones more accurately than the nontonal language speakers (Spanish, Dutch) but less accurately than the Vietnamese speakers.

The results did not support the first and third hypotheses. There was no general advantage for tone language speakers over nontonal language speakers. The Vietnamese-English bilinguals outperformed not only the nontonal L1 groups but also the Bantu group. The third hypothesis, that there might be a graded effect of tone similarity, was not supported, since the Bantu speakers were less accurate than all groups on Mandarin tone identification. These results are consistent with the second hypothesis, that it is the specific tonal cue(s) that that are informative for identifying tone in the L1 that affected the perception of tones in Mandarin. The Vietnamese-English bilinguals were the most likely to identify the Mandarin tones correctly, presumably because in Vietnamese both pitch height and direction are relevant cues. However, in Bantu languages, only pitch height is relevant, since these languages only have register (level) tones. Therefore, when listening to the Mandarin tones, it appears that the Bantu speakers could not reliably attend to both of the relevant pitch cues in Mandarin.

A close examination of the groups' demographic characteristics, which were not perfectly matched, suggests that the Bantu group's performance was unlikely to be due to differences in education, since the Bantu group was more educated than Spanish and Vietnamese groups but still had significantly lower Mandarin tone identification accuracy than both of them. These results support the notion that Bantu speakers' knowledge of languages that use register tones likely resulted in them being tuned to attend to pitch height. This tuning to pitch height may have in turn caused them interference when needing to attend to both pitch direction and pitch height in Mandarin.

One key difference between the groups was in the amount of L1 exposure they reported at the time of the study. Vietnamese, Spanish, and Bantu speakers reported lower L1 exposure than Dutch speakers (31%, 34%, and 32 vs. 67%). Lower L1 exposure was expected for the Vietnamese and Spanish speakers who were English-dominant heritage speakers living in the US. However, it is important to note that even though the Bantu speakers had lower levels of exposure to their L1, they also tended to know multiple tone languages and to be actively using them in their daily lives. Thus, the amount of L1 exposure reported is an underestimate of their tonal language experience. Virtually all of the Bantu speakers' non-English communication time (52%) was in Bantu languages that have register tones. The fact that the Bantu group (who had the most daily tone exposure) was less accurate than the Vietnamese heritage speakers (who were only exposed to Vietnamese 31% of the time) suggests that it is not active exposure to tone but early language experience and L1 tuning that determines how well tones will be perceived in a new tone language. The issue of how early language experience versus current use of the L1 affect tone perception will be further examined when verbal fluency data are coded for all groups.[29]

One prediction that can be made is that if early language experience is what contributed to the Vietnamese speakers' benefit in perceiving Mandarin tones, their language dominance should not affect their ability to perceive Mandarin tones. However, if there is a contribution of active Vietnamese use, this would indicate that it is not just early experience but active use of the L1 that are important for phonetic learning. For the Bantu speakers, data from the language questionnaire will also be extracted that will enable me to analyze if it is not just their L1 but also exposure to and use of other Bantu languages that predicts Mandarin tone perception. For the Dutch and Spanish speakers, who have no prior tone language, there is also an opportunity to examine whether aspects of early language experience (e.g. age of acquisition) versus aspects of active bilingualism (e.g. language dominance) predict Mandarin tone

[29] The data for this experiment were collected a week prior to the writing of this chapter and there has not been enough time to code all measures.

perception. Previous research suggests that early bilingual experience tunes the learner to have greater perceptual openness to new speech sounds (Pettito et al. 2011). However, it is not clear if it is only early language experience or more recent language experience that create this perceptual openness. If there is a benefit of early bilingual experience for new phonetic learning, then we might also expect that early bilinguals will be advantaged relative to later bilinguals. On the other hand, if it is active bilingualism that contributes to this perceptual openness, then we might expect that more balanced bilinguals will be advantaged relative to more imbalanced bilinguals.

There was significant individual variation within each of the language groups. A goal of the study was to consider the role of individual differences in cognitive control and in musical experience for learning tones in a new language. Cognitive control resources were measured using the Simon task. There was no relationship between the cognitive control measure and Mandarin tone identification. The finding that there was no relationship between the Simon score and tone identification might suggest that reactive control is not important for tone perception. However, that conclusion would be inconsistent with previous research that has reported a relationship between reactive control efficiency and speech perception (Lev-Ari & Peperkamp, 2013; 2014).

One possibility is that the measure of cognitive control used was not sensitive enough, especially via online delivery. The other possibility is that the cognitive resources captured by the Simon task were not critical for tone identification. In order to better understand the role of cognitive control in speech perception of a new language, more studies are needed that use a larger set of cognitive measures and that examine the role of cognitive control after continued exposure to the new language.

The other individual difference of interest, musical experience, was predictive of tone identification but only for two of the language groups (Spanish and Vietnamese). A possible explanation for this result is that the ability to detect the effect of musical experience was limited by the range of variability in musical experience scores within groups. An examination of the range of scores for all of the groups suggests that there was sufficient variability in musical experience scores within each group, so this is an unlikely explanation. One difference in the distribution of musical scores, however, is that for the Bantu group, musical experience scores had a bimodal distribution whereas all other groups had a unimodal distribution. The Dutch and Vietnamese groups had similar amounts of variation, but the effect of musical experience was only observed for the Vietnamese group. The Spanish and Bantu groups had the least variation, but the effect of musical experience was detected only for Spanish speakers. Thus other characteristics of the sample may have been confounded with musical experience. Spanish and Vietnamese speakers were the least educated of the four language groups. The finding that they benefitted from musical experience suggests that musical experience confers a greater benefit for tone perception at lower levels of education than at higher levels of education. Due to the limited sample size, fitting more complex models that take into account socioeconomic status (SES) differences was not possible. However, as more data are collected, the role of SES will be considered in future analyses.

One of the more interesting results of the study was the Vietnamese-English bilinguals outperformed the other three groups even though the source of their advantage was their heritage language. Previous research on Mandarin heritage speakers has found that the quality of their production of tones is different from that of speakers of Mandarin who are not heritage speakers (e.g. Chang & Yao, 2016). Thus, it is likely the case that the Vietnamese heritage speakers in our sample may also deviate in their tone representations relative Vietnamese speakers who were born and are currently living in Vietnam. Even so, it was the Vietnamese heritage speakers' experience with a tonal language that uses contour tones that conferred benefits to their learning of Mandarin tones. The advantage the Vietnamese speakers showed adds to a growing body of work showing that early language experience, even if it is not what might be considered typical "native" language experience (Au et al., 2002; Au et al., 2008, Knightly et al., 2003; Pierce et al., 2014) holds consequences for new language learning particularly in the domain of phonetics.

CHAPTER 4: Conclusions

In this doctoral dissertation, I examined new language learning in adult multilinguals. In the first study (Chapter 2), I investigated the learning of Spanish grammatical gender by Chinese international students who were studying abroad in the US. The goal of that study was to uncover whether the degree of immersion in their non-dominant language (English) and ambient exposure to other languages in their environment would predict new language learning and modulate the relationship between cognitive control resources and learning. For the learning of grammatical gender, higher English exposure predicted higher accuracy. Moreover, greater exposure to additional languages also predicted a more sustained decrease in reaction times during the learning task. Active exposure or use of an additional Chinese language (e.g. Cantonese) modulated the relationship between cognitive control and reaction times for learning and generalization of grammatical gender in Spanish.

In the second study (Chapter 3), I investigated whether prior knowledge of a tone language as an L1, individual differences in cognitive control, and musical experience predicted the ability to perceive tones in a new language. Speakers of tonal native languages (Vietnamese or Bantu languages) and speakers of non-tonal native languages (English or Dutch) were asked to identify Mandarin tones. Critically, Vietnamese has a tone system that is similar to Mandarin because it uses the same two acoustic cues as Mandarin (pitch height and pitch direction). In contrast, Bantu languages are more dissimilar because they only use one of the tonal cues from Mandarin (pitch height). All participants had no prior knowledge of Mandarin, but all were L2 speakers of English, also a non-tonal language. L1 speakers of a tone system that was similar to that of Mandarin (Vietnamese) were more accurate in identifying the Mandarin tones relative to speakers of nontonal languages (Spanish or Dutch). However, L1 speakers of Bantu languages, which are tonal but dissimilar to Mandarin, were less accurate in identifying Mandarin tones compared to speakers of nontonal languages. Higher levels of musical experience also predicted higher Mandarin tone identification, but only for Spanish and Vietnamese groups. Cognitive control abilities, as indexed by performance on the Simon task, did not predict tone identification. Although the two studies differed in particular methodology and general approach, here I focus on how the findings of the two studies inform a broader set of theoretical issues about language learning that were reviewed in the introduction to the thesis.

Was there evidence of plasticity and/or constraints?

Both studies involved brief exposure to a new linguistic feature for an hour or less. In the Spanish grammatical gender study, the Mandarin-English bilinguals became more accurate over the course of the learning task and were well above chance. In the generalization phase, designed to examine the learners' more abstract knowledge of the grammatical gender pattern in Spanish, the learners were less accurate for new than for studied items but were still above chance. The results thus indicate that in general, Mandarin-English bilinguals with no prior knowledge of Spanish or of a language with grammatical gender, were able to learn how to pair the gender-congruent definite article (*ese*/*esa*) with both a familiar and novel noun in Spanish. Although a small subset of participants struggled to learn and to generalize (8%), most were able to learn to pair a gender congruent noun and definite article. The result suggests that adult learners have greater plasticity than might have been predicted by some accounts for learning morphosyntax.

The idea that the ability to acquire morphosyntax is lost beyond a period in early childhood was not supported (Lenneberg, 1967). Moreover, a version of the CPH in which parameters for Universal Grammar were set by the native language (Broselow & Finer, 1991) was also not supported because the learners did not have native knowledge of a grammatically gendered language, but the vast majority were able to learn the grammatical gender pattern in Spanish. The fact that the diversity in the language environments of the learners predicted their accuracy in learning Spanish grammatical gender suggests if there are any maturational constraints for learning grammar (e.g. Hartshorne et al., 2018), they are not fixed. The role of the language environment and of individual differences that reflect adaptations to the language environment suggest that the ability to learn morphosyntax in an L2 is at least partially dependent on the language experience and the language context of the learner. The fact that there were effects of the current language for environment for international student learners, who had relatively limited experience in their current language environment, suggests that the ability to acquire some aspects of morphosyntax in an L2 may to some extent be an emergent property of learners' current language context.

In the tone learning study, the Vietnamese speakers, who use the same tonal cues as Mandarin in their L1, were advantaged in perceiving Mandarin tones relative to nontonal L1 speakers (Spanish, Dutch). On the other hand, Bantu speakers, who only use one of the tonal cues used in Mandarin in their L1, perceived Mandarin tones less accurately than the nontonal L1 speakers. These results suggest that the native language played a critical role in the ability to identify the new Mandarin tones. In this experiment, there appeared to be a limit on Bantu speakers' ability to learn the novel Mandarin tones. This difficulty was interpreted as native language interference because the Bantu speakers were less accurate than both nontonal language groups (Spanish-English and Dutch-English bilinguals). Moreover, although there were other group differences in education level and age relative to the Spanish and Vietnamese speakers, the Bantu speakers were matched on these dimensions to the Dutch group. Even so, the Bantu speakers were less accurate than the Dutch speakers. Thus, the findings from the Experiment 2 seem consistent with the interpretation that Bantu language knowledge produced interference in Mandarin tone perception.

From this one session experiment, it is not clear whether this is interference is a temporary stage in the learning process, which might then be resolved when the learners have more Mandarin input (Flege & Bohn, 2021) or if it is a more permanent constraint that results from the learners being tuned to their native (Kuhl et al., 2008). The results of the two studies suggest that despite there being variability in learning outcomes with generally successful learning, there were also substantial individual differences in learning. Some of these individual differences seem to come from the participants' native language knowledge while some may be due to other factors. The theoretical implications of these individual differences are discussed below.

To what extent did the native language shape the ability to learn a new linguistic feature?

In the first study, the fact that the learners did not have prior knowledge of a language with a grammatical gender system did not preclude them from becoming sensitive to the new grammatical gender pattern in Spanish. It is important to note that the skill that was examined (learning to pair the gender congruent noun with a determinate article) is a relatively simple aspect of grammatical gender to learn, especially in a learning situation in which only nouns with transparent gender marking were included. In contrast, previous research shows that learning to use gender predictively has been found to be possible in L2 learners but to require near-native proficiency (Hopp, 2010). Taken together with previous research, the findings of Experiment 1 suggest that at least for the aspect of grammatical gender that was examined, learners' inexperience with grammatical gender prior to the study did not seem to limit their ability to learn.

In the tone learning study, however, there was evidence of a different pattern of L1 influence, such that the features of the native language could either facilitate or constrain the learners' ability to identify tones in Mandarin. Knowing a tone language that also uses contour tones (Vietnamese) was associated with high accuracy in tone identification relative to not knowing a tone language. On the other hand, knowing a tone language with a different kind of tone system (Bantu) than in the target language resulted in at-chance learning performance.

The results of these two experiments suggest that L1 knowledge can interact with learning in different ways for morphosyntax and phonetics. For morphosyntactic learning, knowledge of the same linguistic feature (even if it is implemented differently) predicts better learning of that feature (Sabourin et al., 2016; Sabourin & Stowe, 2018). The facilitative effect of L1 grammatical gender knowledge on L3 learning of a different grammatical gender system is consistent with L1 knowledge being activated in bilinguals even when processing a different language (Luque, Mizyed, & Morgan-Short, 2018; Sanoudaki & Thierry, 2014; 2015). However, in phonetics, having L1 experience using a linguistic cue (pitch height) seems not simply to activate that cue but to block the perception of another cue (pitch direction) in the L3 (Iverson et al., 2013).

Further evidence that the Bantu group experienced L1 interference comes from the distribution of tone their Mandarin tone identification scores. Despite similar sample sizes, the other three language groups had variable Mandarin tone identification scores, but all the Bantu speakers were at or below chance. Their distribution of scores suggests that knowing a different tone system effectively put a limit on Bantu speakers' ability to track pitch direction. Thus, the findings of Experiment 1 indicate that differences between the L1 and the L3 do not pose a hard constraint on learning for morphosyntax. However, the results of Experiment 2 suggest that when the L1 and L3 share a feature that differs in how it is implemented in the two languages, there appears to be a perceptual constraint on the ability to learn the new L3 categories. Whether the L1 interference effect is temporary or more enduring requires further investigation. A future direction is to conduct study in which participants are asked to complete the Mandarin tone identification task multiple times to track whether the Bantu speakers can learn to track pitch direction when given additional learning opportunities.

Was there a consequence of adaptation to bilingual experience for new learning?

This question was addressed in the grammatical gender learning study. I examined learning in a group of L2-immersed Chinese international students who were living in the US. These bilinguals were hypothesized to recruit proactive control for language processing due to their immersion context. Previous research shows that bilinguals who are L1-immersed recruit more reactive control compared to L2-immersed bilinguals who recruit more proactive control for regulating the activation of their two languages (Beatty-Martinez et al., 2020; Zhang et al., 2021).

I hypothesized that higher proactivity would predict better learning of Spanish grammatical gender. This hypothesis was made on two grounds. The first was that proactivity may facilitate the formation of an abstract rule by exploiting the consistent mapping between the final phoneme [-a, -o] and the definite article in Spanish. This rule formation process may be supported by proactive control since this form of control involves tracking contextual biases to form predictions that guide behavior. Because the mapping was consistent, the input provided in the learning task was geared towards maximally engaging proactive control. Second, proactive control was expected to be more likely to be engaged in the learning process because the international student learners were immersed in their L2 and proactive control has been identified in previous research as the critical component of control for language regulation in that interactional context.

The pattern of results differed for those who spoke Cantonese as their L1 and those who spoke Mandarin as their L1. For learners who actively used Cantonese in the US, greater proactivity predicted faster gender congruency decisions whereas for Mandarin-English bilinguals greater reactivity predicted faster responses. The

Cantonese speakers were technically trilinguals because they also spoke Mandarin and English, the nature of the immersion context made Cantonese a much less commonly spoken language than Mandarin. Relative to Cantonese, Mandarin is much a more commonly spoken language among international student circles. Even though all participants were Chinese international students, the Mandarin-English bilinguals had more opportunities to speak Mandarin than the Cantonese-Mandarin-English trilinguals did to speak Cantonese. The Mandarin-English bilinguals may have recruited reactive control because they were not as deeply immersed in English relative to the Cantonese speakers. These findings are consistent with previous research reporting that bilinguals who are immersed in their L1 recruit reactive control for language processing whereas L2-immersed bilinguals recruit proactive control (Beatty-Martínez et al., 2020; Zhang et al., 2021). The different patterns observed for the two different groups of Chinese speakers seems to reflect that learners' adaptation to their interactional context in US also had consequences for new language learning. The fact that proactive control did not predict better learning for all the international students suggests that cognitive resources may not be universally engaged to learn specific linguistic features. The way in which the learner has adapted to their particular language environment may modulate how cognitive control resources are recruited for new language learning.

There is a growing body of work demonstrating that the interactional context of the bilingual shapes which cognitive control resources come to be recruited for processing the known languages (Beatty-Martínez et al., 2020; Green & Abutalebi, 2013; Hartanto & Yang, 2016; Zhang et al., 2021). Although the demands of the language environment have been investigated in the domain of language processing, there is limited research that has investigated the consequences of the language environment for new language learning. By investigating L3 learning in bilinguals, there was an opportunity to examine how cognitive control dynamics emerging from the way in which the two languages are used may shape language learning mechanisms. One interesting question is whether monolinguals, who do not have to manage the uncertainty of language selection in their language environments in the same way as bilinguals, rely on the same cognitive control mechanisms for new language learning. If so, this would indicate that bilingual language experience and the variation in its various forms provide a unique opportunity to understand how language engages cognition in a way that cannot be examined in monolinguals (Fricke et al., 2019; Kroll et al., 2012).

How do linguistic features engage cognitive resources?

In the present studies, two different measures of cognitive control were included as individual differences predictors. In the grammatical gender learning study, the AX-CPT, a continuous measure of proactive/reactive control was used. In the tone learning study, the Simon task, a measure of reactive control efficiency was used (for the rationale behind selecting these two measures see Chapter 1). In the grammatical gender learning study, there was a relationship between cognitive control (AX-CPT) and gender congruency decision reaction times while there was no such relationship between cognitive control (Simon effect) and tone identification.

The finding that there was no relationship between the Simon effect and tone identification in Mandarin could be due to methodological reasons, such as the timing of the Simon task not being as accurate when collected online compared to in the laboratory. The AX-CPT index used (BSI) was based on accuracy, which might have made it a more sensitive assessment of cognitive control and in turn made it more likely to explain variance in learning outcomes. Methodological issues aside, it is possible that the form of control that was examined in the Simon task, reactive control, was not the primary component involved in L3 tone identification. Previous research suggests that speech perception is predicted by individual differences in reactive control (Lev-Ari & Peperkamp, 2013; 2014). I also hypothesized that reactive control efficiency would translate to better rapid adjustment to new cues. However, the studies examined speech perception not of novel phonetic distinctions but of existing distinctions in noise and in ambiguous lexical contexts. It is possible that the component of cognitive control that is drawn upon for learning new phonetic distinctions is largely dependent on the task demands. This would be consistent with research that examines cognitive control engagement during syntactic processing (Hsu, Jaeggi, & Novick, 2017). It could be the case that reactive control is not the relevant aspect of cognitive control for the learning task.

Another possibility is that non-linguistic cognitive control (reactive, proactive or any other form) is not used to perceive tone. There are undoubtedly attentional mechanisms involved in speech perception but the mechanisms for processing tone might be different due to the nature of tone itself. Tone involves tracking pitch, which is an acoustic cue that is used in both linguistic and nonlinguistic auditory processing. Compared to segmentals (consonants, vowels), tone seems to occupy a neutral space with respect to whether it is processed using domain-general and domain-specific mechanisms. Previous research suggests that pitch perception is jointly determined by musical experience (Wong & Perrachione, 2007) and biological predispositions for pitch perception (Wong et al., 2008). These are all factors that are related to perceptual acuity, rather than to the ability to adjust perception. It may be that nonlinguistic auditory pitch processing benefits more from precision (e.g. in music), whereas the processing of segmentals may require more flexibility due to variation that occurs in speech. The ability to adjust to variation, which helpful for segmentals, may not be helpful for processing pitch in a non-linguistic context. Thus, it is still possible that phonetic learning involves cognitive control (Lev-Ari et al., 2013; 2014) but tone might be an exception.

On the other hand, cognitive control ability was predictive of grammatical gender learning outcomes. This finding suggests the ability to adjust to contextual information and to conflicting information to guide behavior in a nonlinguistic context is drawn upon for learning a novel morphosyntactic feature. Taken together, the results of the two studies suggest that the L3 learners were able to draw upon their cognitive resources for new learning in the morphosyntactic domain. In the phonetic domain, how they learned depended on their prior language knowledge and how that prior linguistic knowledge guided their attentional resources during the Mandarin tone identification task. An important consideration for future research will be to use a larger battery cognitive control tasks to better understand which components of control may be implicated in the learning process. After the relevant components of cognitive control are identified, future research can test for causal relationships by examining whether cognitive training improves tones perception.

To what extent did individual differences explain additional variance in language outcomes?

In the Spanish grammatical gender learning experiment, individual differences in exposure to English (L2) and cognitive control (proactivity/reactivity) accounted for different aspects of learning. Higher levels of English exposure predicted better learning accuracy. The association between English exposure and grammatical gender learning ability was explained in terms of the international students who had more exposure to English having a more linguistically diverse circle of people and potentially due to hearing different varieties of English. On the other hand, reaction times were predicted by the interaction between cognitive control and knowledge of an additional Chinese language. English exposure explained 5% of the variance in learning scores while the interaction between cognitive control and knowledge of an additional Chinese language together explained 16% of the variance in learning reaction times and 14% for generalization reaction times.

In the tone learning experiment, there were two variables that predicted tone learning accuracy: language group and musical experience. Together, these variables accounted for 12% of the variance in Mandarin tone identification.[30] Musical experience score was associated with higher tone identification across all language groups. This result replicates previous research showing that musical experience enhances pitch tracking ability in a linguistic context (Bowles et al., 2016; Wong & Perrachione, 2007).

An assessment of the extent to which the various individual differences of interest explained variability in language learning outcomes shows that at most 16% of

[30] Twelve percent is the marginal r squared value and an additional twenty-five percent of the variance is accounted for by the experimental items used.

variation in learning scores is accounted for. This likely indicates that are many other individual differences that contribute to new language learning. In order to form a more dynamic account of language learning, taking into account the way in which individual differences, some which may stem from language experience or other forms of experience and from biological predispositions, will be important to form a nuanced understand of how humans learn language. The findings of this doctoral dissertation thus suggest language learning mechanisms continue to be available to adults, challenging the claim that there is a limited role for language experience. More research is needed to understand which cognitive resources and experiences are important for learning different aspects of language.

Because this doctoral research took place during the COVID pandemic, I did not have the opportunity to carry out the original plan of examining learning using neural measures (EEG). An important consideration is that techniques that are sensitive to online processing can reveal not only plasticity in learning but also differences in processing that are not evident in behavioral measures. The studies I have reported begin to identify individual differences in behavior, but a goal for future research will be to use neural measures to track learning as a complementary approach to better understand what brain and behavior reveal about language learning in adults.

References

Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, *34*(2), 187-214.

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, *59*(2), 249-306.

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*(4), 557-582.

Au, T. K. F., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, *13*(3), 238-243.

Au, T. K. F., Oh, J. S., Knightly, L. M., Jun, S. A., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory and Language*, *58*(4), 998-1011.

Bakker, I., Takashima, A., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Changes in theta and beta oscillations as signatures of novel word consolidation. *Journal of Cognitive Neuroscience*, *27*(7), 1286-1297.

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, *2*, 324

Bates D, Mächler M, Bolker B, Walker S (2015). “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software*, *67*(1), 1–48.

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

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402-409.

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.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener’s native phonological system. The *Journal of the Acoustical Society of America*, *109*(2), 775-794.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech. In O. S. Bohn & M. J. Munro (Eds.), Language experience in second language speech learning: In honor of James Emil Flege, (pp. 13-34).

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233.

Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, *2*(2), 127-145.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. *NeuroReport*, *26*(16), 966.

Bice, K., & Kroll, J. F. (2019). English only? Monolinguals in linguistically diverse contexts have an edge in language learning. *Brain and Language*, *196*, 104644.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56, 9-49.

Birdsong, D., & Vanhove, J. (2016). Age of second language acquisition: Critical periods and social concerns. In E. Nicoladis and S. Montanari (Eds.), 2016, *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 163-182). De Gruyter.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, *118*(2), 245-257.

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus–Stimulus inhibition. *Bilingualism: Language and Cognition*, *17*(3), 610-629.

Bohn, O. S., & Best, C. T. (2012). Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics*, *40*(1), 109-128.

Bowles, A. R., Chang, C. B., & Karuzis, V. P. (2016). Pitch ability as an aptitude for tone learning. *Language Learning*, *66*(4), 774-808.

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Interlanguage Studies Bulletin*, *7*(1), 35-59.

Brunelle, M. (2009). Tone perception in northern and southern Vietnamese. *Journal of Phonetics*, *37*(1), 79-96.

Caffarra, S., & Barber, H. A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, *1605*, 83-92.

Cassimjee, Farida. 1998. *IsiXhosa tonology: An Optimal Domains Theory analysis*. Munich: Lincom Europa.

Cenoz, J., & Valencia, J. F. (1994). Additive trilingualism: Evidence from the Basque country. *Applied Psycholinguistics*, *15*, 195-195.

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*(6), 520-533.

Chang, C. B., & Mishler, A. (2012). Evidence for language transfer leading to a perceptual advantage for non-native listeners. *The Journal of the Acoustical Society of America*, *132*(4), 2700-2710.

Chang, C. B., & Yao, Y. (2016). Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal, 13*(2), 134-160.

Cheng, B., Zhang, X., Fan, S., & Zhang, Y. (2019). The role of temporal acoustic exaggeration in high variability phonetic training: A behavioral and ERP study. *Frontiers in Psychology*, *10*, 1178.

Chomsky, N. (2009). Syntactic structures. In *Syntactic Structures*. De Gruyter Mouton.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3.

Clements, G. N., & Goldsmith, J. (1984). *Autosegmental studies in Bantu tone: Introduction*. De Gruyter.

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

Conboy, B. T., Sommerville, J. A., & Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, *44*(5), 1505.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283.

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.

Declerck, M., Meade, G., Midgley, K. J., Holcomb, P. J., Roelofs, A., & Emmorey, K. (2021). On the connection between language control and executive control—an ERP study. *Neurobiology of Language*, *2*(4), 628-646.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499-533.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565-7574.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496-518.

Downing, Laura J. 1996. *The tonal phonology of Jita*. Munich: Lincom Europa.

Durlik, J., Szewczyk, J., Muszyński, M., & Wodniecka, Z. (2016). Interference and inhibition in bilingual language comprehension: Evidence from Polish-English interlingual homographs. *PloS One*, *11*(3), e0151430.

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism*: *Language and Cognition*, *10*(1), 101.

Eddington, D. (2002). Spanish gender assignment in an analogical framework. *Journal of Quantitative Linguistics*, *9*(1), 49-75.

Eddington, D., & Hualde, J. I. (2008). El abundante agua fría: Hermaphroditic Spanish Nouns. *Studies in Hispanic & Lusophone Linguistics*, *1*(1), 5-32.

Felton, A., Vazquez, D., Ramos-Nunez, A. I., Greene, M. R., Macbeth, A., Hernandez, A. E., & Chiarello, C. (2017). Bilingualism influences structural indices of interhemispheric organization. *Journal of Neurolinguistics*, *42*, 1-11.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: a magnetoencephalography study. *Developmental Science*, *20*(1), e12427.

FindingFive Team (2019). FindingFive: A web platform for creating, running, and managing your studies in one place. FindingFive Corporation (nonprofit), NJ, USA. https://www.findingfive.com

Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory Phonology*, *9*, 353-382.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge University Press.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults’ acquisition of a second language. Studies in *Second Language Acquisition*, *23*(4), 527-552.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78-104.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*(2), 268-294.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*(1), 529-534.

García-Pentón, L., Fernández, A. P., Iturria-Medina, Y., Gillon-Dowens, M., & Carreiras, M. (2014). Anatomical connectivity changes in the bilingual brain. *Neuroimage*, *84*, 495-504.

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., ... & Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, *84*(3), 318-336.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *Journal of Neuroscience*, *33*(2), 387-396.

Goldsmith, J. (1984). Meeussen's rule. *Language Sound Structure*, 245, 259.

Gooch, D., Thompson, P., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. *Journal of Child Psychology and Psychiatry*, *57*(2), 180-187.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia*.

Granena, G., Yilmaz, Y., & Jackson, D. O. (2016). Cognitive individual differences in second language processing and acquisition*. Cognitive Individual Differences in Second Language Processing and Acquisition*, 1-359.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515-530.

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.

Gullifer, J. W., & Titone, D. (2021). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General*, *150*(3), 414.

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.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*(3), 395-421.

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.

Hisagi, M., Tajima, K., & Kato, H. (2014, May). The effect of language experience on the ability of non-native listeners to identify Japanese phonemic length contrasts. In Proceedings of Meetings on Acoustics 167ASA (Vol. 21, No. 1, p. 060003). Acoustical Society of America.

Hofstetter, S., Friedmann, N., & Assaf, Y. (2017). Rapid language-related plasticity: microstructural changes in the cortex after a short session of new word learning. *Brain Structure and Function*, *222*(3), 1231-1241.

Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: neural responses shared across languages. *Journal of Neuroscience*, *32*(44), 15277-15283.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901-93

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*(4), 572-582.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*(1), B47-B57.

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*(2), 169-178.

Jacobs, A., Fricke, M., & Kroll, J. F. (2016). Cross-language activation begins during speech planning and extends into second language speech. *Language Learning*, *66*(2), 324-353.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60-99.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism*: *Langauge and Cognition*, *5*(3), 225.

Kang, C., Fu, Y., Wu, J., Ma, F., Lu, C., & Guo, T. (2017). Short-term language switching training tunes the neural correlates of cognitive control in bilingual language production. *Human Brain Mapping*, *38*(12), 5859-5870.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705-710.

Kenstowicz, M., & Kisseberth, C. (1990). Chizigula tonology: the word and beyond. *The Phonology-Syntax Connection*, 163-194.

Keshavarz, M. H., & Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language (L3*). International Journal of Bilingual Education and Bilingualism*, *7*(4), 295-302.

Kirby, J. (2010). Dialect experience in Vietnamese tone perception. *The Journal of the Acoustical Society of America*, 127(6), 3749-3757.

Kisseberth, Charles W. 1992. Metrical structure in Zigula tonology. In Derek F. Gowlett

(ed.) *African Linguistic Contributions*, 227–259. Pretoria: Via Afrika

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proceedings of the National Academy of Sciences*, 92(7), 2899-2903.

Kleinman, D., & Gollan, T. H. (2018). Inhibition accumulates over time at multiple processing levels in bilingual language control. *Cognition*, *173*, 115-132.

Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood overhearing. *The Journal of the Acoustical Society of America*, *114*(1), 465-474.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556-6560.

Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science*, *23*(3), 159-163.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, *25*(1), 161-168.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259-274). Springer, Dordrecht.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979-1000.

Lenneberg, E. H. (1967). The biological foundations of language. Hospital Practice, 2(12), 59-67.

Lev-Ari, S. (2018). The influence of social network size on speech perception. *Quarterly Journal of Experimental Psychology*, *71*(10), 2249-2260.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*(5), 320-331.

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *Journal of Phonetics*, *47*, 36-46.

Liberman, A. M., Miyawaki, K., JJ, J., & Fujimura, O. (1973). Cross-Language Study of the Perception of the F_3 Cue for [r] vs.[l] in Speech-and Nonspeech-like Patterns. *The Journal of the Acoustical Society of Japan*, *29*(5), 315.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, 15(3), 651-662.

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*(12), 1507-1515.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861-883.

Linck, J. A., & Weiss, D. J. (2011). Working memory predicts the acquisition of explicit L2 knowledge. *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*, 101-114.

Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, *31*(46), 16808-16813.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism*: *Language and Cognition*, *6*(2), 97.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 34(3), 379-413.

Mechelli, A., Crinion, J. T., Noppeney, U., O'doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757-757.

McCawley, J. D. (1970). Some tonal systems that come close to being pitch accent systems but don't quite make it. In *Chicago Linguistic Society* (Vol. 6, pp. 526-532).

McCawley, J. D. (1978). What is a tone language?. In *Tone* (pp. 113-131). Academic Press.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703-704.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25- 40.

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59-65.

Misra, M., Guo, T., Bobb, S. C., & 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*(1), 224-237.

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, *25*(5), 531-546.

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.

Morgan-Short, K., Hamrick, P., & Ullman, M. T. (2022). Declarative and Procedural Memory as Predictors of Second Language Development. In *The Routledge Handbook of Second Language Acquisition and Individual Differences* (pp. 67-81). Routledge.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, *60*(1), 154-193.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*(4), 933-947.

Myers, S., & Carleton, T. (1996). Tonal transfer in Chichewa. *Phonology*, 13(1), 39-72.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. Language, Cognition and Neuroscience, 29(2), 186-217.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca’s area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263-281.

Oetting, J. B., Hartfield, L. R., & Pruitt, S. L. (2009). Exploring LENA as a tool for researchers and clinicians. *The ASHA Leader*, *14*(6), 20-22

Oh, J. S., Au, T. K. F., Jun, S. A., & Lee, R. M. (2019). Childhood language memory in adult heritage language (re) learners. In M. S. Schmid and B. Kopke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 481-492). Oxford University Press.

Oh, J. S., Jun, S. A., Knightly, L. M., & Au, T. K. F. (2003). Holding on to childhood language memory. *Cognition*, *86*(3), B53-B64.

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

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., ... & Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport-International Journal for Rapid Communications of Research in Neuroscience*, *7*(15), 2439-2444.

Perrachione, T. K., Fedorenko, E. G., Vinke, L., Gibson, E., & Dilley, L. C. (2013). Evidence for shared cognitive processing of pitch in music and language. *PLoS One*, *8*(8), e73372.

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The "Perceptual Wedge Hypothesis" as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, *121*(2), 130-143.

Phinney, M. (1987). The pro-drop parameter in second language acquisition. In *Parameter Setting* (pp. 221-238). Springer, Dordrecht.

Pierce, L. J., Klein, D., Chen, J. K., Delcenserie, A., & Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences*, *111*(48), 17314-17319.

Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9-19.

R Core Team (2022). *R: A language and environment for statistical* *computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Robbins, M. L., Karan, A., López, A. M., & Weihs, K. L. (2018). Naturalistically observing noncancer conversations among couples coping with breast cancer. *Psycho-Oncology*, *27*(9), 2206-2213.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly?. *Second Language Research*, *24*(3), 397-430.

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.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*(2), 101-105.

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.

Santillán, J., & Khurana, A. (2018). Developmental associations between bilingual experience and inhibitory control trajectories in head start children. *Developmental Science*, *21*(4), e12624.

Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics*, *21*(1), 23-44.

Sanz, C., Lin, H. J., Lado, B., Stafford, C. A., & Bowden, H. W. (2016). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics*, 37(5), 669-692.

Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, *18*(3), 242-254.

Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. R. (2013). Musical experience influences statistical learning of a novel language. *The American Journal of Psychology*, *126*(1), 95-104.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281-284.

Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: A brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics*, 35(4), 393-417.

Stroop, J. R. (1935). "Studies of interference in serial verbal reactions". *Journal of Experimental Psychology*. *18*(6): 643–662.

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.

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.

Tomasello, M. (2005). Constructing a language: A usage-based theory of language acquisition. Harvard University Press.

US Census Bureau, & United States. Bureau of the Census. (2003). New York, 2000: 2000 Census of Population and Housing. Summary social, economic, and housing characteristics. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923-927.

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*(6), 1781.

Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, *47*(3), 507-546.

von Bastian, C. C., Locher, A., & Ruflin, M. (2013). Tatool: A Java-Based Open-Source Programming Framework for Psychological Studies. *Behavior Research Methods*, *45*(1), 108-115. doi: 10.3758/s13428-012-0224-y.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, *25*(3), 449-466.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.

Wong, P. C. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin*, *59*(2), 83-95.

Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(4), 565-585.

Wong, P. C., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, *28*(10), 995-1006.

Wong, P. C., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cerebral Cortex*, *18*(4), 828-836.

Xu, M., Baldauf, D., Chang, C. Q., Desimone, R., & Tan, L. H. (2017). Distinct distributed patterns of neural activity are associated with two languages in the bilingual brain. *Science Advances*, *3*(7), e1603309.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America*, *120*(2), 1063-1074.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds: perception of American English /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and language experience: issues in cross-language research* (pp. 305–320). Baltimore, MD: York Press.

Yamada, R. A., & Tohkura, Y. I. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics*, *52*(4), 376-392.

Zhang, H., Diaz, M. T., Guo, T., & Kroll, J. F. (2021). Language immersion and language training: Two paths to enhanced language regulation and cognitive control. *Brain and Language*, 223, 105043.

Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *The Journal of the Acoustical Society of America*, *137*(3), 1452-1463.

Appendix A

Recruitment flyer

Appendix B

Screening survey

What is your date of birth?

Please list ALL languages you have studied either on your own or in a classroom setting. For each one, please list the number of years you have studied each language.

Language 1

Language 2

Language 3

Language 4

Language 5

Language 6

Language 7

Where were you born? (Country, City, and State)

In which language(s) does your family usually speak to you?

At which age did you learn English?

Where are you residing at the moment? Please provide the Country, City, and State where you are PHYSICALLY present right now.

Appendix C

Language History Questionnaire

Start of Block: LEAP-Q 1

Q2 What is your date of birth? (mm/dd/yyyy)

____

Q3 What is your gender?

- ◯ Male (1)
- ◯ Female (2)
- ◯ Other (3) ____

Q41 Enter the zip code of the city where you currently reside in the United States.

____

Q44 In which city and state do you reside in China?

____

Q43 Which university are you currently attending? If you're not attending a university, write NA.

____

Q42 Are you an undergraduate or graduate student?

- ○ Undergraduate student (1)
- ○ Graduate student (2)
- ○ Neither (please explain) (4) ____

Q45 Have you ever taken a language course in school that is not English or Mandarin? If yes, please specify which language(s).

- ○ Yes (2) ____
- ○ No (3)

Q46 Are you currently taking a language course that is not English or Mandarin? If yes, please specify which language(s).

- ○ Yes (1) ____
- ○ No (2)

---

*

Q4 Please list all the languages you know **in order of dominance** (Please specify which language you know, i.e. not Chinese but Mandarin):

- ○ 1 (1) ____
- ○ 2 (2) ____
- ○ 3 (3) ____
- ○ 4 (4) ____
- ○ 5 (5) ____

Q6 Please list what percentage of time you are **currently and on average** exposed to each of the languages you listed.
${Q4/ChoiceTextEntryValue/1} : _______ (1)
${Q4/ChoiceTextEntryValue/2} : _______ (2)
${Q4/ChoiceTextEntryValue/3} : _______ (3)
${Q4/ChoiceTextEntryValue/4} : _______ (4)
${Q4/ChoiceTextEntryValue/5} : _______ (5)
Total : ________

Q9 Please name the cultures with which you identify. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc.)

- ○ Culture 1 (1) ____
- ○ Culture 2 (2) ____
- ○ Culture 3 (3) ____
- ○ Culture 4 (4) ____
- ○ Culture 5 (5) ____

Q10 How many years of formal education do you have? (Count grades 1-12 and any years of completed college.)

____

Q11 Please check your highest education level (or the equivalent to a degree obtained in another country):

- ▢ Less than High School (1)
- ▢ High School (2)
- ▢ Professional Training (3)
- ▢ Some College (4)
- ▢ College (5)
- ▢ Some Graduate School (6)
- ▢ Masters (7)
- ▢ Ph.D./M.D./J.D. (8)
- ▢ Other: (9) ____

Q25 Date when you first started college/university in the US

____

Q26 If you have ever immigrated to another country, please provide name of country and date of immigration here.

____

End of Block: LEAP-Q 1

Start of Block: LEAP-Q 2

Q21 The following questions refer to your knowledge of ${lm://Field/1}.

Q22 Age when you...

○ began acquiring ${lm://Field/1} (1) ____

○ began reading in ${lm://Field/1} (2) ____

Q23 Please list the amount of time you spent in each language environment (specify years, months):

○ A country where ${lm://Field/1} is spoken (1) ____

○ A family where ${lm://Field/1} is spoken (2) ____

○ A school and/or working environment where ${lm://Field/1} is spoken (3) ____

Q24 On a scale from zero to ten, please select your level of **proficiency** in speaking ${lm://Field/1}:

- 0 - none (1)
- 1 - very low (2)
- 2 - low (3)
- 3 - fair (4)
- 4 - slightly less than adequate (5)
- 5 - adequate (6)
- 6 - slightly more than adequate (7)
- 7 - good (8)
- 8 - very good (9)
- 9 - excellent (10)
- 10 - perfect (11)

Q39 On a scale from zero to ten, please select your level of **proficiency** in understanding spoken language ${lm://Field/1}:

- 0 - none (1)
- 1 - very low (2)
- 2 - low (3)
- 3 - fair (4)
- 4 - slightly less than adequate (5)
- 5 - adequate (6)
- 6 - slightly more than adequate (7)
- 7 - good (8)
- 8 - very good (9)
- 9 - excellent (10)
- 10 - perfect (11)

Q38 On a scale from zero to ten, please select your level of **proficiency** in reading ${lm://Field/1}:

- 0 - none (1)
- 1 - very low (2)
- 2 - low (3)
- 3 - fair (4)
- 4 - slightly less than adequate (5)
- 5 - adequate (6)
- 6 - slightly more than adequate (7)
- 7 - good (8)
- 8 - very good (9)
- 9 - excellent (10)
- 10 - perfect (11)

Q25 On a scale from zero to ten, please select how much of the following factors contributed to you learning ${lm://Field/1}:

not a contributor				moderate contributor				most important contributor		
0	1	2	3	4	5	6	7	8	9	10

Interacting with friends ()	
Interacting with family ()	
Reading ()	
Language apps or websites/self instruction ()	
Watching TV ()	
Listening to radio/music ()	

Q40 Please rate to what extent you are currently exposed to ${lm://Field/1} in the following contexts:

	never				half of the time				always		
	0	1	2	3	4	5	6	7	8	9	10
Interacting with friends ()											
Interacting with family ()											
Reading ()											
Language apps or websites/self instruction ()											
Watching TV ()											
Listening to radio/music ()											

End of Block: LEAP-Q 2

Start of Block: Block 2

Q30 What is/are your native language(s) (the first languages you learned)? Please be specific. Don't say Chinese. Specify if it's Mandarin or something else.

____

Q31 At which age did you become exposed to English in China?

____

Q33 Growing up in China, were you ever exposed to languages other than Mandarin? (For example, 广东话/粤语, 吴语, 客家话, 湘语, 闽南语, etc)

◯ Yes (1)

◯ No (2)

Q34 If so, to what extent were you exposed to these non-Mandarin languages?

◯ Frequently/On a daily or weekly basis (1)

◯ A couple of times a month at most (2)

◯ A few times a year/rarely (3)

◯ Never (4)

Q40 In which contexts were you exposed to these non-English languages?

- ☐ Home (1)
- ☐ School (2)
- ☐ Extended Family Members (3)
- ☐ On Television (4)
- ☐ With Friends (5)
- ☐ Online (6)
- ☐ Music (7)
- ☐ Other (8)

End of Block: Block 2

Start of Block: Current language environment

Q48 The following questions ask about the place where you are currently living in the United States.

Q49 How often do you hear languages other than English and Mandarin in the city where you live in the United States?

- ○ Every day (1)
- ○ Once to couple of times a week (2)
- ○ A few times a month (3)
- ○ Almost never (4)
- ○ Never (5)

Q50 How often do you hear languages other than English and Mandarin in your circle of friends?

- ◯ Every day (1)
- ◯ Once to couple of times a week (2)
- ◯ A few times a month (3)
- ◯ Almost never (4)
- ◯ Never (5)

Q51
Make sure that your percentages for this question add up to 100%Please indicate the percentage of your friends in the United States who are...

	0 10 20 30 40 50 60 70 80 90 100
International students from China (or countries where Mandarin is widely used) ()	
International students from countries where Mandarin is not spoken ()	
American students ()	

Q52
Responses for this question DO NOT HAVE TO ADD UP TO 100%How much are interested are you in...

0 10 20 30 40 50 60 70 80 90 100

Learning about other cultures ()	
Learning about other languages ()	
Interacting with people from other cultures ()	
Traveling to different countries ()	

Q53 Please rate how important these factors were in your decision to study abroad in the United States.

	Very Important (1)	Somewhat important (2)	Not that important (3)	Not at all important (4)
The desire to improve my English (1)	○	○	○	○
The desire to get an American education (2)	○	○	○	○
The desire to be immersed in the American culture (3)	○	○	○	○
The desire to experience different cultures (4)	○	○	○	○

End of Block: Current language environment

Appendix D

AX-Continuous Performance Task

Trial Type	Cue	Probe	Number of Trials
AX	A	X	160
AY	A	C	2

AY	A	F	2
AY	A	G	2
AY	A	J	2
AY	A	M	2
AY	A	O	2
AY	A	P	2
AY	A	S	2
AY	A	T	2
AY	A	U	2
BX	C	X	2
BX	F	X	2
BX	G	X	2
BX	J	X	2
BX	M	X	2
BX	O	X	2
BX	P	X	2
BX	S	X	2
BX	T	X	2
BX	U	X	2
BY	C	J	2
BY	F	J	2
BY	G	M	2
BY	G	F	2
BY	M	P	2
BY	O	G	2
BY	P	T	2
BY	S	O	2
BY	U	T	2
BY	S	C	2

Appendix E

Noun Phrase Learning

Studied		Novel	
Masculine	Feminine	Masculine	Feminine
ese brazo	esa soga	ese horno	esa naranja

ese lago	esa vela	ese cerdo	esa obeja
ese vino	esa obeja	ese jugo	esa soga
ese pato	esa secadora	ese brazo	esa abeja
ese baño	esa ventana	ese faro	esa corbata
ese lobo	esa niña	ese gato	esa pistola
ese lazo	esa vaca	ese anillo	esa montaña
ese hongo	esa canasta	ese espejo	esa sandia
ese dado	esa uva	ese globo	esa tortuga
ese gato*	esa escoba*		
ese ojo*			

* practice items

Appendix F

Language Recognition

Question	AudioLanguage	Target Response	Trial Type
Is this Spanish?	Italian	No	Spanish

Is this Italian?	Italian	Yes	Other
Is this Arabic?	Tibetan	No	Other
Is this Spanish?	Italian	No	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this Catalan?	Spanish	No	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this Cambodian?	Cambodian	Yes	Other
Is this Cambodian?	Tagalog	No	Other
Is this Spanish?	Spanish	Yes	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this French?	Spanish	No	Spanish
Is this Finnish?	Finnish	Yes	Other
Is this Russian?	Russian	Yes	Other
Is this Spanish?	Spanish	Yes	Spanish
Is this Persian?	Persian	No	Other
Is this Spanish?	Italian	No	Spanish
Is this Tagalog?	Thai	No	Other
Is this Russian?	Mingrelian	No	Other
Is this Malay?	Indonesian	No	Other
Is this Portuguese?	Portuguese	Yes	Other
Is this Thai?	Thai	Yes	Other
Is this Spanish?	Catalan	No	Spanish
Is this Portuguese?	Spanish	No	Spanish
Is this Romanian?	Romanian	Yes	Other
Is this Italian?	Italian	Yes	Other
Is this Spanish?	Spanish	Yes	Spanish
Is this French?	Italian	No	Other
Is this French?	French	Yes	Other
Is this Spanish?	Spanish	Yes	Spanish
Is this Portuguese?	Italian	No	Other
Is this Italian?	Spanish	No	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this Spanish?	Catalan	No	Spanish
Is this Spanish?	Catalan	No	Spanish
Is this Spanish?	Italian	No	Spanish
Is this Spanish?	Spanish	Yes	Spanish
Is this Russian?	Russian	Yes	Other
Is this Hindi?	Persian	No	Other
Is this Spanish?	Spanish	Yes	Spanish

Is this Arabic?	Arabic	Yes	Other
Is this Vietnamese?	Vietnamese	Yes	Other
Is this Spanish?	Spanish	Yes	Spanish
Is this Arabic?	Arabic	Yes	Other
Is this Burmese?	Burmese	Yes	Other
Is this Thai?	Thai	Yes	Other
Is this Portuguese?	Portuguese	Yes	Other
Is this Spanish?	Spanish	Yes	Spanish

Appendix G

Language and Musical Experience Questionaire

Start of Block: LEAP-Q 1

Q41 What is your Prolific ID?

____

Q2 What is your age?

____

Q3 What is your gender?

- ◯ Male (1)
- ◯ Female (2)
- ◯ Other (3) ____

*

Q4 Please list all the languages you know **starting with the language that you use the most to the one you use the least.** Include any languages that you have previously studied or used in the past even if you do not currently use those languages.

- ◯ 1 (1) ____
- ◯ 2 (2) ____
- ◯ 3 (3) ____
- ◯ 4 (4) ____
- ◯ 5 (5) ____

Page Break

*

Q6 Please list what percentage of time you are **currently and on average** exposed to each of the languages you listed.
${Q4/ChoiceTextEntryValue/1} : _______ (1)
${Q4/ChoiceTextEntryValue/2} : _______ (2)
${Q4/ChoiceTextEntryValue/3} : _______ (3)
${Q4/ChoiceTextEntryValue/4} : _______ (4)
${Q4/ChoiceTextEntryValue/5} : _______ (5)
Total : ________

Q11 Please check your highest education level (or the approximate equivalent to a degree obtained in another country):

- ▢ Less than High School (1)
- ▢ High School (2)
- ▢ Professional Training (3)
- ▢ Some College (4)
- ▢ College (5)
- ▢ Some Graduate School (6)
- ▢ Masters (7)
- ▢ Ph.D./M.D./J.D. (8)
- ▢ Other: (9) ____

Q40 Have you ever had a hearing problem? (hearing loss, cochlear implant, etc). **If yes, please specify.**

- ○ Yes (1) ____
- ○ No (2)

End of Block: LEAP-Q 1

Start of Block: LEAP-Q 2

Q38 Please rate your language skills on a scale from 0 to 10 for each of the languages that you know. (0 = None, 1 = Very Low, 2 = Low, 3 = Fair, 4 = Slightly Less Than

Adequate, 5 = Adequate, 6 = Slightly More Than Adequate, 7 = Good, 8 = Very Good, 9 = Excellent, 10 = Perfect)

	Speaking (1)	Understanding (2)	Reading (5)	Writing (6)
⨂ ${Q4/ChoiceTextEntryValue/1} (1)				
⨂ ${Q4/ChoiceTextEntryValue/2} (2)				
⨂ ${Q4/ChoiceTextEntryValue/3} (3)				
⨂ ${Q4/ChoiceTextEntryValue/4} (4)				
⨂ ${Q4/ChoiceTextEntryValue/5} (5)				

Q39 Please provide the age when you first began learning each of the languages that you know. For your native language(s), enter 0.

	Age (1)
${Q4/ChoiceTextEntryValue/1} (1)	
${Q4/ChoiceTextEntryValue/2} (2)	
${Q4/ChoiceTextEntryValue/3} (3)	
${Q4/ChoiceTextEntryValue/4} (4)	
${Q4/ChoiceTextEntryValue/5} (5)	

End of Block: LEAP-Q 2

Start of Block: Musical Experience

Q24 Have you ever played a musical instrument?

○ Yes (1)

○ No (2)

Display This Question:

If Have you ever played a musical instrument? = Yes

Q25 How many instruments have you played?

____

Display This Question:

If Have you ever played a musical instrument? = Yes

*

Q42 Please list all the musical instruments you know **starting with the musical instrument that you play the most to the one you play the least.** Include any musical instrument that you have previously played in the past even if you do not currently play those musical instruments.

○ 1 (1) ____

○ 2 (2) ____

○ 3 (3) ____

○ 4 (4) ____

○ 5 (5) ____

Page Break

Display This Question:

If Have you ever played a musical instrument? = Yes

Q43 At what age did you learn to play a musical instrument?

	Age (1)
${Q42/ChoiceTextEntryValue/1} (1)	
${Q42/ChoiceTextEntryValue/2} (2)	
${Q42/ChoiceTextEntryValue/3} (3)	
${Q42/ChoiceTextEntryValue/4} (4)	
${Q42/ChoiceTextEntryValue/5} (5)	

Display This Question:

If Have you ever played a musical instrument? = Yes

Q44 How long have you been playing a musical instrument?

	Years (1)
${Q42/ChoiceTextEntryValue/1} (1)	
${Q42/ChoiceTextEntryValue/2} (2)	
${Q42/ChoiceTextEntryValue/3} (3)	
${Q42/ChoiceTextEntryValue/4} (4)	
${Q42/ChoiceTextEntryValue/5} (5)	

Page Break

Display This Question:

If Have you ever played a musical instrument? = Yes

Q28 Do you currently play a musical instrument?

- ◯ Yes (1)
- ◯ No (2)

Display This Question:

If Do you currently play a musical instrument? = Yes

Q45 How often do you practice this musical instrument currently? (Ex: Once a week, Three Times a Week, Twice a Month, etc)

- ◯ ${Q42/ChoiceTextEntryValue/1} (1) ____
- ◯ ${Q42/ChoiceTextEntryValue/2} (2) ____
- ◯ ${Q42/ChoiceTextEntryValue/3} (3) ____
- ◯ ${Q42/ChoiceTextEntryValue/4} (4) ____
- ◯ ${Q42/ChoiceTextEntryValue/5} (5) ____

Display This Question:

If Do you currently play a musical instrument? = Yes

Q46 When was the last time you practiced this instrument? (Ex: yesterday, 2 weeks ago, 5 years ago, etc)

○ ${Q42/ChoiceTextEntryValue/1} (1) ____

○ ${Q42/ChoiceTextEntryValue/2} (2) ____

○ ${Q42/ChoiceTextEntryValue/3} (3) ____

○ ${Q42/ChoiceTextEntryValue/4} (4) ____

○ ${Q42/ChoiceTextEntryValue/5} (5) ____

Q31 Can you read music?

○ Yes (1)

○ Some (2)

○ No (3)

Q32 Do you write or compose music?

○ Yes (1)

○ Some (2)

○ No (3)

Q33 Do you teach or have taught others any instrument?

- ○ Yes (1)
- ○ No (2)

Q34 Have you ever taken formal courses in music theory?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Have you ever played a musical instrument? = Yes

Q35 Have you ever had formal, private instruction to learn an instrument?

- ○ Yes (1)
- ○ No (2)

Q36 Have you ever played an instrument (or sang) in an organized group (band, orchestra, etc)?

- ○ Yes (1)
- ○ No (2)

Page Break

Q47
Please translate this audio in English.

____

End of Block: Musical Experience

Appendix H

Pitch Identification stimuli. Each stimulus name reflects the two choices that were visually presented to the participant to the left of the | and to the right the audio clip. The name of the audio clip says the vowel [a, e, i, o, u] and the code of the speaker (f1 = female speaker 1).

tone1-tone2|a_1f1	tone2-tone1|a_1f1	tone4-tone1|a_1f1
tone1-tone2|a_1f2	tone2-tone1|a_1f2	tone4-tone1|a_1f2
tone1-tone2|a_1m1	tone2-tone1|a_1m1	tone4-tone1|a_1m1
tone1-tone2|a_1m2	tone2-tone1|a_1m2	tone4-tone1|a_1m2
tone1-tone2|a_2f1	tone2-tone1|a_2f1	tone4-tone1|a_4f1
tone1-tone2|a_2f2	tone2-tone1|a_2f2	tone4-tone1|a_4f2
tone1-tone2|a_2m1	tone2-tone1|a_2m1	tone4-tone1|a_4m1
tone1-tone2|a_2m2	tone2-tone1|a_2m2	tone4-tone1|a_4m2
tone1-tone2|e_1f1	tone2-tone1|e_1f1	tone4-tone1|e_1f1
tone1-tone2|e_1f2	tone2-tone1|e_1f2	tone4-tone1|e_1f2
tone1-tone2|e_1m1	tone2-tone1|e_1m1	tone4-tone1|e_1m1
tone1-tone2|e_1m2	tone2-tone1|e_1m2	tone4-tone1|e_1m2
tone1-tone2|e_2f1	tone2-tone1|e_2f1	tone4-tone1|e_4f1
tone1-tone2|e_2f2	tone2-tone1|e_2f2	tone4-tone1|e_4f2
tone1-tone2|e_2m1	tone2-tone1|e_2m1	tone4-tone1|e_4m1
tone1-tone2|e_2m2	tone2-tone1|e_2m2	tone4-tone1|e_4m2
tone1-tone2|i_1f1	tone2-tone1|i_1f1	tone4-tone1|i_1f1
tone1-tone2|i_1f2	tone2-tone1|i_1f2	tone4-tone1|i_1f2
tone1-tone2|i_1m1	tone2-tone1|i_1m1	tone4-tone1|i_1m1
tone1-tone2|i_1m2	tone2-tone1|i_1m2	tone4-tone1|i_1m2
tone1-tone2|i_2f1	tone2-tone1|i_2f1	tone4-tone1|i_4f1
tone1-tone2|i_2f2	tone2-tone1|i_2f2	tone4-tone1|i_4f2
tone1-tone2|i_2m1	tone2-tone1|i_2m1	tone4-tone1|i_4m1
tone1-tone2|i_2m2	tone2-tone1|i_2m2	tone4-tone1|i_4m2
tone1-tone2|o_1f1	tone2-tone1|o_1f1	tone4-tone1|o_1f1
tone1-tone2|o_1f2	tone2-tone1|o_1f2	tone4-tone1|o_1f2
tone1-tone2|o_1m1	tone2-tone1|o_1m1	tone4-tone1|o_1m1
tone1-tone2|o_1m2	tone2-tone1|o_1m2	tone4-tone1|o_1m2
tone1-tone2|o_2f1	tone2-tone1|o_2f1	tone4-tone1|o_4f1
tone1-tone2|o_2f2	tone2-tone1|o_2f2	tone4-tone1|o_4f2
tone1-tone2|o_2m1	tone2-tone1|o_2m1	tone4-tone1|o_4m1
tone1-tone2|o_2m2	tone2-tone1|o_2m2	tone4-tone1|o_4m2
tone1-tone2|u_1f1	tone2-tone1|u_1f1	tone4-tone1|u_1f1
tone1-tone2|u_1f2	tone2-tone1|u_1f2	tone4-tone1|u_1f2
tone1-tone2|u_1m1	tone2-tone1|u_1m1	tone4-tone1|u_1m1
tone1-tone2|u_1m2	tone2-tone1|u_1m2	tone4-tone1|u_1m2

tone1-tone2|u_2f1	tone2-tone1|u_2f1	tone4-tone1|u_4f1
tone1-tone2|u_2f2	tone2-tone1|u_2f2	tone4-tone1|u_4f2
tone1-tone2|u_2m1	tone2-tone1|u_2m1	tone4-tone1|u_4m1
tone1-tone2|u_2m2	tone2-tone1|u_2m2	tone4-tone1|u_4m2
tone1-tone4|a_1f1	tone2-tone4|a_2f1	tone4-tone2|a_2f1
tone1-tone4|a_1f2	tone2-tone4|a_2f2	tone4-tone2|a_2f2
tone1-tone4|a_1m1	tone2-tone4|a_2m1	tone4-tone2|a_2m1
tone1-tone4|a_1m2	tone2-tone4|a_2m2	tone4-tone2|a_2m2
tone1-tone4|a_4f1	tone2-tone4|a_4f1	tone4-tone2|a_4f1
tone1-tone4|a_4f2	tone2-tone4|a_4f2	tone4-tone2|a_4f2
tone1-tone4|a_4m1	tone2-tone4|a_4m1	tone4-tone2|a_4m1
tone1-tone4|a_4m2	tone2-tone4|a_4m2	tone4-tone2|a_4m2
tone1-tone4|e_1f1	tone2-tone4|e_2f1	tone4-tone2|e_2f1
tone1-tone4|e_1f2	tone2-tone4|e_2f2	tone4-tone2|e_2f2
tone1-tone4|e_1m1	tone2-tone4|e_2m1	tone4-tone2|e_2m1
tone1-tone4|e_1m2	tone2-tone4|e_2m2	tone4-tone2|e_2m2
tone1-tone4|e_4f1	tone2-tone4|e_4f1	tone4-tone2|e_4f1
tone1-tone4|e_4f2	tone2-tone4|e_4f2	tone4-tone2|e_4f2
tone1-tone4|e_4m1	tone2-tone4|e_4m1	tone4-tone2|e_4m1
tone1-tone4|e_4m2	tone2-tone4|e_4m2	tone4-tone2|e_4m2
tone1-tone4|i_1f1	tone2-tone4|i_2f1	tone4-tone2|i_2f1
tone1-tone4|i_1f2	tone2-tone4|i_2f2	tone4-tone2|i_2f2
tone1-tone4|i_1m1	tone2-tone4|i_2m1	tone4-tone2|i_2m1
tone1-tone4|i_1m2	tone2-tone4|i_2m2	tone4-tone2|i_2m2
tone1-tone4|i_4f1	tone2-tone4|i_4f1	tone4-tone2|i_4f1
tone1-tone4|i_4f2	tone2-tone4|i_4f2	tone4-tone2|i_4f2
tone1-tone4|i_4m1	tone2-tone4|i_4m1	tone4-tone2|i_4m1
tone1-tone4|i_4m2	tone2-tone4|i_4m2	tone4-tone2|i_4m2
tone1-tone4|o_1f1	tone2-tone4|o_2f1	tone4-tone2|o_2f1
tone1-tone4|o_1f2	tone2-tone4|o_2f2	tone4-tone2|o_2f2
tone1-tone4|o_1m1	tone2-tone4|o_2m1	tone4-tone2|o_2m1
tone1-tone4|o_1m2	tone2-tone4|o_2m2	tone4-tone2|o_2m2
tone1-tone4|o_4f1	tone2-tone4|o_4f1	tone4-tone2|o_4f1
tone1-tone4|o_4f2	tone2-tone4|o_4f2	tone4-tone2|o_4f2
tone1-tone4|o_4m1	tone2-tone4|o_4m1	tone4-tone2|o_4m1
tone1-tone4|o_4m2	tone2-tone4|o_4m2	tone4-tone2|o_4m2
tone1-tone4|u_1f1	tone2-tone4|u_2f1	tone4-tone2|u_2f1
tone1-tone4|u_1f2	tone2-tone4|u_2f2	tone4-tone2|u_2f2
tone1-tone4|u_1m1	tone2-tone4|u_2m1	tone4-tone2|u_2m1
tone1-tone4|u_1m2	tone2-tone4|u_2m2	tone4-tone2|u_2m2
tone1-tone4|u_4f1	tone2-tone4|u_4f1	tone4-tone2|u_4f1
tone1-tone4|u_4f2	tone2-tone4|u_4f2	tone4-tone2|u_4f2
tone1-tone4|u_4m1	tone2-tone4|u_4m1	tone4-tone2|u_4m1

tone1-tone4|u_4m2 tone2-tone4|u_4m2 tone4-tone2|u_4m2