Pattern Learning and Accent Familiarity in Monolingual and Bilingual Speakers:

Generalizing Learning Across Accented Speech

by

Samara St. Louis

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Speech Language Pathology

Department of Communication Sciences and Disorders

University of Alberta

© Samara St. Louis, 2023

Abstract

It is increasingly common to encounter speakers with an accented variety of English, especially as society becomes more diverse and multilingual. To begin shedding light on how to improve communication outcomes when interacting with a speaker who has an accent, this research investigates individual cognitive factors which may enhance the ability to process accented speech. These factors include statistical learning ability, accent familiarity, language background, and the relationships therein. *Statistical learning* is the ability to extract regularities from the environment and is highly implicated in processing language. A novel auditory statistical learning (aSL) task with accented phonemes was administered to 43 participants recruited from the University of Alberta, alongside a visual statistical learning task, language background questionnaire, and accent familiarity questionnaire. Linear regression analysis was used to examine the differences in performance between English L1 and L2 participants on the aSL task. It was found that the English L1 participants had a faster reaction time overall, though the English L2 participants were more correct overall. Accent familiarity was a significant predictor of reaction time, especially for L2 participants. Performance on the visual statistical learning task was the only significant predictor of overall aSL score; however, more nuanced analysis showed differences in performance by question type between the two groups.

The findings contribute to understanding how bilingualism can affect the processing of accented phonemes, how statistical learning and bilingualism are related, and how exposure to a variety of accents may be just as beneficial as exposure to a particular accent when processing accented speech. This thesis also provides a lasting contribution to the field by providing a new auditory statistical learning task and Accent Familiarity Questionnaire with preliminary reliability and validity. It also provides an essential foundation for future research.

Preface

This thesis is an original work by Samara St. Louis. Ethics approval was received for this research and its associated methods from the University of Alberta Research Ethics Board 2, Project Name "Pattern learning in monolingual and bilingual English speakers: Generalizing learning across accented speech," Pro00117511, on April 20, 2022. All the work described herein was conducted in the Cognitive Neuroscience of Reading and Speech Production Lab at the University of Alberta.

The research described in this work will contribute to a scholarly publication, with the manuscript currently in preparation. The relative contributions of each co-author are outlined below.

St. Louis, S., Aalto, D., Hodgetts, B., Cummine, J., Gerum, R., & Nisbet, K. (2023). Statistical learning of accented speech in English L1 and L2 speakers. *Cognitive Science.* [Manuscript in preparation].

I was the primary investigator for this research and was responsible for the literature review, stimuli and task creation, study design, data collection, preliminary data analysis, and manuscript preparation and revisions. Daniel Aalto and Bill Hodgetts contributed to study design, study oversight, and manuscript revisions. Jacqueline Cummine contributed her lab space and resources and provided indirect oversight, and Richard Gerum contributed to task coding and programming. Finally, Kelly Nisbet was involved in study design, study oversight, data analysis, and manuscript editing and revisions. The proportion of research conducted by the student, Samara St. Louis, is approximately 60%, while the proportion of writing is 90%.

Acknowledgements

First and foremost, I would like to acknowledge the incredible support and wisdom of my supervisor, Kelly Nisbet. I have been so fortunate to work with you over the last year and a half and feel truly blessed to have had you as a mentor and supervisor. The fact that I always left our meetings feeling more encouraged than before speaks to your remarkable capacity as a mentor and supporter. Thank you for always advocating for me, providing your perspectives, sharing your vast knowledge, and helping me see this project through to completion.

I would also like to thank and acknowledge my committee members, Daniel Aalto and Bill Hodgetts. Thank you for the insightful conversations, for your enthusiasm towards this project, for making me feel like a colleague, and for always sharing your wisdom, whether through literature suggestions, help with Praat and the technical details, and insight into the research process. Thank you also to my co-author, Richard Gerum, for your invaluable assistance with programming my experiment in PsychoPy. I can't express how exciting and gratifying it was to hear the stimuli I had spent so long creating come to life as a fully functioning SL task. Thank you to Kay as well for your coding expertise in finalizing my stimuli. Lastly, thank you to co-author Jacqueline Cummine for the use of your lab space and resources.

I am grateful for the provision of funding throughout my graduate studies from the University of Alberta, and for the opportunity to have a research assistant position as well. I have also received funding from the Government of Canada through the Tri-Council Canada Graduate Master's Scholarship.

To everyone in the Cummine-Hodgetts lab, thank you for your support in various ways. Amber, your help has been invaluable throughout this project. Thank you for your assistance in writing my ethics application, for handling so many of the behind-the-scenes tasks which made this experiment run smoothly, and for always being willing to answer my many questions, even during Audiology lab. To my lovely research assistant Amira, thank you for your enthrallment with this project and for your dedicated help scheduling and running participants; it was a pleasure to share this research with you and have so many insightful conversations about it. A special thanks also to Caitlyn Sharp and Natalie Hanas for showing me the ropes of running participants in CSB, to Siobhan Falk and Dania Awartani for their insight in recruiting and scheduling participants, and to Hala Moraich and Shelley Richmond for always ensuring I had keys and an audiometer when I needed.

I truly would not have been able to complete this thesis without the incredible support of my family and friends. Mom, thank you for always checking in and encouraging me in my work with my 'Sis', and for your support through the final push of writing. Words can't even express how grateful I am for everything you've done to help me. Dad, thank you for matching your renovation efforts to my thesis writing, and for always being so proud even if you couldn't ever quite describe what my research was about. I owe all my work ethic and perseverance to you. To my friends, thank you for your understanding and for giving me the space and time to work on this thesis over the last year and a half, and for always cheering me on and celebrating with me.

Finally, I am indebted to the help and support of my fiancé, Alex, and truly would not be here without him. Thank you for being so steadfast throughout my master's degree, even while I was living away in Edmonton, and for helping with this thesis in any way possible, from working alongside me on virtual study dates to helping ensure my APA citations were formatted perfectly, and everything in between. I am beyond grateful for your constant encouragement and for how you've been right alongside me throughout all the ups and downs of this project. We did it!

Table of Contents

List of Tables

List of Figures

INTRODUCTION

The linguistic landscape of Canada is incredibly diverse, and it is expected to continue to grow and diversify further. With newcomers from around the world choosing to call Canada home comes a marked increase in speakers whose first language is not English. Hence, accented varieties of Canadian English are becoming increasingly ubiquitous, such that it is remarkably common to encounter a speaker with an accent that differs from Canadian English. However, across Canada, individuals differ on their level of familiarity with and exposure to the variety of accents represented; individuals also differ on cognitive factors which may affect the processing of accented speech. This can result in miscommunication or frustration by both the speaker and listener, especially when a listener is not familiar with a particular accent or dialect. What facilitates the ability to understand a speaker with an unfamiliar accent? This research explores the factors which impact an individual's ability to process accented speech, by examining pattern recognition ability, familiarity with accents, and language background.

Motivation

There is no doubt that Canada's linguistic diversity is increasing, as evidenced by the rates of immigration in recent years. According to Statistics Canada (2017), there were 1.2 million immigrants to Canada between 2011 and 2016, which is 3.5% of the total population. Given the global events of the last few years, this number is projected to be much higher at present. For those who have a first language (L1) other than English, becoming proficient in English as a second language (L2) or additional language (LX), is of crucial importance upon arrival in Canada (Dewaele, 2018). As of 2016, the number of people in Canada who do not have English or French as their first language is roughly 7.3 million, or 21% of the entire population (Statistics Canada, 2017). Compared to 2011, this is an addition of nearly one million people who do not speak English as their L1. As these new Canadians learn to become more proficient in English as an LX, they are likely to speak with an accent to some degree (Grosjean, 2010). Furthermore, the majority of immigrants who have arrived in Canada in recent years have been over the age of 15 (Statistics Canada, 2017), meaning they are past the critical period of speaking English without an accent; research has suggested that age 9-12 is the approximate cut-off for unaccented speech in a second language (Dollmann et al., 2020). Altogether, the likelihood of encountering a speaker in Canada with an accent different to one's own is high, due to the increasing rates of immigration and the variety of speakers of English as a second language.

It is established that differences in pronunciation from accented speech are a factor that can lead to breakdowns in communication between a speaker and listener (Grant & Brinton, 2014). The presence of a foreign accent can increase processing time of a sentence, and the divergences in phonetics, phonology, lexicology, or discourse can lead to information being lost (Adank et al. 2009; Floccia et al., 2009; Munro & Derwing, 1995a; 1995b). Further, having an accent affects not only whether one is understood, but also how one is judged or perceived socially. In a meta-analysis conducted by Fuertes et al. (2012), use of a standard accent (i.e. the standard, institutionalized variety of a dominant language with prestige in a society) was associated with much higher ratings of status, dynamism, and solidarity than use of a non-standard accent. For a discussion on the nuance in what is considered a standard variety, why certain accents hold prestige, and the variation in prestige, see Moyer, 2013, and Lippi-Green, 2011. From the perspective of the speaker with an accent, having an accent is significantly associated with feelings of less belonging, expectations of social stigma, and perceived problems in communication (Gluszek & Dovidio, 2010). In an interview of adult English as a Second Language (ESL) students in Canada, Derwing (2003) found that one-third had felt discriminated against because of their accent, and the majority of students felt that they would be more respected if they did not speak with an accent. Indeed, many accents are subject to stigmatization and stereotyping in North American society, with ramifications of linguistic profiling, discriminatory housing practices, biased hiring, political disparity, and over-representation in the legal system (see Lippi-Green, 2011; Baugh, 2003; Hansen & Dovidio, 2016; Kang & Yaw, 2021). These biases can result in a disadvantage for the many speakers of English as an additional language (LX). Hence, there exists a clear need on the part of the listener to mitigate some of this discrepancy. While the sociological factors surrounding accented speech in society are beyond the scope of this study, and there are many complexities surrounding communication across different accents, there may be specific underlying cognitive factors which could influence how accented speech is understood. Investigating these potential cognitive influences is a critical first step towards tackling this larger societal issue.

In this study, we examine the role of individual cognitive factors that may enable a listener to better process accented speech. In communication interactions between an L1 English listener and an LX English speaker, the onus for mutual understanding tends to disproportionately fall on the speaker. The reason for this lack of effort may stem from a variety of sociological factors, but could be partially modulated by certain individual cognitive factors. As a listener, there are several factors which may be able to shift the responsibility for effective communication to a more equal distribution between L1 and LX interlocutors. Aside from overtly challenging some of the biases that can lead to communication breakdowns, these implicit factors include familiarity with the accent being spoken, language background, and language-related pattern recognition ability. Examining these variables provides valuable insight in and of itself, but also speaks to some of the mechanisms which underlie effective communication between L1 and LX English speakers. Herein lies the motivation for the present study: Given the increase in Canadian LX speakers with accents and the social judgements, isolation, and communication barriers they face due in part to accented speech, what cognitive mechanisms of the listener impact attenuation and generalization across different forms of accented speech? If an individual is more familiar with the accent being spoken, how does that facilitate the processing of accented speech? Could speaking another language lead to greater cognitive flexibility in understanding accented speech? Might better pattern-recognition ability lead a listener to pick up on the variation in LX speech more effectively? This study explores some of the factors that may underlie difficulty in the listener's accent perception and generalization abilities. Consequently, mitigating these factors could lead to better navigation of communication breakdown across the groups represented in Canadian society.

Background

Accent Familiarity

Defining Accents

Given the importance of understanding LX speakers with an accent as outlined above, it bears discussing what exactly constitutes accented speech. Here, *accent* is described from the listener's perspective, such that the speaker's phonetic, phonological, and prosodic features differ from what the listener would expect from an L1 speaker (Christia et al., 2012). This does not include differences stemming from dialect or stylistic pronunciation differences (for a discussion on the difference between *accent, dialect,* and *pronunciation,* see Moyer, 2013). Second, accented speech is often measured in terms of intelligibility, comprehensibility, and accentedness. Intelligibility refers to the acoustic-phonetic cues affecting the perception of individual words, and it is often measured through deviations between a listener's transcription and the speech sample. Comprehensibility refers to perceived intelligibility, or how easily an overall message or utterance is understood (Kennedy & Trofimovich, 2008; Moyer, 2013; Munro & Derwing, 1995a). Accentedness is a perceptual judgment of the degree or strength of a foreign accent and is often rated more harshly than comprehensibility and intelligibility (Derwing & Munro, 1997). Intelligibility and accentedness are often related, but they are not necessarily correlated with comprehensibility, as a strong accent can still be perfectly understandable (Derwing &

Munro, 1997; Shintani et al., 2019; for a review of the current field of research in this area, see Poljak, 2019).

While the hierarchy of which features most impact intelligibility in accented speech is not universally agreed upon (see Munro and Derwing, 1995a), there are several salient features that constitute a foreign or regional accent. Schairer (1992) maintains that vowel quality is a significant factor affecting intelligibility, which is concurrent with recent research (see Nihalani, 2010). Accurate production of consonants is a similarly important factor, especially for North American listeners (see Riney et al., 2005), followed by phonological processes such as elision or assimilation. Of critical importance for accurate production of consonants are the factors of voice onset time (VOT) and segment duration (Setter, 2006). Across speakers of various language backgrounds, VOT for word- and syllable-initial consonants leads to differences in categorical perception, which can greatly influence intelligibility (Moyer, 2013). All these phonetic and phonological characteristics have regular distributional and transitional patterns within each language, and differences between L1 and L2 patterns can result in what is perceived as accented speech (Moyer, 2013). There are prosodic features such as stress and intonation that characterize an accent as well, but this research focuses on the phonetic features described above, as they have a more significant impact on intelligibility (Jenkins, 2002). These features will be discussed further in the methods section due to their critical role in stimuli creation.

Accent Familiarity

Accent familiarity is an important factor related to the intelligibility and processing of accented speech. Lack of familiarity with an accent can lead to longer processing times of sentences spoken in that accent on a variety of tasks (Adank et al., 2009; Kang, 2012; Winke & Gass, 2013; see also Christia et al., 2012). Research by Kennedy and Trofimovich (2008) found that participants who worked as ESL teachers and had experience hearing accented L2 speech were more accurate in transcribing Mandarin-accented speech compared to those who did not have experience with L2 speech. These effects extended beyond semantic context, suggesting that experience with the greater variability in L2 speech leads to more effective decoding. Interestingly, both the experienced and non-experienced listeners rated the degree of accentedness in the stimuli to be the same. The effect of familiarity and language experience was identified even 40 years ago, when Gass and Varonis (1984) demonstrated that a listener's familiarity with the accent, and L2 speech in general, was highly correlated with their accuracy in transcribing and summarizing accented speech. They found language experience affected judgments of comprehensibility as well. Much research has since corroborated the findings that exposure and familiarity modulate perceptual challenges with accented speech (Hanulíková & Weber, 2012; Levy et al., 2019; Porretta et al., 2016; Sumner & Samuel, 2009). Huyck et al. (2017) showed that even when the speech signal is degraded, familiarity with accents played a role in how much of the speech was learned, measured by words identified correctly (see also Adank et al., 2009). Lastly, Bradlow and Bent (2008) determined that exposure to multiple speakers of Mandarin aided L1 English participants in understanding a novel Mandarin speaker. There was no advantage for these participants when listening to an unfamiliar Slovakian speaker, though, suggesting that increased understanding of an LX speaker from a certain language background comes from prior exposure to that language background specifically.

Examining the role of familiarity in understanding accented speech further, research has shown that having the same L1 background as the speaker with an accent can be advantageous in understanding (see Shintani et al., 2019). In an ERP study examining vowels, Conrey et al. (2005) showed that listeners are more attuned, and have a different neurological response, to phonological features present in their own L1 when processing a foreign accent. At a fine acoustic level, phonemes are processed differently depending on language background. The phonetic inventory of an individual's L1 determines which acoustic cues are utilized when processing phonemes, and having a shared L1 when listening to accented speech can lead to more efficient processing (Wagner et al., 2006). In general, phonetic similarity between a listener's L1 and a speaker's LX speech has been found to aid in processing, which is why a shared L1 can be advantageous (Major et al., 2002; Winke & Gass, 2013). In sum, it is established that familiarity with a certain accent, due either to exposure or a shared L1, can improve communicative outcomes between speaker and listener. However, the mechanism underlying this advantage is still not fully understood.

Bilingualism

Bilingualism and Understanding Accents

To what extent does fluently speaking another language impact the ability to process and understand accented speech, whether through a shared L1 or otherwise? The interplay between bilingualism and the aforementioned dimensions of familiarity and shared language background is nuanced; however, research has shown that bilingual individuals may have an advantage in processing accented speech. A study by Saito and Shintani (2015) compared bilingual listeners from Singapore and monolingual listeners from Canada in their comprehensibility ratings of unfamiliar Japanese-accented English speech. The researchers found that the bilingual individuals performed much better in understanding the foreign-accented speech, which was likely due to increased exposure to phonetic variation in English compared to the monolingual Canadians. These bilingual individuals also used more strategies for understanding the accented speech, relying on lexicogrammatical cues instead of pronunciation alone. A follow-up study by Shintani et al. (2019) expounds on these findings, noting that many of the Singapore bilinguals were fluent in Mandarin, which is similar to Japanese—underscoring the importance of both language background and accent familiarity. Additionally, those who had more English exposure in their early childhood had more difficulty in comprehending the accented speech, again lending support to the role of varied language input, which often comes from growing up in a bilingual setting. Grey et al. (2019) offers an alternative explanation to the advantage in comprehension that bilingual listeners seem to have. In an EEG study of Dutch-English bilinguals listening to Chinese-accented and American-accented English, the bilingual listeners still rated foreign Chinese-accented English to be highly comprehensible. The authors suggest a "good enough" strategy when bilinguals process foreign-accented speech, based on the lack of neural response to slight grammatical errors. This is likely related to the high degree of variability in English that many multilinguals are accustomed to. Grey et al. (2019) also emphasize the impact of familiarity with an accent, even for a bilingual speaker.

High variability in phonetic input, and the related cognitive mechanisms for addressing such variability, seems to be a crucial factor for bilingual individuals in processing accented speech. Greater exposure to multiple accents, not just multiple speakers of an accent, can lead to faster processing of accented speech (Baese-Berk et al., 2013). For a bilingual individual, dealing with variability is inherent to learning multiple languages, particularly in differences of categorical perception. Indeed, a study by Archila-Suerte et al. (2012) found that early bilingual adults, who had acquired their L2 before age 5, performed very well on both within- and between-categorization of L2 sounds (i.e., within-categorization being the identification of phonetic exemplars to be members of the same phoneme, like /p/, and between-categorization being the identification of exemplars to be different phonemes, like /b/ and /p/). However, bilinguals who had acquired their second language later were only able to accurately identify between-category membership of L2 sounds. It appears that as an early bilingual, being accustomed to hearing multiple slightly different pronunciations of the same phoneme is associated with less difficulty in categorizing these phonemes as the same, which can lead to increased proficiency in understanding accented speech.

Levy et al. (2019) also comments on the connection between phonetic categories and bilingualism, saying that bilingual individuals "may have an advantage at processing unfamiliar accents because of their greater flexibility at mapping variable input onto mental representations, and their ability to deal with the phonology of two languages with different phonetic boundaries" (p. 372). This heightened ability of bilinguals in matching unfamiliar phonetic input to existing representations may additionally be due to the cognitive control advantages of bilingualism (Bialystok et al., 2009). Similarly, bilingual individuals may also be more skilled at identifying subtle differences in pronunciation which can help differentiate words, especially in a foreign accent (Levy et al., 2019). Lastly, Bent and Atagi (2017) found that increased levels of phonological awareness is implicated in the mapping of unfamiliar pronunciations to existing lexical representations, which leads to better processing of foreign-accented speech; while their research examined bilingual children, it has been found that bilingual adults have better phonological working memory, which may confer a similar advantage (Adesope et al., 2010). Bent and Atagi (2017) identify task switching, selective attention, and inhibition as additional important factors in perceiving and adapting to accents—all of which are common traits in bilingual individuals (Bialystok et al., 2009).

Cognitive Correlates of Bilingualism

Bilingualism has long been connected to higher levels of cognitive processing, including those mentioned above. This section provides a brief overview of some of the cognitive correlates implicated in language processing specifically. In a systematic review and meta-analysis by Adesope et al., (2010), it is highlighted that bilinguals have greater metalinguistic and metacognitive awareness, better ability to selectively attend to certain information and inhibit other information, and higher cognitive flexibility, to name a few advantages (see Adesope et al., 2010, for more). Enhanced attentional control stems from bilingual speakers' ability to concurrently hold multiple linguistic representations in mind and selectively attend to only one, and this advantage has been shown for both children and adults (Adesope et al., 2010). Wang and Saffran (2014) provide an additional summary of the benefits associated with being bilingual, citing research that shows improved skills in bilinguals such as mapping between words' form and meaning, implicit learning, phonological working memory, and inhibitory control. It is suggested that these advantages may improve the ability to process novel speech (Wang & Saffran, 2014). Finally, in a review article, Bialystok et al. (2009) echoes the findings that bilinguals have greater executive functioning skills, including cognitive flexibility and control. Taken together, it appears that knowing more than one language confers an advantage in aspects of both cognition and language. The ability to map phonetic input variability to existing representations is likely a crucial aspect in understanding the variation in accented speech, making bilingualism a factor of high interest in the present study.

Statistical Learning

Statistical Learning and Language

The final area this research explores is pattern recognition ability, which is referred to as statistical learning (SL). Statistical learning is defined as "the discovery of structure by way of [analyzing] statistical properties of the input" (Misyak & Christiansen, 2012, p. 302); it is the process of tracking regularly occurring properties in continuous sensory input to determine its underlying structure (Benitez et al., 2020). SL helps individuals to categorize, discriminate, and segment information in multiple modalities to create probability-based representations, enabling rapid predictions of incoming information (Frost et al., 2015). This ability is demonstrated in infants as young as eight months old (Saffran et al., 1996), and is used throughout the lifetime across sensory, behavioral, and cognitive domains (for a review of the current field of statistical learning see Frost et al., 2019). Namely, statistical learning is involved in visual and auditory domains, event processing, motor planning, semantic memory, social cognition, music appreciation, face recognition, and, most importantly for this study, in various aspects of language (Frost et al., 2019; Siegelman, 2020). Picking up regularities within the environment is done implicitly and without conscious awareness; the brain develops representations of statistical relationships between stimuli before an individual is aware that this has happened, and before a behavioural response is possible (Koelsch et al., 2016). A significant body of literature has demonstrated that individuals can extract statistical regularities from environmental input in various modalities, tasks, and across age groups (Abla et al., 2008; Bulf et al., 2011; Christiansen, 2005; Kirkham et al., 2002; Saffran et al., 1996).

Statistical learning is especially implicated within language acquisition and use. Every language has its own inherent set of predictable rules and patterns, which govern its variety of specific sounds, order of words, grammar, and pronunciation of sounds. These linguistic regularities have both transitional and distributional probabilities; transitional probability (TP) refers to the likelihood that items will co-occur sequentially, while distributional probability refers to the frequency with which an item may appear (Thiessen et al., 2013; for more information on which statistics are learned, see Thiessen et al., 2013). Language users recognize and learn both types of probabilities through their SL mechanism, though SL research has focussed more on TP. For example, in orthography, there are certain co-occurrences of letters which are more probable than others; speakers of English know the letters S and E are more statistically probable to occur together than the letters J and A (Chetail, 2017). Letter-to-sound correspondences are governed by probabilistic patterns as well (Treiman & Kessler, 2006). In morphology, there are non-adjacent dependencies such as the structure 'is __ing', where occurrence of the auxiliary 'is' has statistical likelihood of co-occurrence with the present progressive form '-ing,' even though there is a variable verb in between (Pacton et al., 2005). Adjacent dependencies occur frequently as well (i.e., where a structural property immediately predicts another), such as determiners predicting a subsequent noun (Misyak & Christiansen, 2012). Syntax, prosody, semantics, and even gesturing contain similar probabilistic patterns (see Siegelman, 2020). Research has found that in both learners and proficient speakers of a language, SL is associated with learning morphological patterns and orthographic regularities (Pacton et al., 2001), learning the structure of syntactic phrases (Saffran & Wilson, 2003), expanding lexical knowledge and vocabulary size (Shafto et al., 2012), and forming syntactic categories (Gerken et al., 2005), among other aspects; (for an overview, see Misyak & Christiansen, 2012, or Siegelman, 2020). Throughout syntax, morphology, orthography, and other areas of language, the rules and patterns are not deterministic or random. The statistics of these linguistic regularities are processed and learned by individuals, enabling proficient language use (Siegelman, 2020).

Indeed, individuals with greater SL ability have been found to perform better on various measures of language outcomes. For example, in a study by Misyak and Christiansen (2012), it was found that SL ability was strongly related to both verbal working memory performance and language comprehension, with SL being the only cognitive mechanism to predict comprehension accuracy. SL can also predict oral language and vocabulary knowledge (Spencer et al., 2015). Further, much research has positively correlated SL ability with syntactic processing, lexical knowledge, vocabulary acquisition, reading ability, and semantic knowledge, among other aspects of linguistic ability (Ahufinger et al., 2021; Arciuli & Simpson, 2012; Conway et al., 2010; Lany et al., 2018; Misyak et al., 2010; Shafto et al., 2012; Siegelman, 2020; Singh et al., 2012). Additional evidence for the connection between SL and language comes from individuals with language impairments, who have been found to have deficits in SL ability (Ahufinger et al., 2021). In a study of individuals with dyslexia, Ahissar and colleagues (2006) found that general language aptitude was related to sensitivity to the statistics of auditory stimuli. For reviews and meta-analyses of the connection between SL and language impairment, see Bogaerts et al., 2020; Lammertink et al., 2017; and Saffran, 2018. Given the plethora of connections between SL ability and language outcomes, it is evident that SL is a relevant factor in the current study and worth examining further, as it is highly implicated in the processing of language.

Statistical Learning in the Auditory Domain

Of particular relevance is the connection between statistical learning and the auditory domain. Just as morphology, orthography, syntax, and semantics follow predictable statistical distributions, so do auditory elements in speech such as phonetic, phonological, and acoustic features. Relating to the phonetic features of accents in particular, vowel quality (i.e., vowel duration and formants), consonant quality (i.e., voice onset time, segment duration, and spectral energy), and phonological processes (i.e., elision, substitution etc.) were all identified previously as salient features which characterize foreign and regional accents (Moyer, 2013). In the auditory domain of a speech signal, there are distributional statistics relating each of these dimensions to a specific phoneme in a given language or dialect (Kleinschmidt & Jaeger, 2015; Toscano & McMurray, 2010; see Stilp & Lewicki, 2013). For example, speakers of English know that a voice onset time (VOT) of 10-20 ms or less corresponds to a voiced phoneme, while VOT of 40-50 ms or more corresponds to a voiceless phoneme (Toscano & McMurray, 2010; see also Allen & Miller, 1999). This statistical pattern helps give rise to categorical perception, such as identifying the difference between /d/ and /t/. Similarly, vowels have systematic differences in duration, such that the vowel in "send" is usually shorter than the vowel in "sand," for example (Lehet & Holt, 2020). Being sensitive to these statistics helps language users form stable long-term phonetic categorical representations, but also exhibit short-term flexibility when encountering inevitable variance in the speech signal, such as a speaker with an accent (Roark & Holt, 2022).

Given the statistics in the auditory features of speech, it is no surprise that statistical learning has been correlated with performance on auditory measures as well. Of note, SL ability in children and adults has been tied to segmenting speech (Saffran et al., 1996), learning phonotactic constraints (Warker & Dell, 2006), forming phonetic categories (Maye et al., 2008), constructing phonemic representations (Vandermosten et al., 2019), and phonological processing (Spencer et al. 2015). Again, further evidence comes from studies of individuals with dyslexia, who have been found to have reduced ability to learn the statistical cues of speech sounds and identify phonetic contrasts (Vandermosten et al., 2019). Additionally, individuals use SL to process non-linguistic auditory stimuli. This includes pure tones (Creel, Newport, & Aslin, 2004), timbres of the same pitch (Koelsch et al., 2016), and even computer sound effects (Gebhart, Newport, & Aslin, 2009). In a study by Saffran et al. (1999), it was found that adults and eight-month-olds alike could extract regularities and learn the statistical patterns of musical tones equally as successfully as syllables, and these findings were corroborated by Abla et al., (2008). This suggests that SL ability applies to the segmentation of non-linguistic stimuli in addition to linguistic; auditory stimuli are grouped and categorized similarly in the brain, whether words or tones (Saffran, 1999; Vasuki et al. 2016). Finally, research has found a relationship between musical ability and enhanced auditory SL performance, as SL is highly implicated with learning the regularities in music (Vasuki et al., 2016). Notably, musical expertise and experience seems to augment auditory SL, which may have an implication for the role of experience and accent familiarity in the SL of accented speech.

Connections

Accents and Statistical Learning

As mentioned, the acoustic features which characterize speech, regardless of the level of familiarity or accentedness, follow statistical distributions. While there are regular distributions and relationships between features such as VOT and duration, speech contains considerable acoustic variability as a function of dialect, accent, or speaker. A speaker is perceived to have an accent when the acoustic and phonetic features of their speech differ from a listener's representations, and the task for the listener is to map the variability in the speech signal to existing categorical representations (Idemaru & Holt, 2011; Lehet & Holt, 2020). As such, categorical perception seems to be an important mechanism in understanding accents; categorical perception is the parsing of speech sounds into discrete perceptual categories, such as /b/ or /d/ (Hixon et al., 2020). A form of statistical learning has been proposed as a potential model for how speakers adjust to the variability in accented speech and adapt their perceptual categorization of speech sounds. In the case of an unexpected or unclear acoustic feature, like ambiguous VOT for a stop consonant, statistical learning may help the listener shift to rely less on VOT for categorization and rely more on a different feature, like fundamental frequency (Idemaru & Holt, 2020). This allows for short-term flexibility in processing a pattern of acoustic features that differs from expectations, such as in an accent, within the long-term stability of existing categorical representations (Roark & Holt, 2022; for a detailed explanation, see Lehet & Holt, 2020). The statistics of divergent auditory features are learned rapidly, and this helps to accommodate acoustic variability, as speakers of different languages will exhibit different relationships or relative weights between features (Idemaru & Holt, 2011). This process of perceptual statistical learning appears to be a possible link between picking up on the subphonemic statistical regularities that characterize speech and adapting to variation in these regularities as found in a foreign accent.

At a broader level, statistical learning is one potential underlying mechanism supporting accent familiarity and perceptions of intelligibility of accented speech. There are probabilities in the distribution of accents, dialects, and representative speakers that one is familiar with, including the likelihood of encountering a familiar speaker or accent again (Kleinschmidt & Jaeger, 2015). As outlined previously, experience with multiple speakers from the same accent (Baese-Berk et al., 2013; Bradlow & Bent, 2008) can lead to generalization across groups of speakers, which could help guide speech perception. Listeners learn the statistics of how patterns of features characterizing a phoneme can vary across speakers who share the same accent (Kleinschmidt & Jaeger, 2015). As such, the structure of previous experiences informs perceptual or statistical learning about current experiences and interactions. Accent familiarity may then be a function of generalizing experience with other speakers' patterns of acoustic features when encountering a new speaker of that language background. Recall that bilingual individuals were found to have more robust categorical perception (Archila-Suerte et al., 2012), and that their increased experience with speech variability allows them to process accents more effectively (Saito & Shintani, 2015; Shintani et al., 2019). This suggests a possibility that bilinguals may have better perceptual-level statistical learning ability, and it provides an example of how accent, accent familiarity, statistical learning, and language background are all interrelated.

Bilingualism and Statistical Learning

As for more general connections between bilingualism and statistical learning, a relationship has been reported between SL ability and L2 proficiency, yet the directionality of this relationship is still unknown. Frost et al. (2013) found a significant correlation between scores on a visual statistical learning (VSL) task and measures of reading ability in the acquisition of a foreign language. This was in the absence of any significant correlation with general cognitive abilities, showing cognition was not a factor. It seems that greater SL ability may predict better adaptability or learning capability regarding the structure of an additional language (Frost et al., 2013). Similar research by Wu et al. (2012) found a congruent result with learners of Mandarin as an L2, particularly with reading and orthography. Research by Wang and Saffran (2014) suggests that being bilingual may confer an advantage in SL ability. In comparing the performance of English monolinguals, Mandarin monolinguals, and Mandarin-English bilinguals on a SL task using tonal syllables, the Mandarin bilingual speakers exhibited better SL performance than both groups of monolinguals. Interestingly, a subsequent experiment found that bilinguals from non-tonal languages (e.g. English and Spanish, German, Korean, or Polish) were still just as successful at learning the syllables as the Mandarin monolinguals. This suggests that despite a lack of familiarity with tones, bilingualism plays a significant role in SL ability (Wang & Saffran, 2014). The authors posit that because bilingualism has been associated with greater phonological working memory capabilities, this could lead to correlation with better statistical learning of word forms. It is thought that the ability to hold speech in phonological working memory for longer enables bilingual speakers to analyze the transitional properties more effectively (Wang & Saffran, 2014). In sum, there seems to be a relationship between SL ability and bilingualism, in addition to the connection between bilingualism and processing accented speech. The extent of how these three factors overlap and interact in their influence is still largely unexplored.

Summary

Thus far, a broad overview has been provided of the theories and literature concerning accents, the role of accent familiarity and exposure, bilingualism, language- and auditory-based statistical learning, and the connections therein. It seems that greater familiarity with an accent makes processing accented speech easier, that bilinguals may have an advantage in processing accents, and that a finely tuned statistical learning mechanism may explain how accents are processed. However, many of these connections do not have robust empirical support as of yet, and have not been studied together within individuals. The directions of the relationships between all of these factors are still unknown, so this study is positioned to investigate the influence and interaction of all three interconnected factors; previous research has largely only focussed on pairs of factors. Additionally, statistical learning is a relatively recent concept, and the intersection of SL with broader concepts such as accented speech has been relatively unexplored.

Research Questions

Given the paucity of empirical connections in the literature between statistical learning, bilingualism, and accent familiarity amid a linguistically diverse population, there are many avenues to explore within and between each of these topics. Several related research questions have been selected to provide direction to this research. The overall goal of the study is to determine the magnitude of the relationships between statistical learning, bilingualism, and accent familiarity. Particularly, this study aims to examine the following questions:

1) Does familiarity with different accents impact the ability to accurately learn patterns across accented phonemes?
2) Does language background, particularly being bilingual, impact the ability to accurately learn patterns across accented phonemes?
3) Do accent familiarity and language background interact in their influence? Is one a better predictor than the other of pattern learning across accented speech?

METHODS

Tasks

Statistical Learning

The experiment consisted of a series of tasks that assessed participants' ability to learn and recognize both visual and auditory patterns. The visual statistical learning (vSL) task was conducted as a baseline task measuring non-linguistic SL ability, while the auditory statistical learning (aSL) task assessed participants' generalization of learning across accented phonemes.

Statistical Learning Paradigm. Statistical learning refers to how the presence of discrete units is learned based on the probability of stimuli within each unit (for an overview of learning and processing in SL, see Siegelman, Bogaerts, Kronenfeld, et al., 2018). SL tasks are traditionally conducted with triplets of stimuli, and the transitional probabilities (TPs) within each triplet indicate their presence and boundaries. In their seminal study, Saffran et al., (1996) used the phrase ‘pretty baby’ to illustrate this point. The syllable ‘pret’ is much more likely to be followed by ‘ty,’ which is within an established word, than the syllable ‘ty’ is to be followed by ‘ba,’ which is across a word boundary. As another example, if A is exclusively followed by B in the triplet ABC, then the TP between A and B is 1.0. However, if triplet ABC can be followed by either DEF, GHI, or JKL, then the TP between C and D is only 0.33. Hence, the TPs within triplets are greater than the TPs between triplets, and this indicates the presence and boundaries of each triplet. That is, a stimulus can be reliably predicted based on its preceding stimulus (see Saffran et al., 1996).

Statistical learning tasks contain two phases: a familiarization phase and a testing phase. In the familiarization phase, the triplets are concatenated together in a continuous stream with a congruent amount of time both between each stimulus and between each triplet; the triplets are thus embedded within the stream. The only statistical cue to the presence of triplets and to their boundaries is in the transitional probabilities (TPs) between stimuli. In the familiarization stream for both aSL and vSL tasks, no triplets are repeated twice in a row, and no items are repeated within a triplet. The stream is presented for 10 minutes, and there is no cover task. In the testing phase, participants are presented with a series of two-alternative forced-choice (2AFC) questions to assess their offline learning of the TPs between stimuli. Each trial in the testing phase consists of a foil or novel triplet alongside a triplet that appeared in the familiarization phase. The foil triplets are composed of the same stimuli as the target triplets but in a sequence that does not appear in the familiarization phase (e.g., ABC versus CDA). After presentation of the foil and target triplet, participants are asked to identify which triplet they heard or saw previously. Performance is measured by accuracy on discriminating target triplets from foils through correctly identifying previously encountered triplets, and through reaction time in identifying the target triplet. The total number of correct responses is divided by the number of 2AFC questions to yield a percentage of correct responses. Accuracy can be further broken down in the aSL task by the type of triplet, as described below.

Auditory Statistical Learning Stimuli. Previous research with aSL has largely relied on using synthesized speech with no additional acoustic or phonetic information, as in the classic study by Saffran et al. (1996). To date, however, there are few or no studies which utilize natural speech with multiple forms of accented English. This research aimed to fill that gap and thereby increase generalizability of the findings by using natural accented speech from speakers with the following language backgrounds: Tagalog, Punjabi, Mandarin, and Albertan English. According to Statistics Canada (2017), Tagalog, Punjabi, and Mandarin are the three most-spoken non-official languages in Edmonton in both the 'Mother Tongue' and 'Language Spoken Most Often at Home' categories. French is also a frequent mother tongue and language spoken at home; however, this research focuses on languages which are not official languages of Canada. Cantonese also ranked similarly to Mandarin, but Mandarin was chosen as it is spoken by the majority of individuals in China.

Representative speakers from the four aforementioned language backgrounds were chosen from the Speech Accent Archive (SAA), which is a database of carefully recorded and transcribed speech submitted by speakers of English around the world (Weinberger, 2015). The data of the SAA is available for use under a Creative Commons license (CC BY-NC-SA 2.0); for literature using SAA data, see Gao & Weinberger (2018). Each speaker in the archive reads a standardized elicitation paragraph containing most of the consonants, vowels, and consonant clusters in American English, and their speech is recorded on a high-quality device. The speakers chosen for this study have all acquired English after age 12 and resided in an English-speaking country for at least four years; because they learned English past the critical period, they all exhibit an accented variety of English (Dollmann et al., 2020; Moyer, 2013). For detailed demographic information of each speaker, refer to Appendix A.

From the elicitation paragraph used in the SAA, 16 phonemes were chosen alongside their subsequent vowels to create consonant-vowel (CV) syllables (e.g.,"do-re-mi"). Each CV syllable was extracted from the recording and manipulated using Praat (Boersma & Weenink, 2022) to have a duration of 50-60 ms. Only the vowel was manipulated in order to preserve the original VOT, which is a core characteristic of accented phonemes, as discussed previously. The native phonetic inventories of Tagalog, Mandarin, and Punjabi were systematically compared to identify phonemes which were most likely to differ in pronunciation from Albertan English phonemes (Moran & McCloy, 2019); Albertan English will hereafter be referred to as unaccented, since this research is being conducted in Edmonton. Each CV syllable was carefully chosen to minimize coarticulation effects such as nasalization, velarization, or rhoticization, and tense vowels were chosen over lax or schwa vowels. For a complete list of all CV syllables used, a rationale for each, and details on the structure of each triplet, see Appendix B.

Each of the 16 phonemes was extracted from all four varieties of English to create 48 CV syllables, which were then concatenated into 8 triplets using Praat. The triplets consisted of both accented and unaccented CV syllables, with 4 triplets being entirely unaccented and 4 being a combination of accented and unaccented. Some syllables were used in more than one triplet, as there are only 16 CV combinations rather than 24. This created a 'hard' triplet and 'easy' triplet distinction, where the four hard triplets had four phonemes in different orders (e.g., "do-re-fa," "re-do-mi," "mi-fa-do," "fa-mi-re"). The easy triplets had three phonemes each which do not appear in other triplets and are always in a fixed order (e.g. "so-la-ti"). The 8 triplets were repeated 4 times for a total of 32 presentations per cycle. The cycle was repeated 6 times, for 192 total presentations of triplets. This design, as well as the congruent design for the visual SL task, has been used extensively through the SL literature, with established reliability and validity (see review by Frost et al., 2019). The triplets presented in the familiarization phase all had the same accent within each triplet, with a meticulous framework of which triplets were presented in which accent condition. Half of the triplets were only presented in Albertan English, while the other half had an equal proportion of Mandarin, Punjabi, and Tagalog presentation.

In terms of the testing phase, participants were asked to identify the target triplet from foils in a series of 42 questions. There were 18 questions where the triplet was presented with one foil as a two-alternative forced choice (2AFC), 12 questions where the triplet was presented alongside three foils as a four-alternative forced choice (4AFC), and 12 questions where a portion of the triplet was presented and there were three options of syllables to complete it, referred to as 3-blank. All of these questions were presented in a randomized order and there was no time limit. In terms of the makeup of each triplet with the 48 accented phonemes, there were four different question types: Normal, Easy AB, Hetero Acc, and Both Correct. There were 18 'Normal' questions, where each triplet presented had the same accent condition within the triplet, and the target triplet was paired with any of the foils. There were 6 'Easy AB' questions, which had an 'easy' triplet in the Albertan accent as the target, and the foils were all in non-Albertan accents. 'Hetero Acc' meant that the target and foil triplets each had different accents within the triplet, and there were 7 of this type. Lastly, 'Both' meant that there were two correct answers, with a target triplet presented in two accent conditions, and there were 11 of this type; this was included to see if participants were more likely to choose one accent condition over the other when presented with two correct options. The 3-Blank questions were half 'Normal' and 'Both Correct,' while the 2AFC and 4AFC were a proportionate mixture of each question type. Finally, there were some target triplets in the testing phase that appeared in a different accent condition than the familiarization phase. This was included to assess generalization of the accented phonemes.

Visual Statistical Learning Stimuli. The stimuli used in the vSL task is described in Siegelman, Bogaerts, Kronenfeld, et al. (2018), which was patterned after the original design by Kirkham (2002). Complex black shapes on a white background are used instead of spoken syllables, presented in a continuous stream of isochronous triplets in the center of the computer screen. These 16 unique shapes (see Appendix C) are organized in the same manner as the auditory stimuli to create 8 distinct triplets, with four 'hard' triplets and four 'easy.' Each shape is presented for 800ms with 200ms between shapes and between triplets.

Questionnaires

In between the SL tasks, participants completed a series of questionnaires to prevent fatigue. The questionnaires included a brief demographic survey to gather their age, sex, gender, handedness, any cognitive or neurological conditions, education level, and year of study. A language background questionnaire adapted from the Language Social Background Questionnaire (LSBQ) used by Anderson et al. (2018) was also administered. This questionnaire assessed any other languages participants knew or spoke, when they learned these languages, their proficiency in speaking, listening, reading, and writing for each language, and in which domains they used each language. Scoring the LSBQ was done using a pre-made spreadsheet developed by Anderson et al. (2018), and it yielded a composite factor score which indicates the degree of bilingualism. This was a continuous variable, but it was also used categorically with cut-off scores for monolinguals, receptive bilinguals, and bilinguals. Other factor scores yielded from the questionnaire included continuous values for non-English use at home, non-English use socially, and English use proficiency.

Lastly, participants completed a self-reported measure of familiarity with the accents used in the study. An Accent Familiarity Questionnaire (AccFamQ) was developed and study data were collected using REDCap, an electronic data capture tool hosted through the University of Alberta (Harris et al., 2009). Participants were first asked to give a rating from 1 to 10 on their general comfort level in interacting with speakers with an accent. From there, they checked which accents they were familiar with and had the option to type up to five additional accents. The given list comprised the top 10 non-official languages spoken in Edmonton and also included French (Statistics Canada, 2017). If participants checked a box, a sliding scale appeared asking them to rate how familiar they were with that accent of English on a scale of 1-100. Mandarin, Tagalog, and Punjabi were mandatory sliding bars, because those were the accents included in the task, and any additional sliding bars were based on participants' selections. Next, participants were asked to rate how frequently they interacted with non-native speakers of English on a weekly basis, ranging from 'never' to 'daily' on a 10-point scale. Follow-up questions were asked on how frequently those interactions were with Mandarin-, Tagalog-, or Punjabi-accented English on a 10-point scale. Lastly, participants were presented with boxes to check that represented various domains where they may regularly (i.e., at least once a week) encounter Mandarin, Tagalog, or Punjabi accents. The options for domains included 'Home,' 'School,' 'Work,' 'With friends,' 'With family,' 'In the community,' 'In leisure or hobby activities,' 'In social media,' 'In TV/movies,' 'In religious activities,' or 'NA.' The questionnaire took between 5-10 minutes to complete in total and was filled out online through the REDCap website. To date, we have yet to learn of a standard accent familiarity questionnaire with reliable and valid measures in the literature, hence creating the AccFamQ for this study. Similar questionnaires assessing experience with accented speech have been used by Levy et al. (2019), Porretta et al. (2016), and Shintani et al. (2019), which served as a reference.

Participants

There were 43 participants who completed the experiment (34 female, 6 male, 3 other). All participants were adults ranging from age 18 to 54 (*M* = 25.4, *SD* = 7.25, *Md* = 23). Participants were all proficient in reading English, as measured through self-report and years spent in Canada or another English-teaching country. Recruitment was conducted at the University of Alberta, with most participants being undergraduate students. Participants were compensated with a $10 gift for their time. This study was approved through the University of Alberta Research Ethics Board 2, Pro00117511.

Procedure

The experiment was conducted in person at the University of Alberta Clinical Sciences Building. After participants were walked up to the lab and informed consent was obtained for the study, a standardized hearing screening was conducted at 500, 1000, 2000, and 4000 Hz at 25 dB in both ears. If participants failed the hearing screening, they were referred to the audiologist on the study team for a follow-up assessment if they desired; data was not discarded based on screening results. Because the study used primarily auditory stimuli and required sensitivity to subtle changes in these stimuli, the hearing screen ensured participants could correctly and fully participate in the experiment. Participants were seated at a computer for the duration of the experiment, wearing headphones for the auditory stimuli, and they were randomly assigned to one of two order groups to address order effects. Half of the participants completed the auditory SL task first, followed by the language background questionnaire and accent familiarity questionnaire, which included demographic questions, and completed the visual SL task last. The other half completed the visual SL task first, followed by the questionnaires, and the auditory SL task last. The entire experiment took approximately one hour.

Challenges

Creation and implementation of a novel auditory statistical learning task and novel questionnaire was not without its challenges. The complications and solutions described below lend insight into the process of developing and running this experiment.

Stimuli Creation and Task Design

The most significant obstacles were faced while putting together the stimuli and designing the triplets for the aSL task. First, there were challenges with finding appropriate recordings from the Speech Accent Archive (SAA) for each representative accent, as I was looking for each speaker to be male, of a similar age, and have a similar age of English acquisition. The Mandarin speaker had a higher habitual pitch than the other male speakers, so his sample stood out compared to the others; a different Mandarin speaker's sample was chosen partway after comparing syllables to the other speakers. The Albertan English speaker exhibited significant glottal fry throughout his sample, making some syllables unintelligible in isolation. Once chosen, the recordings were uploaded to Praat, and each of the 64 consonant-vowel combinations was segmented, extracted, and exported to an individual file. Some CV combinations were part of a CVC(C) syllable (e.g., "ki" from the word "kids") so had to be carefully segmented to minimize coarticulation effects. Because the audio sample was from a standardized reading passage rather than individually produced sounds, the duration of each sound was extremely short. Consequently, one phoneme needed to be substituted partway through: "la" from the word "Stella" was found to be far too short and low in pitch due to its unstressed position in the word, so "pee" from the word "peas" was chosen instead. The duration tier was manipulated using Praat to make each syllable slightly longer, resulting in an average length of 276.4 ms per CV syllable. Additionally, the frequency (Hz) was manipulated for some syllables, because varying intonation patterns made some syllables stand out substantially, both within and across triplets. The duration and frequency manipulations were documented for each syllable and each accent condition. Additionally, all sounds were manually normalized to 70 dB intensity using Praat. Overall, this process was extremely tedious, time-consuming, and cognitively demanding.

After normalizing the syllables for pitch, duration, and intensity, they were each saved as the ‘edited’ version. I taught myself how to use a script in Praat to add the appropriate amount of milliseconds of silence to each of the syllables so they were approximately 1 second in duration. Then I used another script to concatenate the three syllables into their corresponding triplet as a sample of how the triplet would sound. Eight triplets per speaker were created this way. I also worked with a computer science student who formally added silence before and after each syllable, so the total duration was exactly 1 second. Each syllable ended up being between 999.7 to 999.9 ms due to programming bias from his Python script.

Once the stimuli were finalized, considerable time was spent designing the testing phase for the aSL task as well. This involved deciding how many of each question difficulty type would be asked (2AFC, 4AFC, and 3-blank) and what the different accent conditions or question types would be. I went through systematically to ensure that each triplet in each accent condition was represented a proportionate amount of the time, there was a proportionate number of one-correct and both-correct questions, there was a proportionate number of ‘easy’ and ‘hard’ triplets, and foil triplets had each syllable and accent condition equally represented. Next, I delineated which syllables would go in what order for the foils and target triplets. Due to glottal fry, salient intonation patterns, or audio quality, three triplets were not included as any of the target triplets. These were ‘daycheethi’ in English, ‘toibathee’ in English, and ‘kigoshee’ in Tagalog. I also needed to develop a list outlining which triplets could not follow each other, since the testing phase script and SL paradigm stipulated that no two identical triplets could follow one another. This meant that the same triplet could not follow even if it was in a different accent (e.g., 08a, 09a, 10a and 08c, 09c, 10c, where ‘a’ was the Albertan English condition and ‘c’ was Mandarin English). After this extensive editing and planning, the sounds were ready to be programmed into the aSL task in PsychoPy.

The aSL task was patterned after another aSL task designed for fMRI with non-linguistic stimuli. Many parameters were the same, but several adaptations had to be made, including omitting all the ‘pause’ triplets—a vestige from the fMRI task to wait for the scanner. There were a few iterations of the script necessary to thoroughly adapt both the auditory SL and visual SL to be a behavioral task instead of an imaging task. The sound files also needed to be resampled and changed to exactly 1000ms rather than 999ms to mitigate the script from crashing, which was happening continuously for a few weeks. Lastly, adjustments were necessary to allow facets like text in the participant IDs, a fixation cross during the familiarization phase, numbers to indicate progress monitoring in the testing phase, and to have the resulting output files show all the desired data, including keystrokes, correct/incorrect response, and running time stamps. All programming was generously completed by a postdoctoral fellow from the University of Toronto, whom I worked closely with through the many versions of the script and its output.

Questionnaires

Creating the Accent Familiarity Questionnaire was also a substantial task, especially given that there was no existing literature on which to model this questionnaire; only a portion was informed by Porretta’s work on quantifying accent familiarity (2016). First, I explored various survey sites to find a platform that was accessible, intuitive, and could display sliding scales. REDCap was decided on due to its security, favorability with the Research Ethics Board, and ability to have sliding rating scales. I then had to learn how to navigate REDCap and teach myself to use features such as data piping, branching logic, field embedding, and field matrices. In deciding how many accents to ask participants about, brevity and conciseness had to be balanced with not solely asking about Mandarin, Tagalog, and Punjabi. Hence, Spanish and Arabic accents were included in the questions about domains, as Spanish and Arabic are the next most common non-official languages spoken in Edmonton after Mandarin, Tagalog, and Punjabi. The inclusion of additional accents was essential given the randomized order of tasks, with half of the participants completing the accent familiarity questionnaire before doing the aSL task, to not prime the participants. All in all, these various challenges in creating novel tasks contributed to my understanding and appreciation of the experimental process and its nuances.

Summary of Methods

To recap, 43 participants recruited through the University of Alberta completed two statistical learning tasks, one with novel accented stimuli based on natural speech and the other with visual shapes. Participants were exposed to the stimuli for 10 minutes during the familiarization phase. Then, they were asked 42 questions requiring them to identify a triplet previously presented in the familiarization phase against foil triplets. For the auditory statistical learning task, the questions varied based on the composition of accents in the targets and foils. Participants also completed a standardized hearing screening, demographic questionnaire, accent familiarity questionnaire, and an adaptation of the Language Social Background Questionnaire (LSBQ), modeled after Anderson (2018).

RESULTS

All analysis was run through the statistical software R (v 4.0.3) using the lme4 package (Bates et al., 2015). Linear regression using model fitting was utilized to understand how the various predictor variables are connected to one another in prediction of the criterion variables, which were accuracy and reaction time on the auditory statistical learning task. Model fitting procedures were used throughout to examine which variables best determined the proportion of shared variance for the other factors.

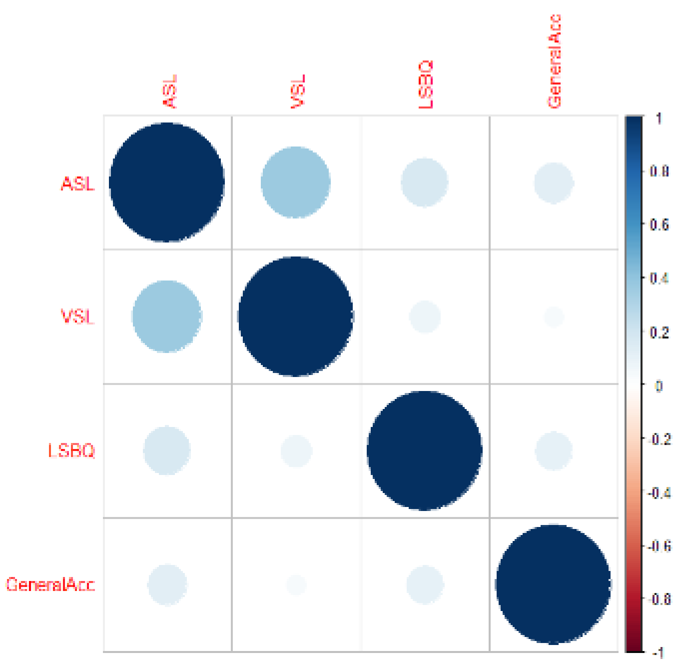
Participants' performance on the auditory statistical learning (aSL) task was first measured according to accuracy, or score, to address the research questions stated above. Accuracy was determined by the number of questions answered correctly in the testing phase, with the total number of questions being 42 (*M* = 24.15, *SD* = 4.04, *Min* = 16, *Max* = 36). The aSL task was subsequently measured by reaction time, and performance was further analyzed according to question difficulty and question type. Visual statistical learning was measured through overall score or accuracy as well (*M* = 27.62, *SD* = 6.49, *Min* = 17, *Max* = 39). Language background was quantified as the total composite factor score on the LSBQ questionnaire, which was a rational number between -5 and 13 (*M* = 1.69, *SD* = 3.86). Those who received a composite factor score above 1.23 were classified as bilingual, those with scores below -3.13 were classified as monolingual, and those with scores in between were classified as receptive bilinguals (Anderson, 2018). Language background was further broken down into the categorical variables of English L1 status, as to whether English was the participant's first language (*n* = 29) or a second or additional language (*n* = 14). The mean LSBQ score for English L1 participants was 0.046, while the mean for English L2 participants was 5.037. LSBQ score was also used to classify by language status, as to whether participants were categorized as monolingual (*n* = 5), receptive bilingual (*n* = 15), or bilingual (*n* = 23). Since the number of monolinguals was so low, language status was reclassified by English L1 and L2. There were 14 English L1 receptive bilinguals, 10 English L1 bilinguals, and 13 English L2 bilinguals; only one participant was an English L2 receptive bilingual. Language background also included the continuous variable of participants' non-English language use at home (*M* = 1.75, *SD* = 6.49) and socially (*M* = 3.24, *SD* = 7.83); a lower score indicates more English use, whether as a first or second language. Lastly, accent familiarity was quantified as a self-reported measure of general comfort with accented speech, on a scale of 1 to 10 (*M* = 7.63, *SD* = 1.62). Accent familiarity was further analyzed according to the frequency of participants' interactions with non-native speakers of English, measured on a Likert scale out of 10 with “Never” corresponding to 0 and “Daily” corresponding to 10 (*M* = 7.77, *SD* = 2.62). Other variables from the questionnaire included participants’ familiarity out of 100 with the accent varieties of Mandarin (*M* = 57.77, *SD* = 28.77), Tagalog (*M* = 54.07, *SD* = 30.18), and Punjabi (*M* = 46.84, *SD* = 28.55), and the number of accents they reported as being familiar with out of 16 (*M* = 6.63, *SD* = 2.79).

Auditory Statistical Learning

Regarding aSL score, correlation analysis was run between the variables of LSBQ score, general accent comfort, and vSL score. The only variable that was correlated with aSL score was performance on the visual statistical learning (vSL) task, $r(41) = .36$, $p = 0.016$, as shown in Figure 1; see also Table S1 in the Supplementary Materials for all the correlation values.

Figure 1

Correlation Plot for aSL Score, vSL Score, LSBQ, and General Accent Familiarity Rating



Note. Color and size of the dots relate to the strength of the correlation. Blue dots indicate a positive correlation, while red dots indicate a negative correlation; the larger the dot, the higher the correlation.

Linear regression was conducted as well between LSBQ score, general accent comfort, aSL, and vSL. Again, the only variable significantly predicting score on the aSL task was score on the vSL task; see Table 1.

Table 1

Regression Analysis for vSL Score, General Accent Familiarity, and LSBQ Predicting aSL Score

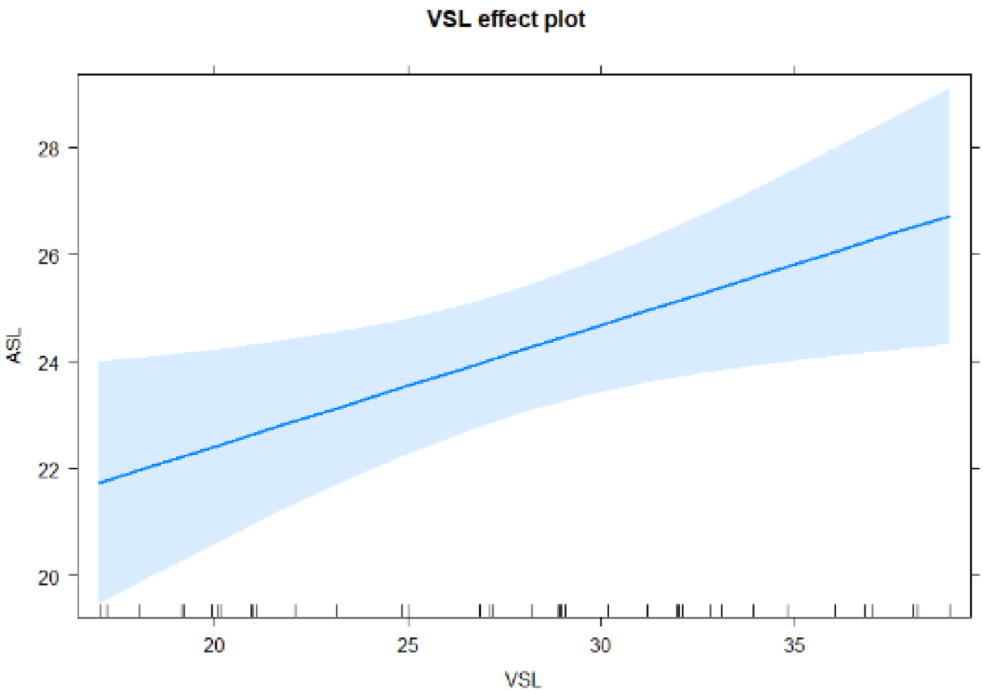
	Estimate	Standard Error	*t*	*p*	
(Intercept)	15.96032	3.75577	4.520	0.000129	***
VSL	0.21956	0.09153	2.399	0.021326	*
GeneralAcc	0.24845	0.36833	0.675	0.503943	
LSBQ	0.13168	0.15468	0.851	0.399782	

Note. $n = 43$, $R^2 = .097$, $p = .073$, model residual standard error is 3.838 on 39 degrees of freedom.

Linear regression between vSL and aSL scores (b = 0.227, SE = 0.091, $t = 2.508$, $p = 0.016$) and the corresponding graph in Figure 2 show the positive relationship between participants' overall scores on both tasks; see also Table S2.

Figure 2

Regression Model for aSL Score as Explained by vSL Score



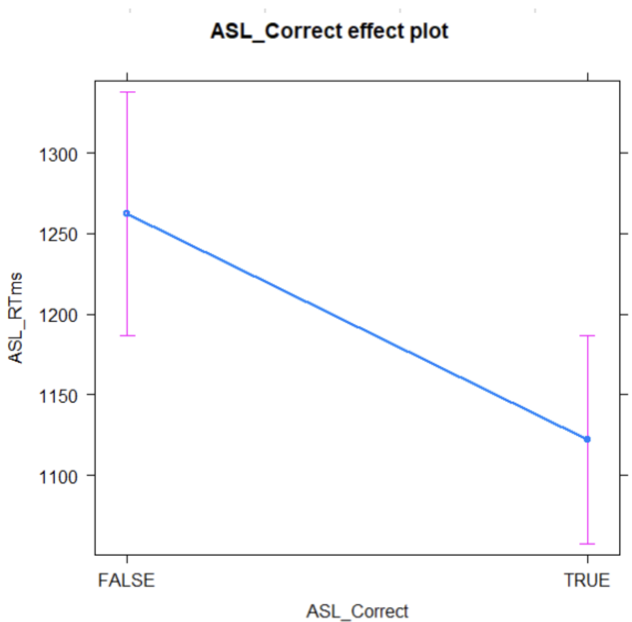
Note. Total score for both aSL and vSL tasks is out of 42. Shaded area denotes confidence intervals. $R^2 = .1118$.

No other factors were found to be predictors of overall aSL score.

Beyond overall score, post-hoc analyses examined participants' reaction time (RT) in milliseconds as a further indication of auditory statistical learning performance. Each individual RT trial across 43 participants and 42 questions was considered rather than 43 grand averages; 106 trials were removed as they were greater than 5000ms (*n* = 1700). Linear regression showed that participants were faster to respond to questions which they got correct, and slower with responses that were ultimately incorrect (b = -139.83, SE = 50.60, *t* = -2.763, *p* = .00579); see table S3 and Figure 3. Similarly, the higher participants' score, the faster their response time was on average (b = -28.321, SE = 6.119, *t* = -4.628, *p* = < .001).

Figure 3

Reaction Time on aSL Task as Explained by Correct Trials



Note. Reaction time in milliseconds is displayed on the *y*-axis. Correct trials are indicated as 'TRUE' and incorrect as 'FALSE' on the *x*-axis. Pink lines denote error bars.

Performance on the aSL task was further analyzed according to question type and difficulty level. As mentioned, question type refers to the composition of accents within the triplets and foils. For 'Normal' questions, each triplet had the same accent condition within, and any target was paired with any of the foils. 'Easy AB' questions had an 'easy' triplet in the Albertan accent as the target, and the foils all contained non-Albertan accents. 'Hetero Acc' meant that the target and foil triplets each had different accents within the triplet. Lastly, 'Both' meant the same target triplet was presented in two accent conditions. Question difficulty ranged from 2AFC, where one target and one foil was presented, to 4AFC, where one target and three foils were presented, and 3-blank, where two-thirds of a target triplet was presented and participants had three options to choose from to complete the triplet. Table 2 shows how many individual trials of each question type and difficulty were presented, after removing 106 outliers based on reaction time.

Table 2

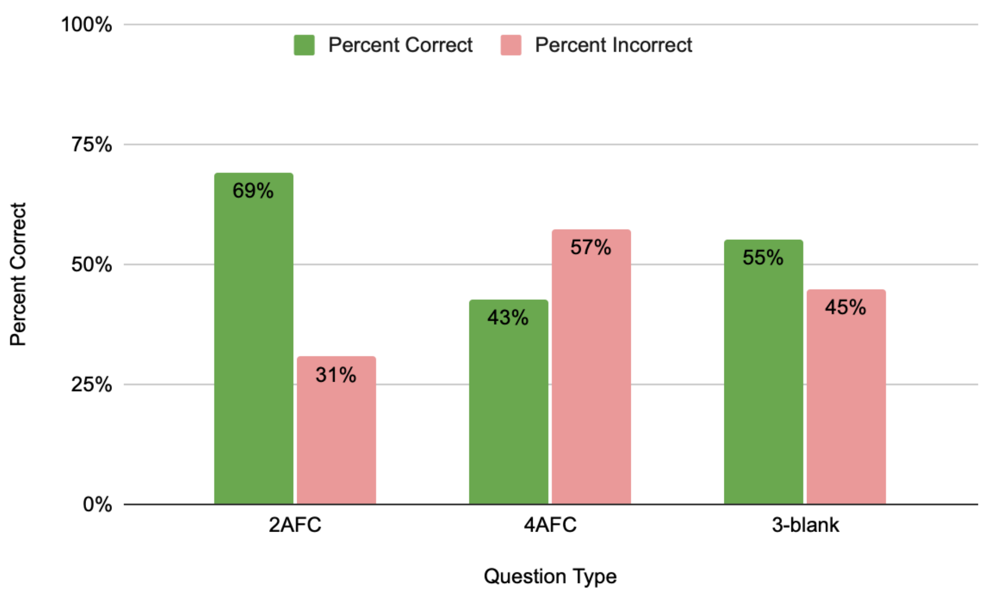
Number of Each Question Presented in Testing Phase of Auditory Statistical Learning Task

	Question Type				
	Normal	Easy AB	Hetero Acc	Both	Total
Number of Each	727	247	284	442	1700
	Question Difficulty				
	2AFC	4AFC	Blank		Total
Number of Each	733	478	489		1700

In terms of question difficulty, Figure 4 shows a general breakdown of performance, with 2AFC having the most correct responses and 4AFC having the least. In line with the trend of correct questions having a faster response time, Figure 5 shows the corresponding average reaction times by question difficulty.

Figure 4

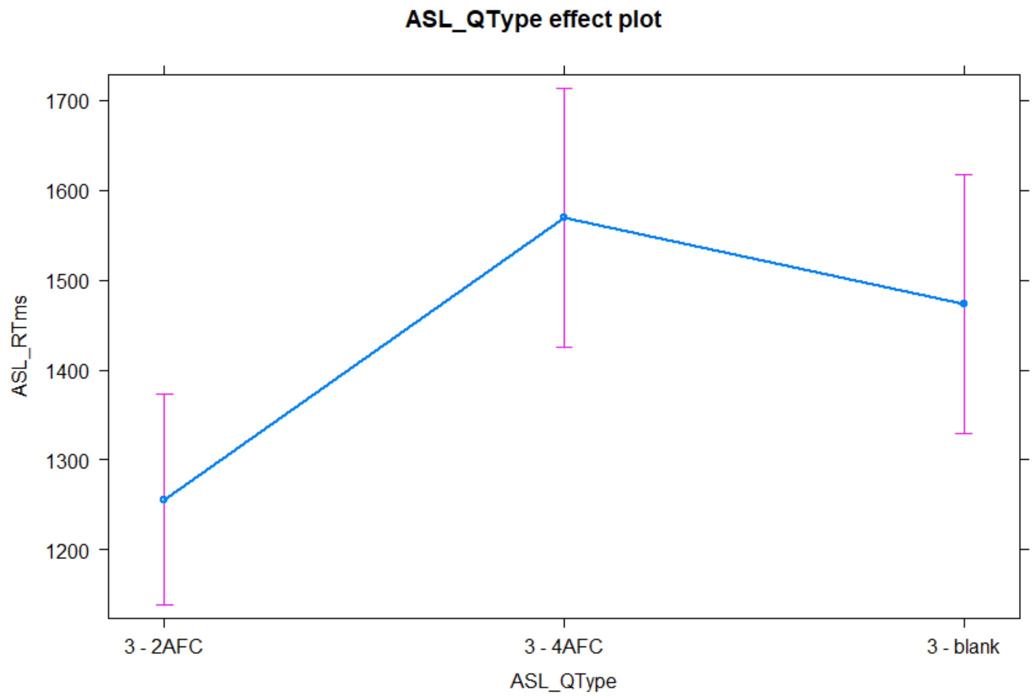
Overall Percent Correct for Each Question Difficulty Type



Note. Based on n = 1700 trials.

Figure 5

Overall Reaction Time for Each Question Difficulty Type



Note. Based on n = 1700 trials.

As for the breakdown and results by question type, this is reported later in the context of language background and accent familiarity findings.

Lastly, aSL performance by reaction time was further analyzed with a full regression model including the following factors: correct response, question difficulty (default response of 2 AFC), question type (default response of Both Correct), language status of English as L1 (default response of English L2), non-English use at home, non-English use socially, general accent comfort, and frequency of interactions with accented speech, as shown in Table 3.

Table 3

All Factors Predicting Reaction Time on aSL Task

	Estimate	Standard Error	*t*	*p*	
(Intercept)	1646.648	168.735	9.756	< .001	***

ASL_Correct	-146.94	55.014	-2.671	0.00764	**
ASL_QType3 - 4AFC	144.949	62.632	2.314	0.02077	*
ASL_QType3 - blank	127.531	65.096	1.959	0.05026	.
ASL_LegendEasy AB	-202.940	86.881	-2.336	0.01962	*
ASL_LegendHetero Acc	-160.162	83.845	-1.910	0.05627	.
ASL_LegendNormal	-56.333	65.443	-0.861	0.38947	
EnglishL1_binaryY	-252.642	87.012	-2.904	0.00374	**
NonEng_home	-4.487	7.216	-0.622	0.53414	
NonEng_social	-6.493	4.255	-1.526	0.12722	
GeneralAcc	-43.674	17.683	-2.470	0.01362	*
FreqInteractionsTotal	19.772	11.301	1.75	0.08035	.

—

Significance codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 ' . ', 0.1 ' '
Note. R^2 = .018, p = < .001, model residual standard error is 1023 on 1688 DF.

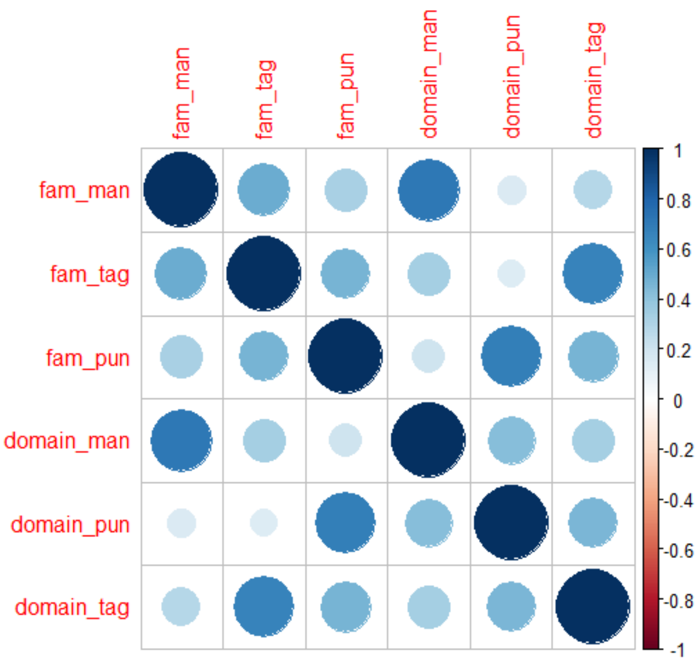
Of all the factors analyzed, correctness and L1 status have the most significance in predicting reaction time, with general accent comfort as a significant predictor as well. Question difficulty of 4AFC and question type of 'Easy AB' also were significant. Model fitting procedure examined which variables most impacted the variance predicted in order to optimize the model.

Accent Familiarity Questionnaire

Participants all rated their familiarity with Mandarin-, Tagalog-, and Punjabi-accented English on a sliding scale to 100, and they also rated familiarity with additional accents that they checked off as being familiar. Participants checked off the domains in which they regularly encountered speakers with a Chinese, Tagalog, and Punjabi accent as well. Domains of encountering accented speech were tabulated out of 10 options for each accent condition and included home, school, work, and in the community. Figure 6 shows the correlation plot between familiarity ratings and number of domains checked for each accent.

Figure 6

Correlations Between Familiarity and Domains for Mandarin, Tagalog, and Punjabi Accents

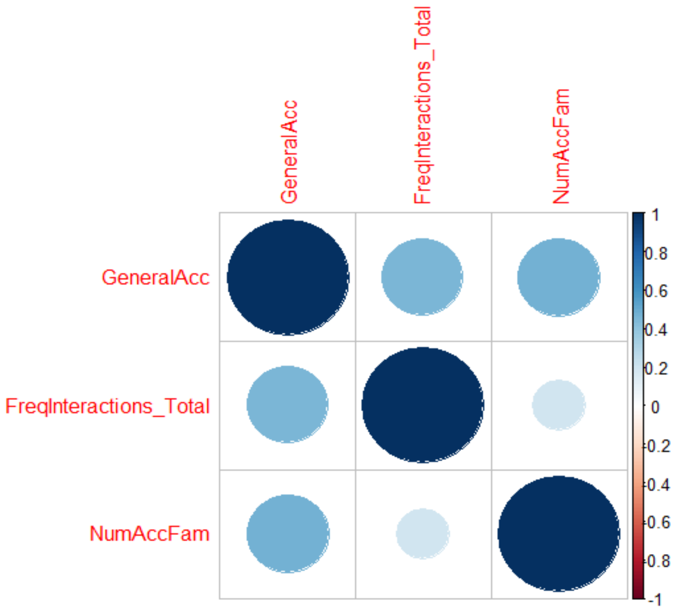


Note. Color and size of the dots relate to the strength of the correlation. Blue dots indicate a positive correlation, while red dots indicate a negative correlation; the larger the dot, the higher the correlation.

As shown by the size and shade of the dots, familiarity ratings and number of domains have a strong positive correlation for each accent. Similarly, participants' general accent comfort, frequency of interactions with accented speech in general, and number of accents identified as familiar were all positively correlated as well, as shown in Figure 7.

Figure 7

Correlations Between General Accent Comfort, Frequency of Interactions, and Number of Familiar Accents

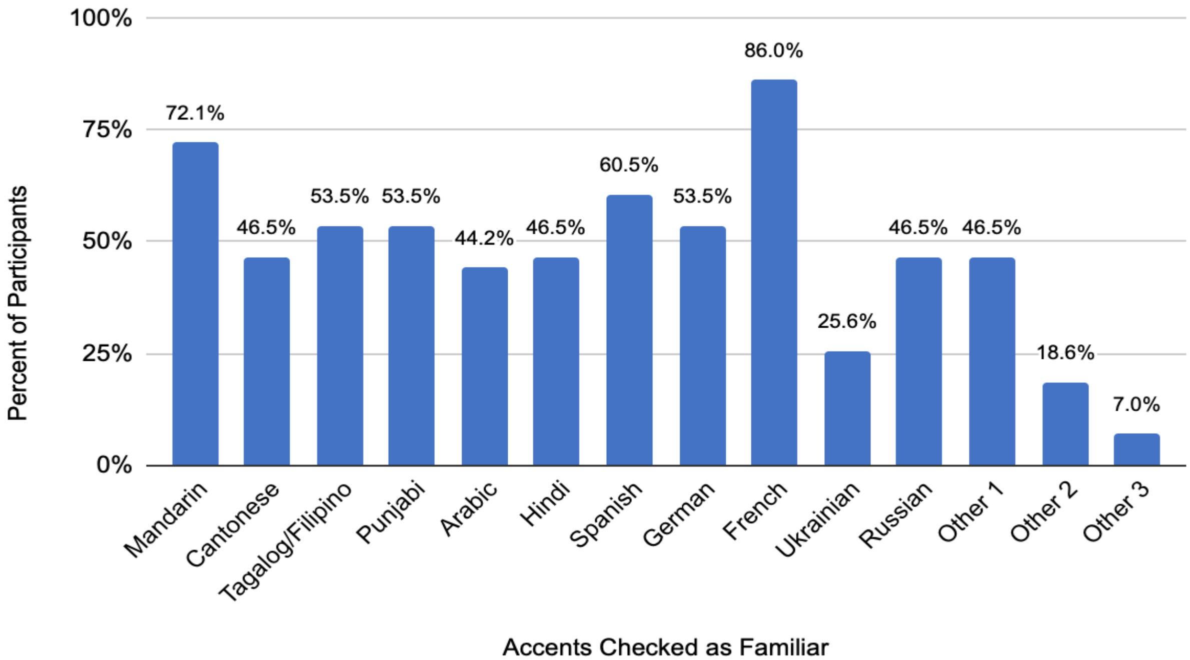


Note. Color and size of the dots relate to the strength of the correlation. Blue dots indicate a positive correlation, while red dots indicate a negative correlation; the larger the dot, the higher the correlation.

As for the accents that participants were the most familiar with, French was the most common with 86% of participants indicating familiarity, followed by Mandarin at 72%. See Figure 8.

Figure 8

Responses from "Which Accents Are You Familiar With? Check All that Apply"

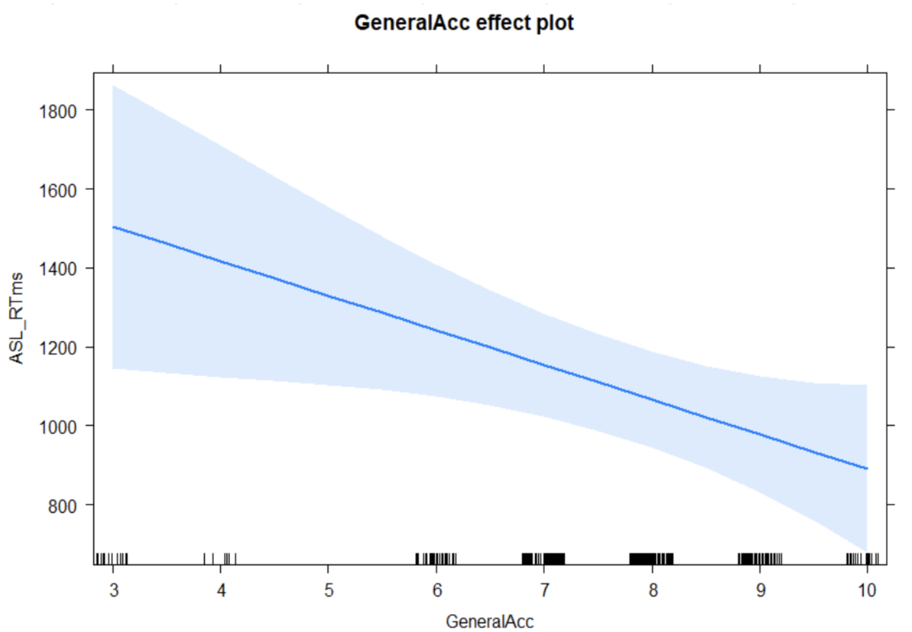


Accent Familiarity

Beyond the general findings of the Accent Familiarity Questionnaire, a few analyses were conducted to examine the role of accent familiarity in predicting performance on the aSL task, to address the first research question querying the same. First, it was found that neither general accent comfort nor frequency of interactions with accented speech alone were significant predictors of reaction time on the aSL task. However, in the full model described above, accent comfort was a significant predictor, showing that individuals more comfortable with accents were faster to respond to all questions. Moreover, when looking at reaction time on individual question types, a higher general accent comfort predicted faster responses on the 'Hetero Acc' questions specifically (b = -87.69, SE = 37.5, t = -2.338, p = .0201); see Table S4. Figure 9 shows this relationship.

Figure 9

General Accent Comfort Predicting Reaction Time for 'Hetero Acc' Questions



Note. General accent comfort is plotted on the *x*-axis and was self-reported on a scale from 1 to 10. Model residual error is 1012 on 282 DF, R^2 = .0155.

A higher frequency of interaction with accented speech was also a predictor for faster reaction time on the 'Hetero Acc' questions (b = -50.31, SE = 23.34, *t* = -2.155, *p* = .032). However, general accent comfort or frequency of interaction was not a significant predictor for the other question types.

To better understand the variables of general accent comfort and frequency of interactions with accented speech, we examined how participants' ratings of their familiarity with Mandarin-, Tagalog-, and Punjabi-accented English influenced their comfort and frequency of interaction ratings. It was found that Mandarin familiarity primarily influenced accent comfort, while Tagalog familiarity primarily influenced frequency of accent interactions. A linear regression model with Mandarin, Tagalog, and Punjabi familiarity predicting general accent comfort found that Mandarin familiarity was trending towards significance as a predictor (*p* = .0574), though the overall model was significant (R^2 = .1906*, p* = .0103). Mandarin on its own significantly predicted general accent comfort (b = .0237, SE = .0079, *t*

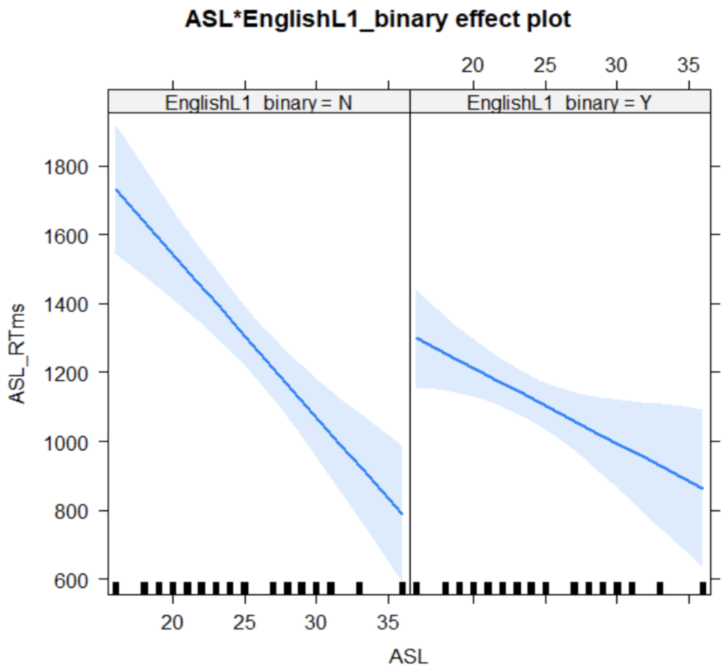
= 2.979, p = .0048). As for frequency of interactions, Tagalog familiarity trended towards significance when examining all three accents (p = .0676), and this model was significant as well (R^2 = .1767, p = .0141). Tagalog on its own did significantly predict frequency of interactions (b = .039, SE = .012, t = 3.247, p = .0023). The role of accent familiarity, whether general comfort or frequency of interactions, will be detailed further in connection to language background findings.

Language Background

To further investigate the role of language background on aSL performance as queried in the second research question, and because English L1 status was a highly significant predictor in the full model, several analyses were conducted to delve into these complex relationships more thoroughly. First, a linear regression model was conducted with overall aSL score (b = -47.168, SE = 8.855, t = -5.236, p = < .001) and English L1 status (b = -837.957, SE = 313.958, t = -2.669, p = .007), and both factors were found to be significant predictors of aSL reaction time. The interaction between these factors was also significant (b = 25.379, SE = 12.590, t = 2.016, p = .043); see Table S5 in Supplementary Materials. Figure 10 shows the differences in reaction time by language background and as a function of total score.

Figure 10

Interaction Between aSL Score and English L1 Status on Reaction Time

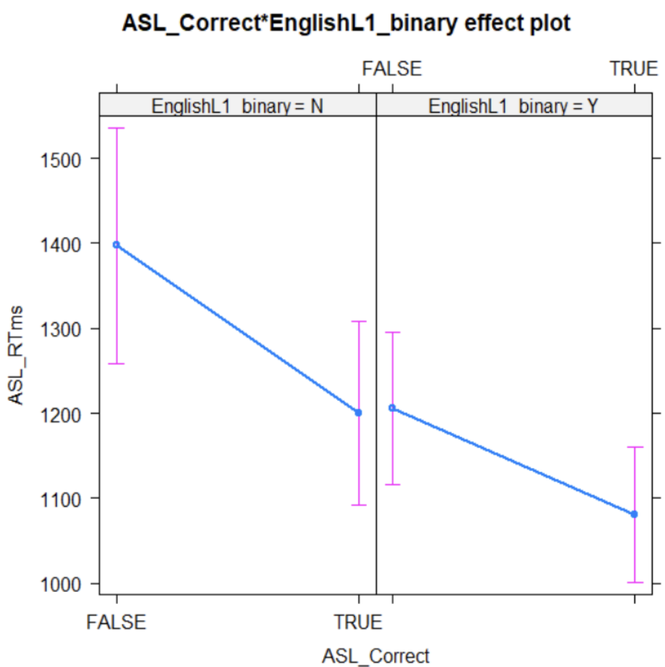


Note. Reaction time in milliseconds is displayed on the *y*-axis, score on the aSL task out of 42 is displayed on the *x*-axis, and the shaded area denotes the confidence interval. The graph on the left shows participants who do not have English as a first language, while the graph on the right shows participants who do have English as their first language.

The primary difference is in participants with low aSL scores, as the English L2s have a much slower response time than English L1s. However, for the good statistical learners with high scores, the graph shows little difference in reaction time, regardless of first language. As for the general trend between faster reaction time and correct trials, the same held true when examining English L1 versus L2 status separately; however, L1s were still faster overall for both correct and incorrect responses, as shown in Figure 11. Correctness (b = -197.15, SE = 89.92, t = -2.192, p = .0285) and English L1 status (b = -190.80, SE = 84.27, t = -2.264, p = .0237) were both significant predictors in this linear regression model as well; see Table S6.

Figure 11

Interaction Between Correct and Incorrect Trials and English L1 Status on Reaction Time



Note. The pink lines denote error bars. Correct trials are marked as TRUE on the *x*-axis.

The graph shows that those with English as an L1 responded at approximately the same speed to questions which they got incorrect compared to questions that the English L2 group got correct.

However, even though the English L2 participants had a slower reaction time overall, they had a higher percent correct for individual trials compared to English L1 participants. Table 4 shows the number of trials answered correctly by English L1 status.

Table 4

Correct Trials by English L2 and L1

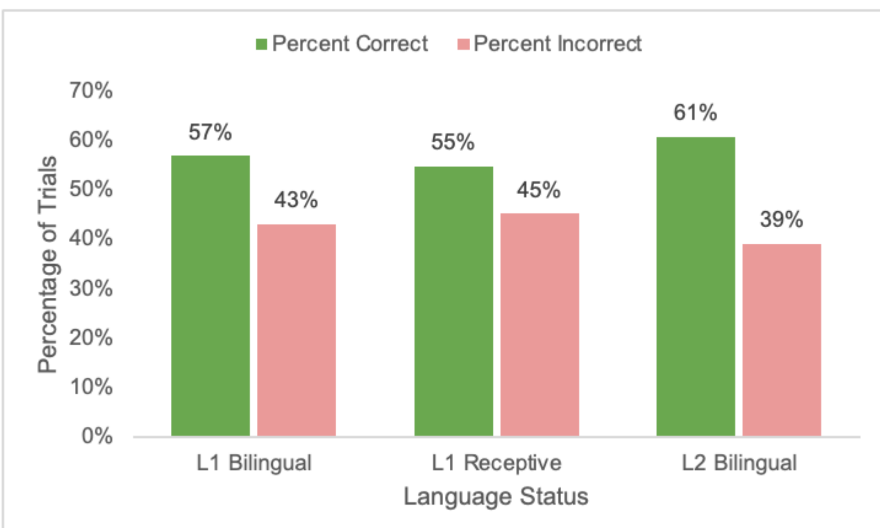
	English L2	English L1
Correct Trials	345	636
Incorrect Trials	211	508
Total	556	1144

Percentage Correct	62.05%	55.59%

A chi-square test of independence was run between the L1 and L2 groups, and there was a significant relationship between the variables, χ^2 (1, n = 1700) = 6.39, p = .011, such that English L2s indeed had a significantly higher proportion of correct trials. Overall correct trials were also examined by language status, whether L1 English bilingual, L2 English bilingual, or L1 English receptive bilingual. Figure 12 shows that English L2 bilinguals had the highest percentage correct overall on individual trials; however, the chi-square test between the three groups was not significant, χ^2 (2, n = 1554) = 4.305, p = .116.

Figure 12

Overall Percent Correct on Individual Trials by Language Status

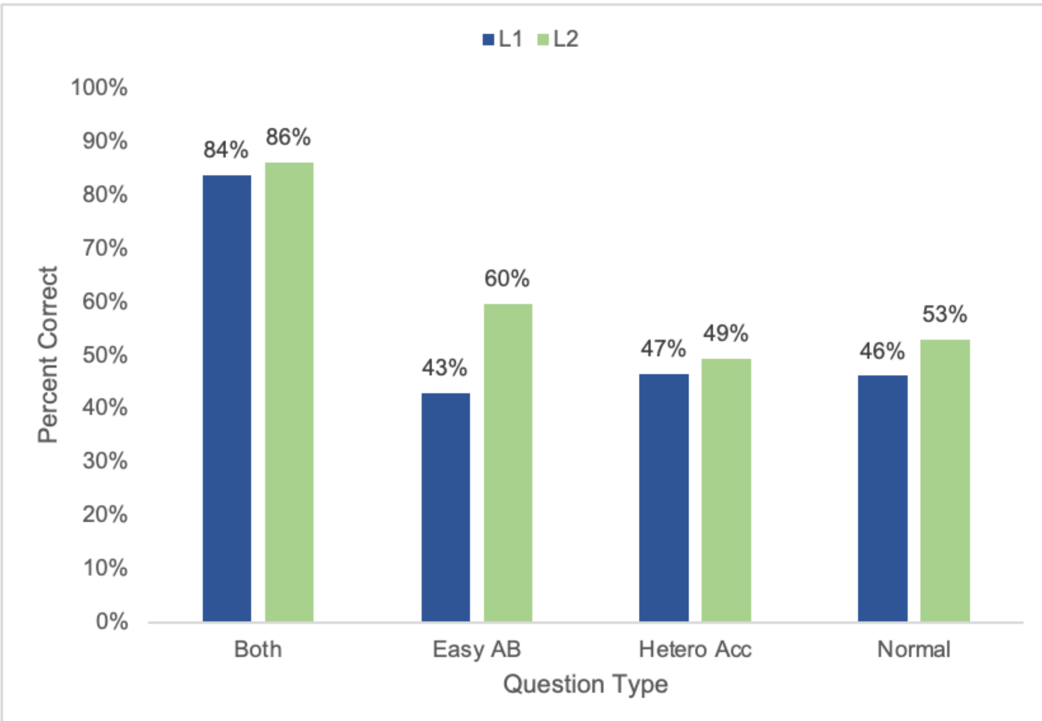


Note. n = 1554 trials.

Further to overall correct trials, examining specific aSL question type and difficulty by language background also provides more nuanced results. Performance on each question type was analyzed by percentage correct according to English L1 status. Figure 13 shows that participants were most successful on the 'Both' question type, which had two correct answers. A chi-square test was run between groups for each question type, and English L2 participants' performance was significantly better than English L1s' for the 'Easy AB' question type, χ^2 (1, n = 1700) = 6.13, p = .013.

Figure 13

Percent Correct for Each Question Type by English L1 Status

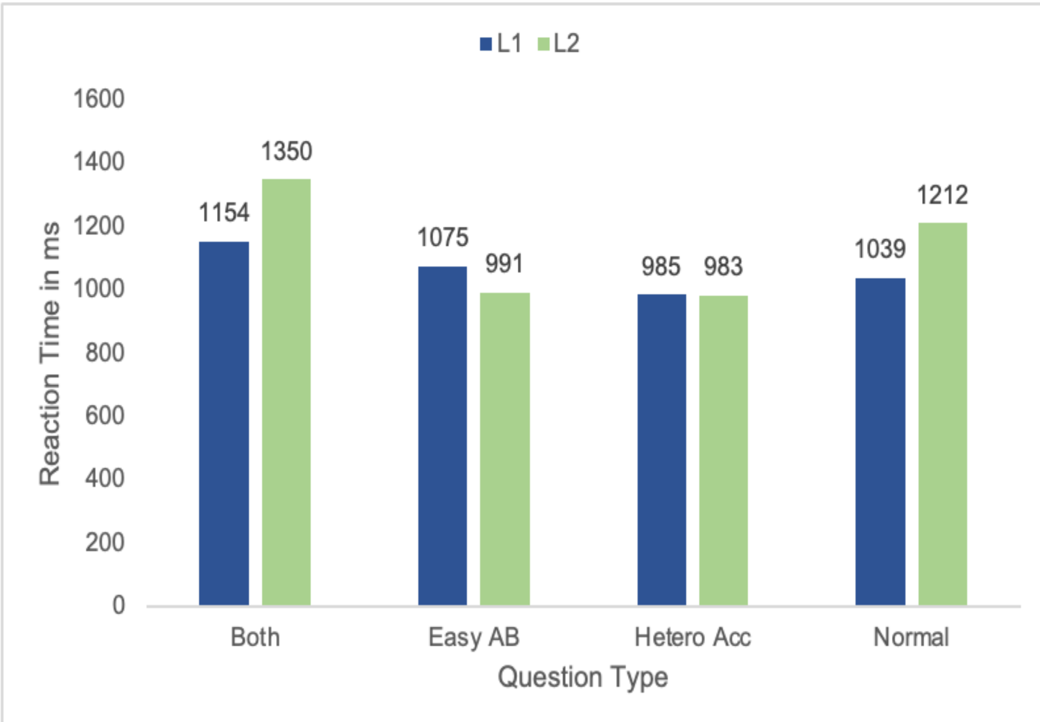


Note. n = 1700 trials.

In comparison, reaction time on each question type by L1 status is shown in Figure 14.

Figure 14

Average Reaction Time for Each Question Type by L1 Status



Note. This is the average reaction time for correct trials, shown in milliseconds, n = 981 trials.

These graphs show that even though English L1s were faster to respond to the 'Both' and 'Normal' questions, they were less correct overall. Linear regression confirmed that English L1 participants were significantly faster in responding to the 'Normal' questions (b = -193.44, SE = 80.63, t = -2.399, p = .0167). As for question difficulty, English L1 participants were also faster to respond to the fill in the blank questions compared to English L2 (b = -208.53, SE = 98.57, t = -2.115, p = 0.0349).

Lastly, the full regression model with all factors predicting reaction time on the aSL task was disaggregated to compare English L1 and L2. For English L1 participants, the only significant factor was the question difficulty level of 4AFC (b = 147.0725, SE = 73.2134, t = 2.01, p = .0448); correctness of trials, general accent comfort, and question type of 'Easy AB' did not significantly predict reaction time in this model unlike the full model; see Table S7 in the Supplementary Materials. When examining correct trials alone, only question difficulty type of 4AFC (b = 297.116, SE = 97.320, t = 3.053, p =

.00236) and 3-Blank (b = 243.868, SE = 90.963, t = 2.681, p = .00753) were significant predictors of reaction time, though the model itself was still significant (R^2 = .0164, p = .022). However, for incorrect trials alone, no factors were significant and the model was not significant (R^2 = -0.0086, p = .8615); see Tables S8 and S9.

Yet, for English L2 participants, the linear regression model showed that general accent comfort (b = -216.69, SE = 56.54, t = -3.833, p = < .001), frequency of interactions with accented speech (b = 81.15, SE = 28.02, t = 2.896, p = .0039), non-English social use (b = -28.49, SE = 11.39, t = -2.502, p = .0126), and correctness of trials (b = -213.43, SE = 103.86, t = -2.055, p = .0403) were all significant predictors of reaction time; see table S10. This indicates that in the original aggregated L1 and L2 model, factors more related to question type and difficulty were predicting variance from the L1s, while factors related to accent comfort and interactions were driving variance prediction from the L2s. When examining correct trials alone for English L2s, general accent comfort, frequency of interactions with accented speech, and non-English use socially are still significant, while question type of 'Hetero Acc' became significant (b = -397.957, SE = 190.584, t = -2.088, p = .03754); see Table S11. For incorrect trials, question difficulty level of 4AFC and 3-Blank significantly increase response time, while general accent comfort still significantly decreases response time (b = -218.564, SE = 92.18, t = -2.371, p = .0187). The model itself is significant as well (R^2 = .0429, p = .0359); see Table S12.

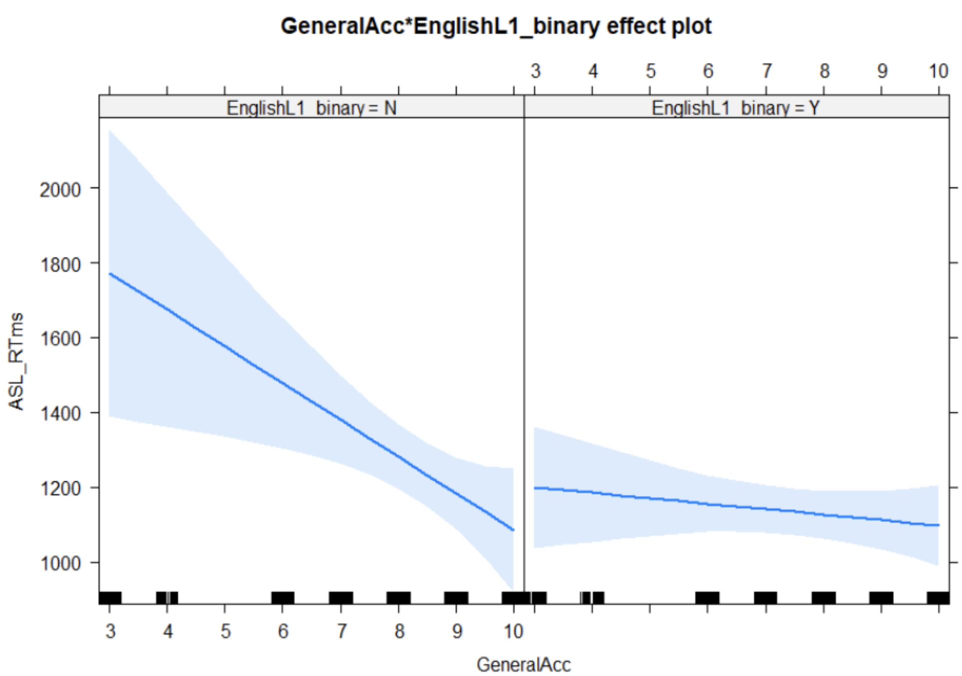
Connections Between Accent Familiarity and Language Background

Finally, to address the third research question of whether accent familiarity and language background interact in their influence, the connection between these two factors was analyzed through a number of measures. Before examining the impact on aSL reaction time, linear regressions were conducted between language and accent variables. Receptive bilinguals were found to have a significantly lower general comfort with accents compared to bilinguals (b = -0.224, SE = 0.088, t = -

2.541, p = .0111). There was also a significant difference between receptive bilinguals and bilinguals on frequency of interactions with accented speech, with receptive bilinguals interacting with accented speakers much less often (b = 1.5, SE = 0.133, t = -11.27, $p < .001$, R^2 = .0776); see Figure S1 in the Supplementary Materials. Next, linear regression was conducted between general accent comfort and English L1 status predicting reaction time on the aSL task; both accent comfort (b = -98.28, SE = 37.53, t = -2.619, p = .009) and English L1 status (b = -823.55, SE = 333.76, t = -2.467, p = .0137), were significant, leading to faster response times; see table S13. The interaction between these factors was significant as well, and the relationship is shown in Figure 15.

Figure 15

Interaction Between Accent Comfort and English L1 Status on Reaction Time



Note. Participants rated their general comfort in interacting with speakers with an accent out of 10, as shown on the *x*-axis.

The graph shows that especially for English L2s, the higher participants rated their comfort in interacting with non-native speakers of English, the faster their response time was. This finding provides an apt summary of the interconnectedness and relationships between the three variables investigated in this study.

DISCUSSION

This thesis aimed to investigate the relationships between accent familiarity, language background, and statistical learning ability, using a variety of measures and analyses to examine how they interact. A novel auditory statistical learning paradigm and an accent familiarity questionnaire were developed in order to help test these relationships. Thus, the discussion will focus on the following four objectives: commenting on the methodological implications and validity of the novel aSL task and accent familiarity questionnaire, and speaking to each of the three research questions presented above: whether accent familiarity, language background, or a combination thereof impact the ability to accurately learn patterns across accented phonemes.

Task and Questionnaire Validity

Auditory Statistical Learning

While the structure of the auditory statistical learning task was similar to the visual statistical learning task, which was developed by Kirkham (2002), described in Siegelman, Bogaerts, Kronenfeld, et al. (2018), and validated throughout the literature, the stimuli and question types were novel to this experiment. Therefore, a preeminent question was whether this task would perform as expected compared to other statistical learning tasks, whether participants would demonstrate any indication of learning and perform above chance level, and whether participants might experience ceiling or floor effects based on the difficulty. The results showed that the aSL task did perform as expected, based on the normal distribution of scores and normal skewness (.53) and kurtosis (.74). The mean was roughly three points lower than the mean for the vSL task, and many participants verbally indicated that they found the aSL task more difficult. However, the reaction time data was strongly correlated with accuracy data, showing that even if participants thought they did not perform well, the test still functioned as expected. Participants were faster to respond to trials which they got correct and slower for responses which were incorrect; similarly, those with higher scores had faster responses overall. This is congruent with other statistical learning tests that have been validated in the literature, which have found that good statistical learners are faster to respond in general (Hunt & Aslin, 2001). The relationship with response time and accuracy held true across different question difficulty levels as well, whether 2AFC, 4AFC, or 3-blank. The 4AFC questions proved to be the most difficult based on percentage correct, and they also had the slowest reaction time overall.

The other aspect that speaks to the validity of the aSL task was its relationship with performance on the vSL task. The only variable to significantly predict overall aSL score was vSL score, as they were positively correlated and the regression model was significant. This shows that the sample's performance on the visual task, which has been validated, was congruent with the performance on the auditory task. While correlation between overall scores is not enough to fully speak to the theoretical implications of the statistical learning mechanism, particularly regarding the domain-general versus domain-specific argument of statistical learning ability, it does point towards SL skill being stable and consistent within individuals. For these purposes, it helps to confirm that the aSL task functioned similarly to how an established SL task functions. Lastly, it bears noting that throughout the analyses run with aSL score and reaction time, the adjusted r-squared values were all quite small. However, because statistical learning is so multi-componential in nature, it is typical to have a small r-squared value. The models described above are still significant, and the small r-squared is still informative such that it points to the existence of a relationship. With so many unknown factors in the burgeoning field of statistical learning, any variance that we can predict and explain is worthwhile. As such, the adjusted r-squared value for vSL score predicting aSL was .1118, and this was one of the highest values throughout the models. To have 11% of the variance in aSL performance predicted by vSL performance speaks to the unitary facets of statistical learning in general, in addition to helping validate the aSL task.

Accent Familiarity Questionnaire

The other methodological contribution from this research was the creation of a novel questionnaire assessing accent familiarity. While further research and revision is needed to establish this as a reliable and valid measure, there are proponents that suggest preliminary reliability and utility of this tool. First, many of the measures of the Accent Familiarity Questionnaire (AccFamQ) were correlated with each other, showing that participants responded consistently across different questions. General accent comfort, frequency of interactions with accented speech in general, and number of accents identified as familiar were all positively correlated. There was even a positive correlation between the number of additional languages participants knew, as identified on the LSBQ, and general accent comfort; this also contributes to reliability as congruent answers were noted even across questionnaires. Familiarity with each accent, whether Mandarin, Punjabi, or Tagalog, was also strongly correlated with the number of domains participants reported encountering that accent variety in, which is what would be expected. Second, it also is congruous that receptive bilinguals had a lower general accent comfort and frequency of interactions than productive bilinguals, who are more likely to be actively involved in language communities and interacting with other non-native English speakers. Finally, the responses on the AccFamQ were also in concordance with the sample's demographic and language characteristics. The highest percentage of people checked French as an accent they were familiar with (86%), followed by Mandarin (72%), which aligns with the fact that many of the receptive bilinguals had French as an L2, while many of the bilinguals had Mandarin or Cantonese as an L1 or L2 as reported on the LSBQ. Consistency across the measures of the AccFamQ and with items on the LSBQ points to aspects of reliability even in this first draft of the questionnaire.

Furthermore, the findings about Mandarin and Tagalog familiarity and their relationship to accent comfort and frequency of interactions also suggest aspects of ecological validity. It was found that Mandarin was a primary driver of accent comfort ratings, which can be explained by the relatively higher proportion of Mandarin-speaking individuals in the sample. Comfort interacting with accented speakers is likely derived in part from regular interactions with friends and family and from a language community speaking a similar variety of accented English, which Mandarin or Cantonese-speaking participants would experience. Secondly, Tagalog was found to be the primary driver behind participants' ratings of their frequency of interaction with accented speech. This is notable because Tagalog is the top non-official language spoken in Edmonton, where this research was conducted. Frequency of interactions reflects what is present in the environment, and there are many Tagalog speakers in the larger community which participants would frequently encounter and have interactions with. The fact that the results of the questionnaire reflect the characteristics of participants in the sample and their environment also points to preliminary reliability and validity of this novel measurement tool.

Accent Familiarity

In light of the discussion of the AccFamQ itself, what bearing do its results have on the first research question posed by this study: "Does accent familiarity impact the ability to accurately learn phonemes across accented speech?" Based on the parameters originally intended to answer this question, and the sample size, there was null effect. Neither general accent comfort or frequency of interactions on their own were significant predictors of either score or reaction time on the aSL task. However, it is worth noting that in the full linear regression model, accent comfort (GeneralAcc) was a highly significant predictor. In the full model for English L2 participants, accent comfort and frequency of interactions were significant predictors as well, particularly for correct trials. From a model fitting perspective, frequency of interaction was not originally included as a variable in the analysis. However, once it was added, general accent comfort became significant; frequency of interaction predicted a certain amount of variance, allowing general accent comfort to predict more. As with many aspects of this research, accent familiarity is a complex concept that interacts with many other variables, even variables within the context of accent familiarity. Additionally, research has shown that a lack of accent familiarity leads to longer processing times (Adank et al., 2009; Kang, 2012; Winke & Gass, 2013; see also Christia et al., 2012), which was also shown by our linear regression model between general accent comfort and L1 status in predicting reaction time (see Figure 14 in the results). The graph shows that a lower familiarity overall is related to slower response times for all participants, whether L1 or L2, and that response speed did increase with familiarity. This shows that accent familiarity does have some relationship with performance on the aSL task, even in reaction time alone.

An accent familiarity finding of note was that a higher level of general accent comfort and frequency of interactions both predicted a faster reaction time on the 'Hetero Acc' triplets. This means that participants with a higher level of familiarity were able to generalize learning across different accents presented in the same triplet. The 'Hetero Acc' question type was designed to examine exactly this—to see if the statistics of the transitional probabilities between phonemes could be learned even when presented in different accents, or if participants would treat the phonemes as discrete units based on accent condition. Although multiple accents in the same word is not something that occurs naturally in a real-world context, the fact that a relationship exists between accent familiarity and highly accented stimuli is noteworthy. However, participants with a greater level of familiarity were not necessarily more correct on these questions, just faster to respond to them. As with all the variables predicting a faster reaction time, lack of a higher overall score or greater accuracy does not mean that learning has not occurred. Faster reaction time points to an intermediate measure of learning, through better recognition, more confidence, and an increased processing speed, which are all online measures of learning rather than offline. Because a higher level of accent familiarity meant a faster reaction time for the most accented triplets, there does indeed seem to be an impact on learning across accented phonemes, just not learning how it was expected to be measured.

The findings in relation to the research question at hand also illuminate the complexity in defining and quantifying familiarity, especially with accented speech being such a variegated and subjective concept. To operationally define accent familiarity for this study, two aspects were chosen: comfort and frequency of interaction. While both subjective self-report measures, comfort was more of a qualitative variable while frequency of interactions was a more quantitative variable. The variables were positively correlated with one another, $r(41) = .45$, $p = < .001$, but they did behave differently in the analyses in terms of the variance predicted by each. Of course, there are many ways that familiarity can be defined, and this shows the influence of question wording on trying to measure an abstract concept objectively. Question wording has implications for many facets of a questionnaire, including designing rating scales, ascertaining numeric quantities, minimizing or introducing bias, and optimizing responses (Krosnick & Presser, 2009; Glendall & Hoek, 1990). Instead of having participants rate their general comfort in interacting with speakers with an accent, there could have also been a question that had participants rate their general familiarity with accented varieties of English. This might have yielded a variable closer to what was originally intended. However, it was telling to have two aspects of accent familiarity to analyze, both qualitative comfort and quantitative frequency of interaction, even if they may have been too broad to predict aSL score or RT on their own.

Language Background

The second research question asked whether language background, particularly being bilingual, similarly impacts the ability to generalize learning across accented phonemes. Again, based on overall aSL score alone, there was null effect from the LSBQ, which was the composite factor score indicating bilingualism. This could be due to a small sample size, but likely is also due to the LSBQ score being too broad. When language background was broken down particularly into English L1 versus English L2 status, and the dependent variable was broken into RT and into question type subsets, there were significant findings, which do point to language background impacting learning. As with accent familiarity, the nuanced findings related to language background also speak to the complexity of trying to operationally define and quantify bilingualism. From the various facets of the LSBQ, we analyzed L1 status, non-English use at home and socially, and language status as English L1 bilingual, English L1 receptive bilingual, and English L2 bilingual. As Kałamała et al. (2023) discuss, bilingualism is highly complex to define, depending on indices such as age of acquisition, age of active communication, daily language use, diversity of daily language use, proficiency, and code-switching. Any of these variables, or interactions between them, may have had more of an impact on overall aSL score.

The overarching important finding related to language background was that those with English as their first language had faster reaction times overall, for correct and incorrect trials, but those with English as their second or additional language actually had a higher percent correct overall. The faster RT of English L1s suggests that they were more confident, had a faster processing speed, and could access representations of the phonemes more quickly. However, the L2s showed more offline evidence of learning, as even though they had to think about the questions longer, they were ultimately more correct. It is worth noting that even though the English L1 group contained receptive bilinguals and bilinguals, their mean LSBQ score was 0.046, and the cutoff to be considered bilingual is 1.23; the mean LSBQ for L2s was 5.034. While it is not a perfect comparison of monolinguals versus bilinguals, the L2s can be generalized as more bilingual than the L1s. The fact that the L2s were more accurate could relate to the relationship between proficiently using more than one language and innate statistical learning ability, which will be discussed in greater detail below. However, another explanation for the slower RT of English L2s could be the processing differences and cognitive load of knowing and actively using two languages. According to a review by Roman and Gomez-Gomez (2022), research has shown that both languages are co-activated in the bilingual brain even when only one is needed, which places a demand on cognitive resources, inhibition skills, and mechanisms of language control. Needing to inhibit one's L1 in an L2 setting, which is the case of the English L2 participants, may explain the slower reaction time.

The other key finding with language background was the different factors predicting performance by English L1 and L2. The full linear regression model predicting reaction time for English L1s showed that only question difficulty had any impact, with 4AFC and 3-Blank significantly predicting an increase in RT. These questions had the slowest response time and lowest percentage correct. When splitting between correct and incorrect trials, there were no significant findings for the incorrect trials. This indicates that the variables of interest only have an effect on learning the statistical regularities, since they are only significant in the correct trials. The negative adjusted r-squared also confirms that the explanatory variables have no effect for incorrect trials. For L1s, it seems like none of the language and accent variables affect their RT, suggesting that their inherent SL ability is fixed – or influenced by other outside factors not captured in this experiment.

For the English L2s, however, both general accent comfort and frequency of interactions explained a faster reaction time, and their use of a language other than English socially also predicted faster reaction times on correct trials. Even across both correct and incorrect trials, general accent comfort was a significant predictor. This shows that accent familiarity matters much more for L2 speakers and is in fact associated with reduced processing times for the L2s alone. There was less of an impact of question difficulty or type, showing that L2s may have been able to learn more efficiently across numerous types of targets and foils than the L1s. Again, there were more significant variables of interest in the correct trials than the incorrect, showing their effect on learning the statistical regularities. For L2s, these included non-English social use and the ‘Hetero Acc’ question type. It is intriguing that participants who use English less often in societal contexts such as shopping, accessing services, school, and interacting with friends show a faster reaction time to accented speech. It could be argued that this represents a higher use of code-switching in daily life, which may have an impact on processing different ‘codes’ of accented phonemes. Further investigation into code-switching as a predictive variable would be necessary to examine this further. Additionally, faster RT on the ‘Hetero Acc’ questions could be due to a higher exposure to variability in accented varieties of English as an L2 speaker.

In addition to the ‘Hetero Acc’ question type, the other question subset that resulted in statistically different performance by English L1 and L2 participants was the ‘Easy AB’ question. This finding was counter to what had been predicted with this particular question type. The ‘Easy AB’ question was designed to be the easiest and show the most evidence of learning, as the target was always an Albertan-English triplet with 1.0 transitional probability between each syllable (considered an ‘easy’ triplet). It was predicted that if L1 English participants, monolingual or otherwise, did not have much familiarity with a variety of accents, let alone the accents in the study, they should be able to at least pick out the ‘unaccented’ triplet. However, the opposite seemed to be true, with the L1s performing worse on this question type. This could be because Alberta English was the most frequent accent presented throughout the familiarization phase, so the L2s may have been more effective at learning it as a salient statistical characteristic, contributing to their higher rates of overall correctness. The L2s are also used to navigating different patterns in one place within English more so than the L1s, which may contribute in general to their more accurate performance. Perhaps more presentations of each accented type would have led to different results. Interestingly, this was also the only question where L2s had a faster reaction time than L1s, but this is likely because there were more correct trials to calculate an overall average reaction time from. The question type ‘Easy AB’ was also statistically significant in the full model with both L1s and L2s, pointing to the salience of the performance differences, with variance driven by English L2s. The fact that L2s, who were more bilingual in general, performed significantly better on the ‘Hetero Acc’ and ‘Easy AB’ questions—and were more accurate overall despite being slower— shows that there is indeed an impact of language background on learning across accented phonemes.

Connections Between Accent Familiarity and Language Background

Finally, the third research question inquires whether accent familiarity and language background interact in their influence, or if one is a stronger predictor than the other. The results showed that there is indeed an interaction between language background and accent familiarity. As has been discussed from the full linear regression models already, general accent comfort and frequency of interaction significantly improved reaction time for L2s. With the linear regression between accent comfort and English L1 status predicting reaction time (see Figure 14 in the results), it was shown that L2s’ reaction time was highly sensitive to change in accent familiarity levels. Less of an effect was seen for English L1s, as their RT did not change as drastically; this corroborates the idea that other variables are driving performance for the L1s, and the mechanism they use to process phonetic variation is still an area to be explored. The L2s, on the other hand, relied more on their exposure to multiple different varieties of English. At high familiarity, L2s even surpassed the L1s in overall reaction time. Underlying this finding could be the L2s’ exposure to variability and their need to inhibit and shift languages constantly, which is likely also due to the connection between bilinguals and accent familiarity. As discussed in the introduction, bilinguals may have an advantage in processing accented speech due to their greater flexibility of mapping input onto representations, the high instance of phonetic variability they are exposed to, inhibition, selective attention, and task switching proficiencies, and even a heightened categorical perception ability (Archila-Suerte et al., 2012; Bent & Atagi, 2017; Bialystok et al., 2009; Levy et al., 2019). Greater exposure to accents, thus greater familiarity, interacted with the advantages inherent to bilinguals and created an additive effect. This led to a greater boost in processing speed for L2s compared to accent familiarity as a factor on its own, which was seen in the English L1s' more invariant performance. It is difficult to conclude whether accent familiarity or language background is a *better* predictor of learning across accented phonemes, as they are intertwined in their prediction, as are many of the variables explored by this research.

General Discussion

This study focused on numerous unanswered questions in the fields of accent perception, bilingualism, and statistical learning ability. Overall, the findings were more nuanced and complex than what was originally predicted, opening the door to several other areas for further discussion and future research. The results of this study illuminated new ways to address the research questions presented, rather than a simple accepting or rejecting of the proposed hypotheses. Several uncertainties and considerations inherent to the topics outlined in this paper remain, including the necessary amount and type of exposure to accents, the role of bilingualism with an individual's phonological representations, the connections between bilingualism and statistical learning, and the nature of statistical learning ability. Overall, the rationale for this research was to explore individual cognitive factors that enable a listener to better process accented speech. While the results suggest that inhibition and linguistic control as a function of bilingualism may play a role, alongside statistical learning ability, there are still many other considerations and avenues left to explore.

Accents

The first broad area implicated in this research is the type and amount of accent familiarity necessary to improve processing of accented speech. The accents used in the auditory statistical learning task were Mandarin, Punjabi, Tagalog, and Albertan English, and the sample did not have any participants who were highly familiar with Punjabi or Tagalog. There were, however, several participants who had high familiarity with Mandarin, and it was found that Mandarin familiarity was a driver of general accent comfort ratings. Consequently, higher accent comfort was significant in the full model as a predictor for faster reaction time. However, the greatest number of participants identified they were highly familiar with French; Tagalog and Punjabi familiarity was tied with German familiarity, and Spanish was higher than all three of these. It seems that for this population, overall familiarity with the variability found in accents was more of a factor than specific familiarity with Mandarin, Punjabi, and Tagalog. This seems to suggest that familiarity with accents in general, rather than specifically the accents used in the study, impacts processing. However, this hypothesis would have to be verified through methodology such as utilizing French-accented stimuli, having accented stimuli longer than three syllables, and testing comprehensibility rather than implicit pattern learning. As discussed in the introduction, research by Bradlow and Bent (2008) found that increased understanding of a speaker from a certain language background comes from prior exposure to that language background specifically. Yet, there were also studies which found that experience hearing L2-accented speech in general aided transcription and comprehension, and Baese-Berk et al. (2013) maintained that greater exposure to multiple accents, not just multiple speakers of a single accent, can lead to faster processing of accented speech. The results of this study corroborate that finding; at the level of accented phonemes, it seems that exposure to accents in general was enough to lend an advantage, rather than exposure to the particular accent the phonemes were from. This advantage was most notable in speakers of English as a second or additional language, for reasons discussed above.

Secondly, how much exposure is enough to be considered ‘familiar’? While familiarity with each Mandarin, Punjabi, and Tagalog was correlated with the number of domains participants encountered those accents in, the average number of domains identified was no greater than 3 out of a possible 10 for any of the three accents. Oftentimes, the domains indicated for these accents were ‘In Social Media’ and ‘In the Community.’ Frequency of interaction with Mandarin, Punjabi, and Tagalog was also low on average. This means that generally speaking, though participants did not frequently interact with speakers of Mandarin, Punjabi, and Tagalog nor did they encounter aforementioned speakers in numerous domains of their daily life, their general accent familiarity was still high enough to show intermediary evidence of learning with Mandarin-, Punjabi-, and Tagalog-accented phonemes. A study by Adank et al. (2009) came to a similar conclusion, suggesting that being highly familiar with an accent was not contingent on interacting with speakers of that accent on a regular basis. They found that participants who were exposed to the accent through social media and the university had similar processing times to accented stimuli compared to speakers of that accent (Adank et al., 2009). This opens up the question of how much exposure to an accent is enough and where does it need to come from to have an impact. Is consuming media with actors that have a certain accent enough, or does one need to know and interact with speakers from a particular accent? The literature is divided, as Bradlow and Bent (2008) found that short-term, explicit training with an accent has an effect on processing, while Adank & McQueen (2007) found that it does not, though these findings were in the context of regional accents. The results of the current study suggest that perhaps exposure even through social media and the community is enough to affect processing, though further research is indubitably necessary to see how far-reaching the effect of this familiarity is, especially with comprehension of accented speech.

Language

Another area for further discussion is the differences in phonological representations between L1 and L2 speakers and how this may have affected performance. Research by Vroomen et al. (1998) found that listeners use the phonological properties of their native language to aid in speech segmentation, relying on different phonetic and acoustic cues depending on their L1. Speech segmentation is an area highly implicated in statistical learning ability, as established by the seminal study by Saffran et al., (1996). This means that the English L2s in the present study could have been relying on their native phonological cues for segmenting and learning the regularities of the accented stimuli. Hence, the slower RT exhibited by the L2s could have been related to having to shift which cues they relied on and adjust to the phonetic characteristics of the stimuli presented. As for the English L1s, the inverse may be true in explaining their faster reaction time. The phonemes presented in the aSL task were largely specific to English, especially since phonemes not in the phonetic inventories of Mandarin, Tagalog, or Punjabi were used intentionally in triplets. For those who had English as a first language, accessing one's native phonetic inventory and phonological representations should have been advantageous, at the very least with the Albertan-accented triplets.

However, though the processing was faster for English L1s, the fact that L2s showed more evidence of learning through a greater overall percentage correct indicates that there is more going on. This is especially the case given the L2s were more accurate on learning the Albertan-English accented triplets than the L1s were. Vroomen et al. (1998) also cite the fact that monolinguals are often unable to shift their representation of phonological cues as they segment speech from a foreign language. While the English L1s in the present study were not solely monolingual, they did have a lower composite LSBQ score compared to the L2s, trending towards monolingualism. Perhaps the accuracy advantage demonstrated by the L2s points to their ability to shift cues more effectively, which may be related to the finding that bilinguals have greater flexibility in categorical perception (Archila-Suerte et al., 2012). It also could be due to additional factors affecting speech segmentation ability, such as inherent SL ability or exposure to the greater variation in speech, or factors not captured in this study. More research is needed to determine how one's underlying phonological representations differ as a function of L1, what cues individuals may be relying on, how this is affected by interference from one's L2 (Roman & Gomez-Gomez, 2022), and what the interactions might be with SL ability across accented phonemes.

Language and Statistical Learning

A question central to the interpretation and implications of this study is the directionality of the relationship between statistical learning ability and bilingualism. As was outlined in the introduction, bilinguals have been found to have a potential advantage on SL tasks compared to monolinguals (Wang & Saffran, 2014), likely due to the cognitive and neural differences associated with bilingualism, namely inhibition, task control, selective attention, and cognitive flexibility (Bialystok et al., 2009). The results of this study found that for good statistical learners with higher overall scores on the aSL task, language background did not seem to confer an advantage (see Figure 10 in the results), though the English L2s had a slightly faster reaction time than the L1s in the highest aSL scores. While the directionality of SL ability and language cannot be elucidated from these findings alone, they do lend support to the idea that SL ability could precede bilingual ability. As in, it could be SL ability driving the acquisition of and proficiency in an additional language, rather than an additional language driving SL performance. Research finding that SL ability has a distribution even in infants, who have not acquired language yet, corroborates this idea (Bulf et al., 2011; Saffran et al., 1999; Shafto et al., 2012). Research has also found that SL ability is a predictor of grammar acquisition in adults learning an L2 language (Chen et al., 2022). Perhaps those with a higher inherent SL ability are more likely to learn English as a second language. Among those who have had to learn English for academic purposes or after moving to

Canada, perhaps those with higher SL are the individuals most likely to be pursuing post-secondary academic studies at an English-speaking university and voluntarily participating in research. Since the neural basis of SL ability is still relatively unexplored, as is its longitudinal progression as individuals learn languages, it is unclear whether SL ability does modulate language proficiency.

Another question raised by these findings pertains to what level statistical regularities are learned at in L1 and L2 speakers. Does identifying and learning the statistical regularities happen at a more subconscious level than matching a phoneme to an individual's L1 or L2 representation? Are the regularities learned before needing to inhibit a competing language? For the less effective statistical learners with lower scores, reaction time was slower for L2s than L1s, suggesting that perhaps the cognitive demands of linguistic co-activation and inhibition were slowing their processing time and affecting learning of the regularities. Inhibition might have been occurring on a more subconscious level than statistical learning. For the good statistical learners with higher scores, though, the reverse may be true. Perhaps these learners were able to extract the regularities before needing to consider language classification, resulting in no impedance from a competing L1 phonological representation. As for the English L1 learners, this could be why the slope of the graph was less steep; because no inhibition of their dominant language of English was necessary, the difference in performance is not as prominent. Previous studies examining statistical learning in bilingual individuals have not demarcated by first or second language, nor have they considered differences in performance between low and high SL scorers within the bilingual group. Moreover, research has largely considered how SL ability affects acquisition of one's first language, not a second or additional language (Romberg & Saffran, 2010). As such, it remains to be determined whether SL ability informs bilingual ability, both in language proficiency and in the effect of inhibition on language processing.

Lastly, questions remain regarding how age of acquisition affects both bilingual proficiency and statistical learning performance. In the research about categorical perception and bilingualism, it was also found that those who learned their second language before the age of five performed better on within-categorization tasks of L2 sounds. That is, they were able to more effectively identify slightly different pronunciations of the same phoneme, such as /p/, to be members of a single category (Archila-Suerte, et al., 2012), which may relate to more effectively processing accented speech. For the L1 and L2 bilinguals in this study, the age at which they learned their second language was not further analyzed, but this would be a variable of interest for future research and analysis. As for statistical learning performance, the literature has not yet come to a consensus regarding the stability or flexibility of SL within an individual. Some studies have found SL ability to be a stable measure, while others have shown improvements with age or in certain domains but not others (Siegelman & Frost, 2015; Saffran et al., 1997; Raviv & Arnon, 2018). If bilingualism does precede statistical learning ability, does learning a language later in life relate differently to SL than learning a second language from birth? Further, research has shown that an L2 learned after the critical period is related to different activation patterns and areas in the brain compared to a language learned in early childhood (Xu et al., 2017; see review by Tao et al., 2021). Does brain activation for SL differ for L2 sequential bilinguals as compared to simultaneous bilinguals, and what might these ramifications be for processing accented speech? Though the research regarding the neuroimaging of SL is still in its infancy, and there are no empirical answers to these questions as of yet, the relationships between SL and language in the context of this study's results provide ample room for further query.

Limitations

Two of the most apparent limitations of this study relate to the auditory statistical learning task design and the sample size. As for the novel aSL task, the decisions made in its creation raise some potential limitations and confounds. Because natural conversational speech was used, which contained very brief phonemes, the duration of the vowel in each syllable was elongated with the consonant burst left unaltered. The decision to alter the vowel portion of the CV phoneme was made primarily because of the repetition of vowels across the stimuli. There was no repetition of any consonants, meaning that they were the most informative aspect of each CV syllable, and they also carried features of accentedness, such as VOT. A related potential confound comes from the splicing of the phonemes from existing words. The prosody, whether rising or falling intonation, word stress, and duration, varied depending on where the word was in the context of the reading passage; it also varied depending on the speaker. When splicing a phoneme from a word, there were additional acoustic features coming from a discourse level that may have been at odds with the segmental features inherent to particular accents. As outlined in Vroomen et al. (1998), vowel harmony and word stress are salient cues to phonetic boundaries, and each triplet had varying vowels and stress as they were spliced together from a continuous reading passage. Decisions to make the triplets were based more on the presence or absence of phonemes across native phonetic inventories rather than considering the harmony of the vowels or a predictable stress pattern. Similarly, while all efforts were made to minimize coarticulation effects, there may have been some elements of consonant assimilation or nasalization, particularly when extracting a phoneme from a CVC word. The acoustic manipulations necessary to create the aSL stimuli could have been potential confounds or affected what was being learned in the task. It is possible that the task was not the most representative of or sensitive to accents.

Another significant limitation comes from the sample size and composition. First, recruitment through the University of Alberta undergraduate and graduate channels meant that the age, sex, and general characteristics of the sample were perhaps not the most representative of or generalizable to a larger population. Not many participants were highly familiar with Tagalog or Punjabi, and there were no speakers of either language (one receptive bilingual of Tagalog). A wider sample more representative of the linguistic population of Edmonton may have yielded different results. With only 43 participants overall, it would be beneficial to replicate and extend the experiment with more participants, and this would allow for more in-depth analysis as well. The intent was to compare equal numbers of monolinguals and bilinguals; however, there were only five monolinguals. Given the language backgrounds of the participants collected, decisions were made to stratify the sample according to English L1 and L2 status instead. The English L1 group was made up of 5 monolinguals, 14 receptive bilinguals, and 10 productive bilinguals, while the L2 group was made up of 13 productive bilinguals and 1 receptive bilingual. With the size of the L1 group, there was not enough statistical power to compare L1 monolinguals, L1 receptive bilinguals, and L1 bilinguals as separate groups. Underscoring the conclusions drawn from this research regarding L1s versus L2s is the caveat that the L1s were a heterogeneous group, and there was not an accurate representation of pure monolinguals to compare to other language backgrounds. However, as was discussed earlier, the average LSBQ score of the L1s was near the level to be considered monolingual, while the L2s' score was at the level of being considered bilingual.

Furthermore, there were eight simultaneous bilinguals in the sample who learned English and another language both from birth. Because there was not enough statistical power to run simultaneous bilinguals as a separate group, these participants' L1 was considered based on the dominant language of where they grew up. For most participants, this was Canada, meaning that some of the English L1s were not sequential bilinguals. It is recognized that classifying simultaneous bilinguals as English L1 bilinguals is an oversimplification of a complex linguistic issue, colored by the linguistic Anglo-hegemony of Alberta and the prestige that English holds as a dominant language in society. There are also implications here with the idea of age of acquisition (AoA) affecting bilingual proficiency.

Classifying participants as L2 implies that their additional language(s) were learned after the critical period of language acquisition, and this is not necessarily the case for all participants. If there were more participants, stratification based on different linguistic variables could have been more nuanced, and more participants in each group would have allowed for more accurate comparisons and generalizable conclusions. With a larger sample size, perhaps variables such as the LSBQ composite score would become significant predictors of aSL score as well.

Implications and Future Directions

Despite the limitations, the significance of this project in its contribution to the field is not to be understated. First and foremost, this is the first study of its kind to use natural accented language as statistical learning stimuli, with four different accent conditions. Previous statistical learning research has focussed largely on synthetic speech and certainly has not focussed on the prevalent minority languages of a given community. Given the considerations in designing the novel stimuli and task, the fact that participants did still learn the regularities is an important contribution; performance was above chance level, showing the task can be used in follow-up studies. Because accented speech has rarely been used in SL paradigms, this research provides an important foundation. The statistical learning of accented speech can be investigated along various different levels, from the phonetic-acoustic features such as VOT and vowel formants to prosodic cues such as word stress and intonation patterns. There are many linguistic aspects that comprise an accented variety of English, and this project opens the door for future research to examine which factors contribute the most to efficiently processing an accent from a statistical regularities standpoint.

Moreover, the Accent Familiarity Questionnaire addresses a major gap in the literature, as no such tool existed before this project to the best of our knowledge. Very few, if no, studies have focussed on quantifying familiarity with minority languages. Poretta's (2016) work quantifies familiarity with Mandarin alone, so this questionnaire surpasses previous contributions by having participants quantify familiarity in detail with Mandarin, Punjabi, Tagalog, and other languages. Furthermore, this research shows that when analyzing English bilingualism, it is paramount to delineate whether participants are L1 English bilinguals or L2 English bilinguals. This distinction is not always made in research and there are significant implications and differences between groups that may otherwise be overlooked. Together, the significance of the tasks designed for this study is not only in filling a gap in the field, but also in the focus on languages other than English. In general, there is a lack of linguistically representative testing materials and a bias toward English in the literature, so this study helps to begin pointing the field in a more diverse and inclusive direction.

As such, there are many directions for future research stemming from the outcomes of this study, particularly as solutions to some of the limitations previously discussed. An initial undertaking would be further validation of the aSL task. This would be done by having participants complete both the novel task with accented stimuli and a previously validated aSL task with non-linguistic sounds, and comparing their performance. Alternatively, performance could be compared to another aSL task using linguistic stimuli, similar to the task originally described by Saffran et al. (1996). Furthermore, replicating or extending performance on the novel accented aSL task would allow the questions posed by this study to be explored even more deeply. For example, all participants had the highest percentage correct on the question type of 'Both.' Even if participants merely got this correct because they heard the same answer twice, this means that they did indeed generalize learning across phonemes because they did not conceptualize the triplets as being different. It would be intriguing to do further analysis to see what participants' preference was on the 'Both' question—whether they consistently chose the Albertan English option, an accented option aligned with their familiarity, or if it was deterministic.

Similarly, since it was found that a faster reaction time was associated with correct trials, perhaps a few more trials of each accent presented in the familiarization phase would have allowed participants to solidify their learning, resulting in the more overt outcome of greater accuracy during the testing phase. On the other hand, an online measure of learning would also provide valuable insight to the impact of accent familiarity and language background. This could include neural imaging examining the ease in learning through the familiarization phase, for example. How does brain activation differ while learning the statistical regularities of accented stimuli as a function of familiarity or bilingualism? While there has not been much research as of date considering the neural basis of SL, some research is already planned and underway; a continuation or follow-up to this project with online measures could provide valuable insight. Finally, when designing the task, it was originally intended that performance could be analyzed by triplet and by accent, though the number of triplets from each accent prevented this. If statistical power would allow, questions could be investigated such as whether greater individual familiarity with Tagalog, for example, leads to faster RT and greater accuracy on Tagalog target triplets. Likewise, does reaction time to Mandarin triplets differ if one is a Mandarin L1 or L2 bilingual? This aSL task and data set open the door for a multitude of further studies and detailed analysis.

Secondly, the Accent Familiarity Questionnaire would also need to be further validated and re-administered to a larger sample size. This measure can also have many aspects added to it or explored more deeply, such as examining general familiarity rather than general comfort; quantifying exposure to individual accents through the various domains; delineating exposure to accents through language studies, whether immersion or university language classes, from exposure through other methods; and examining whether participants have lived in a place where English was primarily a second or additional language. Because the LSBQ asks about many of these concepts, but from a language standpoint rather than accented speech, future endeavors could look at aligning the two questionnaires more closely.

Lastly, as this study examines the processing of accented speech and the listener's ability to generalize learning across accented phonemes, a natural next step would be examining comprehension of accented speech. Intelligibility and comprehensibility are a step beyond the learning of statistical regularities in accented speech, and they have more predominant implications for the repair of communication breakdowns. A follow-up study is planned to extend this research and examine the comprehension of accented speech. Since much has yet to be explored regarding the relationships between accent familiarity, language background, and statistical learning, the current study provides an important starting point, laying the foundation for future research into improving the communication outcomes of groups represented in Canadian society.

Conclusion

In summary, the overarching goal of this research was to explore the individual cognitive factors that enable a listener to better process accented speech, by examining the relationships between statistical learning ability, accent familiarity, and language background. A novel auditory statistical learning task with accented varieties of English was utilized, alongside a novel Accent Familiarity Questionnaire. The results indicated that greater accent familiarity facilitated reaction time and processing of the accented stimuli, particularly for those who spoke English as a second or additional language. In comparing English L1 and English L2 participants, those with English as first language had a faster reaction time overall, while those who had another language as a first language were significantly more correct overall. English L1s' performance was more affected by the composition of accents in the target triplets and the question types, while the English L2s' performance was facilitated by their accent familiarity and code-switching, or use of another language socially. It is proposed that L2s were more accurate because of their ability to shift between phonological representations and their cognitive flexibility, though this may have resulted in decreased processing speed compared to the L1s. English L2s also seem to be more sensitive to changes in accent familiarity levels and SL ability than English L1s. Overall, the variables of interest had more impact on reaction time than overall score on the aSL task, pointing to the need for an online measure of learning. The results speak to the complexities in defining accent familiarity and bilingualism, and have implications for questions such as the necessary amount and type of exposure to an accent and the relationship between statistical learning and bilingualism. As for improving the communication outcomes between speakers of various language backgrounds in Canadian society, there is much more to explore. Through contributing two novel tasks by which to measure accent familiarity and auditory statistical learning with accented speech, this research is a significant step forward.

References

Abla, D., & Okanoya, K. (2008). Statistical segmentation of tone sequences activates the left inferior frontal cortex: A near-infrared spectroscopy study. *Neuropsychologia, 46*(11), 2787–95. DOI: 10.1016/j.neuropsychologia.2008.05.012.

Adank, P., Evans, B., Stuart-Smith, J., & Scott, S. (2009). Comprehension of familiar and adankunfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520–529. DOI: 10.1037/a0013552.

Adank, P., & McQueen, J. M. (2007). The effect of an unfamiliar regional accent on spoken-word comprehension. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (ICPhS 2007) (pp. 1925-1928). Dudweiler: Pirrot.

Adesope, O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207–245. DOI: 10.3102/0034654310368803

Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience, 9*, 1558–1564. DOI: 10.1038/nn1800

Ahufinger, N., Ferinu, L., Sanz-Torrent, M., Andreu, L., & Evans, J. (2021). Statistical word learning in Catalan-Spanish and English-speaking children with and without developmental language disorder. *International Journal of Language and Communication Disorders, 57*(1), 42–64. DOI: 10.1111/1460-6984.12673

Allen, J., & Miller, J. (1999). Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *Journal of the Acoustical Society of America, 106*(4), 2031–2039. DOI: 10.1121/1.427949

Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior research methods, 50*(1), 250–263. DOI: 10.3758/s13428-017-0867-9

Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2012). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism, 15*(1), 190–201. https://doi.org/10.1017/S1366728911000125

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive science*, *36*(2), 286–304. doi:10.1111/j.1551-6709.2011.01200.x

Baese-Berk, M., Bradlow, A., & Wright, B. (2013). Accent-independent adaptation to foreign accented speech. *Journal of the Acoustical Society of America, 133*(3), EL 174–80. http://dx.doi.org/10.1121/1.4789864

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.

Baugh, J. (2003). Linguistic Profiling. In S. Makoni, G. Smitherman, A. F. Ball and A. K. Spears (eds.) *Black Linguistics: Language, Society, and Politics in Africa and the Americas*, New York: Routledge, pp. 155–168.

Benitez, V., Bulgarelli, F., Byers-Heinlein, K., Saffran, J., & Weiss, D., (2020). Statistical learning of multiple speech streams: A challenge for monolingual infants. *Developmental Science, 23*(e12896), 1–12. https://doi.org/10.1111/desc.12896

Bent, T., & Atagi, E. (2017). Perception of nonnative-accented sentences by 5- to 8-year-olds and adults: The role of phonological processing skills. *Language and Speech, 60*(1), 110–122. https://doi.org/10.1177/0023830916645374

Bialystok, E., Craik, F., Green, D., & Gollan, T. (2009). Bilingual Minds. *Psychological Science in the Public Interest, 10*(3), 89–129. DOI: 10.1177/1529100610387084

Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14. http://www.praat.org/

Bogaerts, L., Siegelman, N., & Frost, R. (2020). Statistical learning and language impairments: Toward more precise theoretical accounts. *Perspectives on Psychological Science, 16*(2), 319–337. DOI: 10.1177/1745691620953082

Bradlow, A., and Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707–729. DOI: 10.1016/j.cognition.2007.04.005.

Bulf, H., Johnson, S., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*(1), 127–132. https://doi.org/10.1016/j.cognition.2011.06.010.

Chen, Y., Li L., Wang, M., & Wang, R. (2022). Which Cognitive Factors Predict L2 Grammar Learning: Cognitive Control, Statistical Learning, Working Memory, or Attention? *Frontiers in Psychology, 13.* DOI: 10.3389/fpsyg.2022.943988

Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition*, *163*, 103–120. DOI: 10.1016/j.cognition.2017.02.015

Christia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology, 3,* 479. https://doi.org/10.3389/fpsyg.2012.00479

Christiansen, M. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 24–39. DOI: 10.1037/0278-7393.31.1.24

Conrey, B., Potts, G., and Niedzielski, N. (2005). Effects of dialect on merger perception: ERP and behavioral correlates. *Brain and Language , 95*(3), 435– 49. DOI: 10.1016/j.bandl.2005.06.008

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356–371. DOI: 10.1016/j.cognition.2009.10.009

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1119–1130. DOI: 10.1037/0278-7393.30.5.1119

Derwing, T. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59*(4)*,* 547–566. https://doi.org/10.3138/cmlr.59.4.547

Derwing, T., & Munro, M. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1–16. DOI:10.1017/S0272263197001010

Dewaele, J. (2018). Why the dichotomy 'L1 versus LX user' is better than 'Native versus FranciNon-native speaker.' *Applied Linguistics; Oxford University Press, 39*(2), 236–240. DOI: 10.1093/applin/amw055

Dollmann, J., Kogan, I., & Weißmann, M. (2020). Speaking accent-free in L2 beyond the critical period: The compensatory role of individual abilities and opportunity structures, *Applied Linguistics*, *41*(5), 787–809. https://doi.org/10.1093/applin/amz029

Floccia, C., Butler, J., Goslin, J. et al. (2009). Regional and foreign accent processing in English: Can listeners adapt?. *Journal of Psycholinguist Research, 38*, 379–412. DOI: 10.1007/s10936-008-9097-8

Frost, R., Armstrong, B., & Christiansen, M. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin, 145*(12), 1128–1153. http://dx.doi.org/10.1037/bul0000210

Frost, R., Armstrong, B., Siegelman, N., & Christiansen, M. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Science, 19*(3), 117–125. http://dx.doi.org/10.1016/j.tics.2014.12.010

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language?. *Psychological science*, *24*(7), 1243–1252. https://doi.org/10.1177/0956797612472207

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*(1), 120–133. DOI: 10.1002/ejsp.862

Gao, Z., & Weinberger, S. (2018). Which phonetic features should pronunciation instructions focus on? An evaluation on the accentedness of segmental/syllable errors in L2 speech. *Research in Language*, *16*(2), 135–154. https://doi.org/10.2478/rela-2018-0012

Gass, S. M., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*(1), 65–89. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review, 16*, 486–490. http://dx.doi.org/10.3758/PBR .16.3.486

Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language, 32*, 249–268. DOI: 10.1017/S0305000904006786

Glendall, P., & Hoek, J. (1990). A question of wording. *Marketing Bulletin, 1*(5), 25-36.

Gluszek, A., & Dovidio, J. F. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology, 29*(2), 224–234. DOI: 10.1177/0261927X09359590

Grant, L., & Brinton, D. (2014). *Pronunciation myths: Applying second language research to classroom teaching*. Ann Arbor: University of Michigan Press.

Hansen, K., & Dovidio, J. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology, 22*(4), 544–551. doi:10.1037/cdp0000101.

Hanulíková, A., & Weber, A. (2012). Sink positive: Linguistic experience with *th* substitutions influences nonnative word recognition. *Attention, Perception, and Psychophysics 74*, 613–629. https://doi.org/10.3758/s13414-011-0259-7

Harris, P., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. (2009). Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-81. DOI: 10.1016/j.jbi.2008.08.010

Hixon, T., Weismer, G., & Hoit, J. (2020). *Preclinical Speech Science: Anatomy, Physiology, Acoustics, and Perception* (3rd ed., pp. 467-474). Plural Publishing.

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130*(4), 658–680. https://doi.org/10.1037/0096-3445.130.4.658

Huyck, J., Smith, R., Hawkins, S., & Johnsrude, I. (2017) Generalization of perceptual learning of degraded speech across talkers. *Journal of Speech, Language, and Hearing Research, 60,* 3334–3341. https://doi.org/10.1044/2017_JSLHR-H-16-0300

Idemaru, K. & Holt, L. L. (2011). Word recognition reflects dimension-based statistical learning. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1939-1956. DOI: 10.1037/a0025641.

Idemaru, K. & Holt, L. L. (2020). Generalization of dimension-based statistical learning. *Attention, Perception, Psychophysics 82*, 1744–1762. DOI: 10.3758/s13414-019-01956-5

Jenkins, J. (2002). A sociolinguistically-based, empirically-researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*(1), 83–103. DOI: 10.1093/applin/23.1.83

Kałamała, P. , Chuderski, A. , Szewczyk, J. , Senderecka, M. & Wodniecka, Z. (2023). Bilingualism Caught in a Net. *Journal of Experimental Psychology: General, 152*(1), 157-174. doi: 10.1037/xge0001263

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly*, *9*, 249–269. https://doi.org/10.1080/15434303.2011.642631

Kang, O., & Yaw, K. (2021). Social judgment of L2 accented speech stereotyping and its influential factors. *Journal of Multilingual and Multicultural Development*, 1–16. doi:10.1080/01434632.2021.1931247

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *The Canadian Modern Language Review, 64*(3), 459–489. DOI:10.3138/cmlr.64.3.459

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. *Cognition, 83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148–203. https://doi.org/10.1037/a0038695

Koelsch, S., Busch, T., Jentschke, S., & Rohrmeier, M. (2016). Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Scientific Reports, 6*(19741), 1–11. DOI: 10.1038/srep19741

Krosnick, J., & Presser, S. (2009). Question and questionnaire design. In J. D. Wright & P. V. Marsden (Eds)., *Handbook of Survey Research* (2nd ed., pp. 263-313). Emerald Group Publishing.

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical learning in specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3474–3486. https://doi.org/10.3389/fpsyg.2016.01245

Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of child language*, *45*(2), 368–391. DOI: doi:10.1017/S0305000917000253

Lehet, M., & Holt, L. (2020). Nevertheless, it persists: Dimension-based statistical learning and normalization of speech impact different levels of perceptual processing. *Cognition: International Journal of Cognitive Science, 202*, 1–14. DOI: 10.1016/j.cognition.2020.104328

Levy, H., Konieczny, L., & Hanulíková, A. (2019). Processing of unfamiliar accents in monolingual and bilingual children: Effects of type and amount of accent experience. *Journal of Child Language, 46*(2), 368–392. DOI: 10.1017/S030500091800051X

Lippi-Green, R. (2011). *English with an Accent: Language, Ideology and Discrimination in the United States*. (pp. 55–65, 149–181, 248–251, 322–331). Taylor & Francis Group. https://ebookcentral.proquest.com/lib/ualberta/detail.action?docID=958316

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, *36,* 173–190. https://doi.org/10.2307/3588329ing

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*, 122–134. DOI: 10.1111/j.1467-7687.2007.00653.x

Misyak, J., & Christiansen, M. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62*(1), 302–331. DOI: 10.1111/j.1467-9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in psychology*, *1*, 31. https://doi.org/10.3389/fpsyg.2010.00031

Moran, S., & McCloy, D. (2019). *PHOIBLE 2.0.* Jena: Max Planck Institute for the Science of Human History. http://phoible.org, DOI: 10.5281/zenodo.2593234

Moyer, A. (2013). *Foreign Accent: The Phenomenon of Non-native Speech.* (pp. 9–20, 85–124). Cambridge University Press. http://ebookcentral.proquest.com/lib/ualberta/detail.action?docID=1139568.

Munro, M., & Derwing, T. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of 2nd-language learners. *Language Learning, 45*(1), 73–97. DOI: 10.1111/j.1467-1770.1995.tb00963.x

Munro, M., & Derwing, T. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289–306. https://doi-org.login.ezproxy.library.ualberta.ca/10.1177/002383099503800305

Nihalani, P. (2010). Globalization and international intelligibility. In M. Saxena and T. Omoniyi (eds.), *Contending with globalization in world Englishes* (23–44). Bristol, UK: Multilingual Matters.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development, 76*, 324–339. https://doi.org/10.1111/j.1467-8624.2005.00848_a.x

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130,* 401–426. https://doi.org/10.1037/0096-3445.130.3.401

Poljak, L. (2019). Trends in accentedness and comprehensibility research, with respect of L2 speech ratings: A literature review. *Simon Fraser University Educational Review, 12*(2), 36–63. https://doi.org/10.21810/sfuer.v12i2.929

Porretta, V., Tucker, B., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics, 58*, 1–21 https://doi.org/10.1016/j.wocn.2016.05.006

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science, 21*(4), e12593. https://doi.org/10.1111/desc.12593

Riney, T., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly , 39*(3), 441– 66. DOI: 10.2307/3588489

Roark, C., & Holt, L. (2022). Long-term priors constrain category learning in the context of short-term statistical regularities. *Psychonomic Bulletin and Review, 141,* 1–13 https://doi.org/10.3758/s13423-022-02114-z

Roman, P., & Gomez-Gomez, I. (2022). Changes in Native Sentence Processing Related to Bilingualism: A Systematic Review and Meta-Analysis. *Frontiers in Psychology, 13*, doi: 10.3389/fpsyg.2022.757023

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley interdisciplinary reviews. Cognitive science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R. (2018). Statistical learning as a window into developmental disabilities. *Journal of neurodevelopmental disorders*, *10*(1), 1–5. https://doi.org/10.1186/s11689-018-9252-y

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. DOI: 10.1126/science.274.5294.1926

Saffran, J., Johnson, E., Aslin, R., & Newport, E. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1)*,* 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J., Newport, E., Aslin, R., Tunick, R., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105. https://doi.org/10.1111/j.1467-9280.1997.tb00690.x

Saffran, J., & Wilson, D. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy, 4*, 273–284. https://doi.org/10.1207/S15327078IN0402_07

Saito, K., & Shintani, N. (2015). Do native speakers of North American and Singapore English differentially perceive comprehensibility in second language speech? *TESOL Quarterly, 50*(2). 421–446 DOI:10.1002/tesq.234

Schairer, K. (1992). Native speaker reaction to non-native speech. *Modern Language Journal, 76*(3), 309–319. DOI: 10.2307/330161

Setter, J. (2006). Speech rhythm in world Englishes: The case of Hong Kong. *TESOL Quarterly, 40*(4), 763–782. DOI: 10.2307/40264307

Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, *17*(3), 247–271. DOI: 10.1111/j.1532-7078.2011.00085.x

Shintani, N., Saito, K., & Koizumi, R. (2019). The relationship between multilingual raters' language background and their perceptions of accentedness and comprehensibility of second language speech. *International Journal of Bilingual Education and Bilingualism 22*(7), 849–869. DOI: 10.1080/13670050.2017.1320967

Siegelman, N. (2020). Statistical learning abilities and their relation to language. *Language and Linguistics Compass, 14*(3), e12365. DOI: 10.1111/lnc3.12365

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining "learning" in statistical learning: What does an online measure reveal about the assimilation of visual regularities? *Cognitive Science, 42*(S3), 692–727. https://doi.org/10.1111/cogs.12556

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition, 177,* 198-213. https://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105–120. http://dx.doi.org/10.1016/j.jml.2015.02.001.

Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental science*, *15*(4), 482–495. DOI: 10.1111/j.1467-7687.2012.01141.x

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy related skills. *Reading and Writing, 28(*4), 467–490. https://doi.org/10.1007/s11145-014-9533-0

Statistics Canada. (2017). *Edmonton, CY [Census subdivision], Alberta and Division No. 11, CDR [Census division], Alberta* (table). *Census Profile*. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Released November 29, 2017. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

Stilp, C., & Lewicki, M. (2013). Statistical structure of speech sound classes is congruent with cochlear nucleus response properties. *Proceedings of Meetings on Acoustics, 20*(1), https://doi.org/10.1121/1.4865250

Sumner, M., & Samuel, A. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*, 487–501. https://doi.org/10.1016/j.jml.2009.01.001

Tao, L., Wang., G., Zhu, M., & Cai, Q. (2021). Bilingualism and domain-general cognitive functions from a neural perspective: A systematic review. *Neuroscience and Biobehavioral Reviews, 125,* 264-295. https://doi.org/10.1016/j.neubiorev.2021.02.029

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*(4), 792–814. https://doi.org/10.1037/a0030801

Toscano, J., & McMurray, B. (2010). Cue integration with categories: Weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cognitive Science, 34*(3), 434–464. https://doi.org/10.1111/j.1551-6709.2009.01077.x

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98*(3), 642–652. https://doi.org/10.1037/0022-0663.98.3.642

Vandermosten, M., Wouters, J., Ghesquiere, P., & Golestani, N. (2019). Statistical learning of speech sounds in dyslexic and typical reading children. *Scientific Studies of Reading, 23*(1), 116–127. DOI: 10.1080/10888438.2018.1473404

Vasuki, P., Sharma, M., Demuth, K., & Arciuli, J. (2016). Musicians' edge: A comparison of auditory processing, cognitive abilities, and statistical learning. *Hearing Research, 342,* 112–123. https://doi.org/10.1016/j.heares.2016.10.008

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, *38*(2), 133-149.

Wagner, A., Ernestus, M., & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *The Journal of the Acoustical Society of America, 120*(4), 2267–2277. https://doi.org/10.1121/1.2335422

Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in Psychology, 5*(953), 1–9. https://doi.org/10.3389/fpsyg.2014.00953

Warker, J., & Dell, G. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 387–398. DOI: 10.1037/0278-7393.32.2.387

Weinberger, S. (2015). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu

Winke, P., & Gass, S. (2013). The influence of second language experience and accent familiarity on oral proficiency rating: A qualitative investigation. *TESOL Quarterly 47*(4), 762–789. DOI:10.1002/tesq.73.

Wu, D., Y Shih, E., Frost, R., Lee, J.-R., Lee, C.-Y., Tsai, J.-L., . . . Tzeng, O. (2012). *Statistical learning of implicit regularity correlates with acquisition of Chinese literacy as a second language.* Paper presented at the annual meeting of the Psychonomic Society, Minneapolis, MN.

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. https://doi.org/10.1126/sciadv.1603309

Appendix A

Speakers from Speech Accent Archive

	Age	Sex	Birthplace	Age of English Acquisition	Length of English Residence	Phonetic Transcription
Albertan English speaker	25	M	Calgary, Alberta, Canada	0	25 years	[pʰliːz̥ kʰɑlˠ stɛlə æsk hɚ ɾə bɹĩŋ niz θĩŋz wɪθ hɚ fɹʌ̃m ðə stɔɹ sɪks spũːnz əv fɹɛʃ ʃnoʊ pʰiːz faɪv θɪk slæbz̥ əv bluː ʧiːz ɛ̃n mebi ə snæːk fɚ hɝ bɹʌðɚ bɑːb wi ɑlˠso niːd ə smɑlˠ pʰlæstɪk sneɪk ɛ̃n ə bɪg̚ tʰɔɪ fɹɑg fə ðə kʰɪdz ʃi kʰə̃n skup̯ ðis θɪŋz̥ ĩntə θɹi ɹɛd bæːgz̥ ɛ̃n wi wɪlˠ goʊ mit hɝ wɛ̃nzdeɪ æt ðə tʰɹeĩn steɪʃə̃n]
Mandarin-accented English speaker	29	M	Jingmen, Hubei, China	12	5 years	Phonetic transcription available upon request.
Punjabi-accented English speaker	26	M	Hoshiarpur, India	10	.6 years	Phonetic transcription available upon request.
Tagalog-accented English speaker	60	M	Manila, Philippines	6	22 years	Phonetic transcription available upon request.

Table A1. Demographics of speakers from Speech Accent Archive (Weinberger, 2015).

Appendix B

Stimuli Used in Auditory Statistical Learning Task

The standard elicitation paragraph containing most of the consonants, consonant clusters, and vowels in English is provided below. The phonemes chosen for the auditory statistical learning task are thus based on the phonemes available in this paragraph, with careful consideration in regard to linguistic context and relationship to the L1 phonetic inventory of the languages represented.

> "Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station."

Figure B1. Standard elicitation paragraph for natural speech stimuli. The sixteen phonemes used are highlighted.

IPA	Orthography	Rationale
/kɪ/	ki	This syllable was chosen over "call" and "can" to eliminate the effect of velarization and nasalization on the vowel.
/tɔɪ/	toi	The duration of this syllable is longer than the alternative of "to," and the diphthong may provide an additional indication of accentedness.
/ði/	thee	This syllable was chosen over "the" to have a tense vowel rather than a lax schwa.
/θɪ/	thi	This syllable was chosen over "thing" to eliminate the effect of nasalization on the vowel.
/wi/	wee	This syllable was chosen over "will" and "Wen" to eliminate the effect of nasalization and velarization on the vowel. The duration will

		be longer than extracting /wɪ/ from the word "with."
/sɪ/	si	This syllable was chosen over the second syllable in "also" to contain a tense vowel rather than vowel reduction.
/faɪ/	fai	This syllable was chosen over "for" to eliminate the effect of rhoticization on the vowel.
/tʃi/	chee	This was the only exemplar of this phoneme.
/ni/	nee	This was the only exemplar of this phoneme.
/bɑ/	baw	This syllable was chosen over "big" and "bag" for the tense vowel and to avoid the variation in velar-raising found in "bag."
/pi/	pee	Even though this phoneme is present in all phonetic inventories, it was chosen over /lə/ because the word it was in is stressed, while "la" was unstressed, resulting in short duration and low pitch. This was the only exemplar of the phoneme /p/.
/ʃi/	shee	This syllable was chosen over the second syllable in "station" for its tense vowel and longer duration and to eliminate the effect of nasalization.
/deɪ/	day	This was the only exemplar of this phoneme, and the diphthong may provide an additional indication of accentedness.
/ɹɛ/	re	This was the only exemplar of this phoneme.
/goʊ/	go	This was the only exemplar of this phoneme, and the diphthong may provide an additional indication of accentedness.
/mi/	mee	This syllable was chosen over the first syllable in "maybe" for its tense vowel, as "maybe" tends to be reduced in conversational speech.

Table B1. Syllables used in the auditory statistical learning task.

Triplet number	Stimulus	Rationale
1	mee-fai-we	The phonemes /m/, /f/, /w/, and /p/ are found in all four native phonetic inventories. These would be considered "easy" phonemes to constitute the "hard" triplets. The hard triplets are denoted by red font, while the easy triplets below are denoted by green font.
2	fai-mee-pee	
3	pee-we-mee	
4	we-pee-fai	
5	day-chee-thi	The final phonemes in each of these triplets are only present in Albertan English. These would be considered "hard" phonemes, due to their absence in the other phonetic inventories and likely accentedness. The final portion of an input sequence is learned best through aSL, so it is anticipated that by being at the end of the triplet, these phonemes will be a more telling measure of statistical learning (Christiansen, 2005). The remaining phonemes are present in all inventories except Mandarin or have considerable allophonic variation from the congruent Albertan English phonemes, making them "medium" difficulty phonemes.
6	toi-ba-thee	
7	ki-go-shee	
8	nee-si-re	

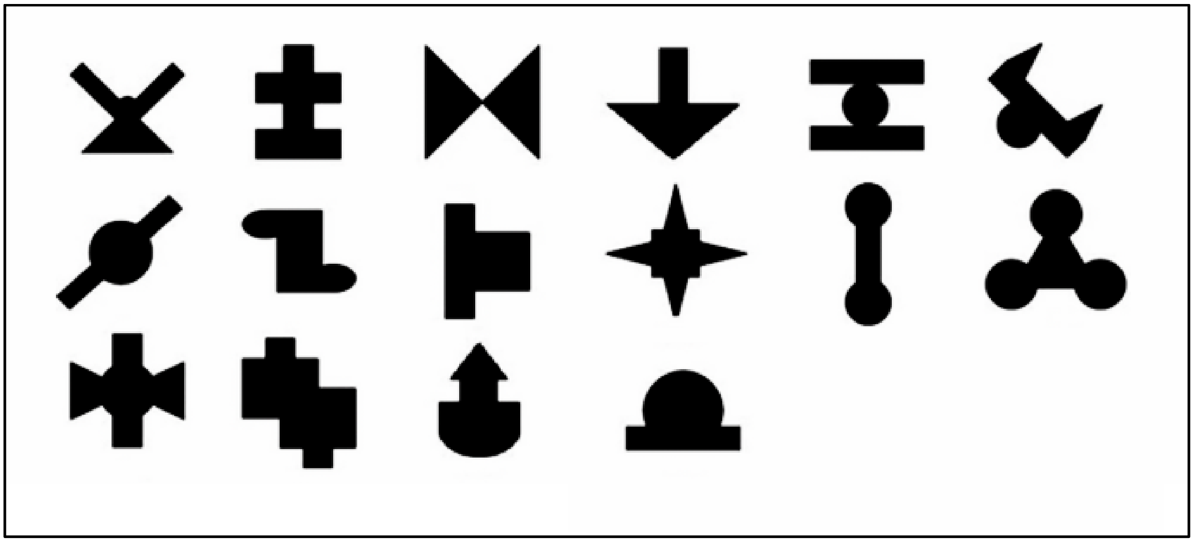
Table B2. These are the triplets used in the auditory statistical learning task.

Figure B2 shows the structure of accented and unaccented triplets. Albertan English is denoted by *a,* Mandarin-accented English by *m,* Punjabi-accented English by *p,* and Tagalog-accented English by *t.* There are eight total triplets. Each of the eight triplets repeats for a total of four times (e.g., "mee-fai-we" will appear in up to four different accents). The four "hard" triplets are denoted in red font, where four phonemes are permuted. Four "easy" triplets are denoted in green font, where phonemes appear in a fixed order only.

1 aaa	1.1 aaa	1.2 aaa	1.3 aaa
2 aaa	2.1 aaa	2.2 aaa	2.3 aaa
3 aaa	3.1 mmm	3.2 ppp	3.3 ttt
4 aaa	4.1 mmm	4.2 ppp	4.3 ttt
5 aaa	5.1 aaa	5.2 aaa	5.3 aaa
6 aaa	6.1 aaa	6.2 aaa	6.3 aaa
7 aaa	7.1 mmm	7.2 ppp	7.3 ttt
8 aaa	8.1 mmm	8.2 ppp	8.3 ttt

Figure B2. Pattern of accents in the eight triplets used in the auditory statistical learning task. For triplet 1, the repetition is denoted by 1.1, 1.2, and 1.3. The same phonemes are used but the accent may or may not change.

Appendix C



Appendix C. Shapes used in the visual statistical learning test, based on Siegelman, Bogaerts, Kronenfeld et al., (2018).

Supplementary Materials

Table S1

Correlations Between aSL Score, vSL Score, LSBQ, and General Accent Familiarity Rating

	ASL	VSL	LSBQ_Total	GeneralAcc
ASL	1	0.36	0.17	0.12
VSL	0.36	1	0.08	0.03
LSBQ_Total	0.16	0.08	1	0.11
GeneralAcc	0.12	0.03	0.11	1

Table S2

Regression Model for aSL Score as Explained by vSL Score

	Estimate	Standard Error	*t*	*p*
(Intercept)	17.874	2.565	6.967	< .001 (***)
VSL	0.227	0.091	2.508	0.016 (*)

Note. The model residual standard error is 3.806 on 41 degrees of freedom, $R^2 = .1118$, $p = .016$.

Table S3

Regression Model for aSL Reaction Time as Explained by Correct Trials

	Estimate	Standard Error	*t*	*p*
(Intercept)	1262.21	38.44	32.835	< .001 (***)
ASL_Correct	-139.83	50.60	-2.763	0.00579 (**)

Note. The model residual standard error is 1031 on 1698 degrees of freedom, $R^2 = .00389$, $p < .005786$.

Table S4

Regression Model for General Accent Familiarity Predicting Reaction Time on 'Hetero Acc' Question Type

	Estimate	Standard Error	*t*	*p*
(Intercept)	1767.24	291.34	6.066	<.001 (***)
GeneralAcc	-87.69	37.50	-2.338	0.0201 (*)

Note. The model residual standard error is 1012 on 282 degrees of freedom, R^2 = .016, p = .02.

Table S5

Interaction Between aSL score and English L1 Status on Reaction Time

	Estimate	Standard Error	*t*	*p*
(Intercept)	2485.525	231.414	10.741	<.001 (***)
ASL	-47.168	8.855	-5.236	<.001 (***)
EnglishL1_Y	-837.957	313.958	-2.669	0.00768 (**)
ASL: EnglishL1_Y	25.379	12.590	2.016	0.04398 (*)

Note. The model residual standard error is 1021 on 1696 degrees of freedom, R^2 = *.022*, $F(3, 1696) = 13.72$, p = < .001.

Table S6

Interaction Between Correct versus Incorrect Trials and English L1 Status on Reaction Time

	Estimate	Standard Error	*t*	*p*
(Intercept)	1397.02	70.83	19.723	<.001 (***)
CorrectTRUE	-197.15	89.92	-2.192	0.0285 (*)

EnglishL1_Y	-190.80	84.27	-2.264	0.0237 (*)
CorrectTRUE: EnglishL1_Y	71.28	108.78	0.655	0.5124

Note. The model residual standard error is 1029 on 1696 degrees of freedom, R^2 = *.007,* $F(3, 1696) = 5.269$, p = .00128

Table S7

All Factors Predicting Reaction Time on aSL Task for English L1s

	Estimate	Standard Error	*t*	*p*	
(Intercept)	1287.9483	158.7282	8.114	< .001	***
ASL_Correct	-122.8711	64.1604	-1.915	0.0557	.
ASL_QType3 - 4AFC	147.0725	73.2134	2.009	0.0448	*
ASL_QType3 - blank	112.6058	75.7396	1.487	0.1374	
ASL_LegendEasy AB	-147.5743	102.0520	-1.446	0.1484	
ASL_LegendHetero Acc	-143.1182	97.6454	-1.466	0.1430	
ASL_LegendNormal	-60.6354	76.2948	-0.795	0.4269	
NonEng_home	-3.8318	7.3634	-0.520	0.6029	
NonEng_social	0.8983	4.6971	0.191	0.8484	
GeneralAcc	-17.3653	18.0385	-0.963	0.3359	
FreqInteractions_Total	4.6093	12.7645	0.361	0.7181	

—

Note. R^2 = .004, p = .1207, model residual standard error is 975.1 on 1133 DF.

Table S8

All Factors Predicting Reaction Time on Correct Trials for English L1 Participants

	Estimate	Standard Error	*t*	*p*	
(Intercept)	1216.078	182.903	6.649	< .001	***

ASL_QType3 - 4AFC	297.116	97.320	3.053	0.00236	**
ASL_QType3 - blank	243.868	90.963	2.681	0.00753	**
ASL_LegendEasy AB	-26.363	127.583	-0.207	0.83636	
ASL_LegendHetero Acc	-93.563	117.477	-0.796	0.42608	
ASL_LegendNormal	-87.806	85.030	-1.033	0.30217	
NonEng_home	-11.982	9.271	-1.292	0.19670	
NonEng_social	2.915	5.993	0.486	0.62689	
GeneralAcc	-37.195	24.005	-1.549	0.12178	
FreqInteractions_Total	5.013	16.430	0.305	0.76038	

—

Note. R^2 = .0164, p = .022, model residual standard error is 916.9 on 626 DF.

Table S9

All Factors Predicting Reaction Time on Incorrect Trials for English L1 Participants

	Estimate	Standard Error	t	p	
(Intercept)	1367.7857	276.7111	4.943	< .001	***
ASL_QType3 - 4AFC	-28.3254	113.7191	-0.249	0.803	
ASL_QType3 - blank	-120.1976	131.5546	-0.914	0.361	
ASL_LegendEasy AB	-304.2119	195.8696	-1.553	0.121	
ASL_LegendHetero Acc	-245.5776	193.1611	-1.271	0.204	
ASL_LegendNormal	-63.1047	166.0425	-0.380	0.704	
NonEng_home	6.6852	11.8700	0.563	0.574	
NonEng_social	0.1855	7.4564	0.025	0.980	
GeneralAcc	1.2682	27.6036	0.046	0.963	
FreqInteractions_Total	3.5039	20.0100	0.175	0.861	

—

Note. R^2 = −0.008621, p = .8615, model residual standard error is 1041 on 498 DF.

Table S10

All Factors Predicting Reaction Time on aSL Task for English L2s

	Estimate	Standard Error	*t*	*p*	
(Intercept)	2632.83	399.12	6.597	< .001	***
ASL_Correct	-213.43	103.86	-2.055	0.040358	*
ASL_QType3 - 4AFC	135.13	116.97	1.155	0.248464	
ASL_QType3 - blank	158.82	122.71	1.294	0.196125	
ASL_LegendEasy AB	-309.46	160.94	-1.923	0.055019	.
ASL_LegendHetero Acc	-200.90	157.61	-1.275	0.202985	
ASL_LegendNormal	-53.26	122.86	-0.434	0.664812	
NonEng_home	9.31	23.04	0.404	0.686270	
NonEng_social	-28.49	11.39	-2.502	0.012640	*
GeneralAcc	-216.69	56.54	-3.833	0.000142	***
FreqInteractions_Total	81.15	28.02	2.896	0.003926	**

—

Note. R^2 = .0496, p = < .001, model residual standard error is 1105 on 545 DF.

Table S11

All Factors Predicting Reaction Time on Correct Trials for English L2 Participants

	Estimate	Standard Error	*t*	*p*	
(Intercept)	2473.542	472.932	5.230	< .001	***
ASL_QType3 - 4AFC	-1.868	147.391	-0.013	0.98990	
ASL_QType3 - blank	95.413	146.622	0.651	0.51566	
ASL_LegendEasy AB	-309.211	186.184	-1.661	0.09769	.
ASL_LegendHetero Acc	-397.957	190.584	-2.088	0.03754	*
ASL_LegendNormal	-152.372	137.720	-1.106	0.26935	

NonEng_home	-6.368	28.118	-0.226	0.82098	
NonEng_social	-29.316	13.389	-2.189	0.02925	*
GeneralAcc	-221.785	72.077	-3.077	0.00226	**
FreqInteractions_Total	111.470	34.695	3.213	0.00144	**

—

Note. R^2 = .055, p = < .001, model residual standard error is 1071 on 335 DF.

Table S12

All Factors Predicting Reaction Time on Incorrect Trials for English L2 Participants

	Estimate	Standard Error	*t*	*p*	
(Intercept)	2270.400	697.515	3.255	0.00133	**
ASL_QType3 - 4AFC	397.224	195.795	2.029	0.04380	*
ASL_QType3 - blank	458.784	226.418	2.026	0.04406	*
ASL_LegendEasy AB	9.033	348.343	0.026	0.97934	
ASL_LegendHetero Acc	385.740	333.275	1.157	0.24847	
ASL_LegendNormal	405.294	285.286	1.421	0.15697	
NonEng_home	32.180	40.496	0.795	0.42776	
NonEng_social	-30.439	21.607	-1.409	0.16045	
GeneralAcc	-218.564	92.180	-2.371	0.01868	*
FreqInteractions_Total	29.731	48.029	0.619	0.53661	

—

Note. R^2 = .0429, p = .0359, model residual standard error is 1151 on 201 DF.

Table S13

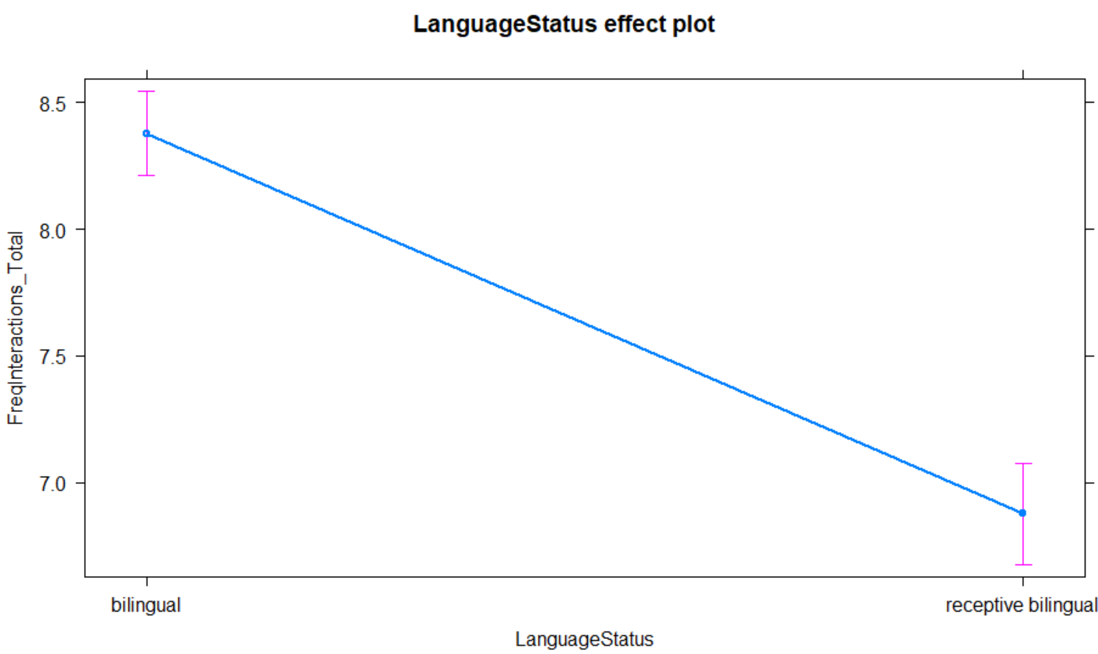
Interaction Between General Accent Familiarity and English L1 Status on Reaction Time

	Estimate	Standard Error	*t*	*p*
(Intercept)	2067.81	306.00	6.758	<.001 (***)
GeneralAcc	-98.28	37.53	-2.619	0.00891 (**)
EnglishL1_Y	-823.55	333.76	-2.467	0.01370 (*)
GeneralAcc: EnglishL1_Y	83.62	41.45	2.017	0.04383 (*)

Note. The model residual standard error is 1029 on 1696 degrees of freedom, R^2 = .007, $F(3, 1696)$ = 4.773, p = .0026.

Figure S1

Language Status Predicting Frequency of Interaction with Accented Speech



Note. The y-axis shows frequency of interaction with accented speech, measured as self-report on a scale of 1 of 10.