The Pennsylvania State University

The Graduate School

Department of Psychology

TRACKING MULTIPLE INPUTS: AN INVESTIGATION OF THE PRIMACY EFFECT IN MONOLINGUALS AND BILINGUALS

A Thesis in

Psychology

by

Federica Bulgarelli

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Psychology

Master of Science

August 2015

The thesis of Federica Bulgarelli was reviewed and approved* by the following:

Daniel J. Weiss
Associate Professor of Psychology and Linguistics
Thesis Advisor

Judith F. Kroll
Distinguished Professor of Psychology, Linguistics and Women's Studies

Paola E. Dussias
Professor of Spanish, Italian and Portuguese, Linguistics and Psychoogy

Melvin M. Mark
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

ABSTRACT

A fundamental challenge of statistical learning is to determine whether variance observed in the input signals a change in the underlying structure. When asked to segment two consecutive artificial languages, learners show a reliable tendency to only learn the first of two presented languages unless the change is correlated with a contextual cue or exposure to the second structure is lengthened (Gebhart, Aslin, & Newport, 2009). In this thesis, I explored whether the primacy effect can be attenuated by manipulating the amount of exposure to the first language. I presented participants with five one-minute blocks of two artificial languages, each followed by a test. In one condition, learners received fixed input, whereas in the other they advanced to the second language immediately after learning the first. When participants advanced to the second language as soon as learning occurred, the primacy effect was attenuated. Notably, contextual cues did not boost performance using this latter paradigm, suggesting that without becoming entrenched in the first language there is no additive effect for such cues. In a second set of experiments, I compared monolinguals to advanced second language learners, to determine whether experience learning and managing two linguistic systems impacts learner's ability to learn two artificial languages presented in succession. Results thus far reveal a greater percentage of second language learners acquiring the second presented language relative to the monolingual group. Overall, the findings suggest that anchoring effects may be due to additional exposure after a single structure has been learned, possibly as a function of learners reducing their sampling rate once mastering a set of statistics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Dan Weiss, for help on this proposal; my committee members, Dr. Dussias and Dr. Kroll for agreeing to serve on my committee and providing feedback that has and will continue to help me improve this line of work; I would also like to thank my lab mates, Dr. Lisa Heimbauer, Dr. Nate George, and Tim Poepsel, for being supportive lab members. Lastly, I would like to thank all of the undergraduate RAs that have helped me with this project: Cara Theoret, Maura Petrulsky, Madison Young, Lindsay Rauch, Sarah Miller and Nick Andereeg.

Chapter 1

Introduction

Statistical learning is the process by which learners track rudimentary statistics in their sensory input in order to extract the structures found in their environment (Saffran, 2003). This process can be achieved in an environment where the learner receives immediate feedback (often in the form of rewards), or in an environment where the learner must keep track of the co-occurrence of specific events without explicit information aiding them in their discovery. Over the past few decades, research in the field has demonstrated statistical learning to be a domain general learning mechanism that operates in both linguistic (e.g. Saffran, Aslin & Newport, 1996) and non-linguistic domains (Fiser & Aslin, 2002; Saffran, Johnson, Aslin, & Newport, 1999).

Statistical learning takes hold early in development. Research on visual statistical learning has provided evidence that infants as young as 3 days old can detect statistical information in their input, but that these abilities are limited at this stage and develop over the course of time and experience (Bulf, Johnson, & Valenza, 2011; Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). Infants and children have been shown to use the process of statistical learning to determine probability distributions, sequential structures, and correlations between stimulus dimensions (Lany & Saffran, 2013). This process, however, does not stop once the structure of one's environment is discovered, and adults demonstrate the same abilities when tested in many of the same contexts (Conway & Christiansen, 2006; Saffran et al., 1999).

While statistical learning operates in many areas of learning, one of the primary areas of application has been language acquisition. Statistical learning has been shown to be involved in phonetic learning (e.g., Maye, Weiss, & Aslin, 2008; Janet F. Werker & Tees, 1984), speech segmentation (e.g., Saffran, Aslin & Newport, 1996), mapping words to meaning (e.g., Smith & Yu, 2008), grammar learning (e.g., Gomez and Gerken, 2002; Reeder, Newport, & Aslin, 2013) and late second language acquisition (Frost, Siegelman, Narkiss, & Afek, 2013). The variety of language learning processes that can be informed by this learning mechanism demonstrate that statistical learning is involved in a variety of cognitive processes across development.

Statistical learning in language acquisition

For the past two decades, research on statistical learning has focused on investigating which types of underlying structures learners can extract using this method. The following section provides a review of studies that investigated the variety of language acquisition processes that can be informed by statistical learning. Critically, these studies focus on demonstrating that the distribution of statistics in an environment can serve as a cue to the underlying structure, but are not intended to imply that statistics are the only cue learners use to determine the structure of their environment. Further, this section will provide examples of the methods by which statistical learning has typically been studied; with infants, children, and adults.

Arguably, one of the first requirements for a child to acquire the language spoken in their environment is the process of phonemic discrimination; the discovery of which sounds in their native language convey meaning. Until a child gains sufficient exposure, discrimination of speech sounds seems to be universal and not impacted by the categories found in their native language (Werker, Gilbert, Humphrey, & Tees, 1981). However, as a child begins to receive linguistic input, their discrimination of speech sounds becomes tuned to the native categories. Werker and Tees (1983) found that by age 4, children discriminate some non-native speech sounds as poorly as adults. In order to determine the developmental time period of this perceptual tuning, Werker and Tees (1984) tested infants’ ability to discriminate non-native speech sounds in 6-8 month olds and 10-12 month olds, and found that most of the infants ages 6-8 months correctly discriminated non-native contrasts, while those ages 10-12 months were largely unsuccessful at discrimination. This finding demonstrates that prior to gaining experience with their native language, infants possess a relatively language-general pattern of speech discrimination, but after exposure to the native language they begin to form language-specific patterns of discrimination.

While the aforementioned research demonstrates that children's speech discrimination abilities are affected by exposure to the native language over the first year of life, it does not provide insight into the mechanism that plays a role into these developmental changes. Maye, Werker, and Gerken (2002) tested whether statistical learning could play a role in explaining the perceptual tuning phenomenon. They tested whether 6 and 8-month-old infants could discriminate the contrast between voiced and voiceless unaspirated alveolar stops [t] and [d] following exposure to an 8-point voicing continuum between the syllables [ta] and [da]. Following an intuition based on distributional patterns in language shown by Lisker and Abramson (1967), half of the infants were familiarized with a unimodal sound distribution (normally distributed around a single peak), and half with a bimodal sound distribution (a continuum consisting of two peaks). They discovered that the distribution of the continuum that the infants heard impacted their ability to discriminate the speech contrast, with the infants trained in the bimodal distribution continuing to discriminate the speech sound contrast, while the infants in the unimodal distribution showed no evidence of discrimination. This study therefore provides some evidence that statistical learning may be involved in the perceptual tuning during the first year of life (though see Werker, 2013, for an argument that statistical learning alone cannot account for the entire perceptual tuning phenomenon, as after 10 months of age social interactions appear to facilitate learning more than distributional characteristics of the input).

While distributional learning may facilitate speech sound discrimination, sequential statistical learning appears to play a role in another fundamental task confronting early language learners, namely speech segmentation. Some of the first evidence that statistical information might be involved in recognizing words in fluent speech came from a study by Jusczyk and Aslin (1995). They familiarized 7.5 month-old infants to different monosyllabic words and then presented the infants with fluent speech that either did or did not contain the familiarized words. They found that infants listened longer to the passages that contained the words heard during familiarization. This experiment demonstrates young infants' ability to recognize familiar words in fluent speech, providing evidence that they are able to recognize syllables that occurred together to make words from the initial exposure in the fluent stream of speech.

In a seminal study, Saffran, Aslin, and Newport (1996) asked whether 8-month old infants could track statistics in fluent artificial speech in order to parse the stream into words. They exposed 8-month-old infants to two minutes of an artificial speech stream that was comprised of four statistically-defined words. The stream was devoid of any cues to segmentation other than the transitional probabilities between syllables. Following familiarization, they tested the infants' ability to discriminate the statistically-defined words that occurred during familiarization from non-words that did not. They found that the infants demonstrated a novelty preference during the testing phase that could only emerge if the infants were aware of the statistical structure of the stream. The same findings were replicated using words and part-word test items consisting of syllables that spanned word boundaries. Further, the same results were shown with adults as well (Saffran et al., 1996b). This study was the first to demonstrate that infants could use the statistical information in their auditory input to determine where word boundaries are found in fluent speech. It was the study that catapulted statistical learning as a viable mechanism for extracting structure from speech input.

Subsequent studies investigated this phenomenon using natural language stimuli. For example, Pelucchi et al. (2009) exposed 8-10 month olds to fluent Italian speech containing words that had high and low transitional probabilities. The speech stream was controlled for frequency as both types of words occurred the same number of times. They found that the infants could use the statistical information to segment the fluent Italian speech in both conditions, thereby providing evidence that statistical learning of transitional probabilities can take place in more naturalistic conditions.

Another aspect of language acquisition that may be informed by statistical learning is mapping words to their meaning. A fundamental part of language acquisition is determining which sounds map onto which items in the environment. Since this varies cross-linguistically, and is fairly arbitrary, this line of research investigates how naïve learners acquire which sounds refers to which object. This is called the word-to-world mapping challenge and has thus far been studied by investigating the constraints that learners may bring to bear to alleviate the problem. One of these constraints may be mutual exclusivity, the idea that one label can only refer to one item. For example, Markman (1990) hypothesized and supported the claim that if children viewed two objects and heard two words, only one of which was known, they would immediately attribute the novel word to the novel object. Mutual exclusivity implies some prior word knowledge on behalf of the learners, but how does the learner acquire these mappings in the first place?

Researchers have investigated the role of tracking the statistics of co-occurrences between sounds and referents to solve this word-to-world mapping problem. Yu, & Smith, (2007) tested whether adults could use statistical information to map novel words to novel items by presenting adult participants with trials during which they heard novel words while looking at pictures of novel objects. During the experiment, participants heard and saw up to 4 words and objects, with no information about which words corresponded with which objects. Each individual trial was ambiguous, such that the mapping task could only be accomplished if the statistics of co-occurrences between words and objects were tracked. They found that participants were able to map the correct word to the correct item above chance regardless of the number of items found in the trials. Smith and Yu (2008) then investigated whether infants could also track the statistics in their input to correctly learn word-referent mappings. In their experiment, 12-14 month old infants were exposed to trials during which they saw two items and heard two words, similarly with no indication of which word belonged to which item. They found that infants were able to use statistical information gained across trials to form accurate word-referent mappings, demonstrating that statistical learning mechanisms also apply to the formation of word-to-world mappings.

Another challenge faced by language learners is the process of grammar learning. Reeder and Aslin (2013) tested whether adult participants could learn grammatical information in an artificial language only by using the statistical information found within the language. They presented participant with 20 minutes of exposure to an artificial language and then asked participants to judge the grammaticality of a series of strings that had either occurred in the familiarization stream, novel structures that conformed to the grammatical rules, or ungrammatical strings. They found that participants rated the familiar and novel grammatical strings as more familiar than the ungrammatical strings, demonstrating that adult learners could in fact use the distributional information from their input to acquire grammatical categories and generalize to novel words, all while allowing for exceptions in the input.

Taken together, the studies discussed in this section demonstrate the centrality for statistical learning at many levels of language acquisition. Statistical learning has also been studied using a variety of methodologies and across development, which demonstrates its ability to play a role in different stages of language acquisition. However, statistical learning is only one of many mechanisms that can inform and aid in language acquisition. The review presented here is meant to demonstrate that the statistics of the environment are one of the cues used by learners when they are presented with an environment that still needs to be deciphered.

Learning in a non-stationary environment

The aforementioned research has oversimplified the problem faced by many learners by exposing participants to a single language with a consistent set of statistics governing the structure. Critically, all of the research presented thus far has made the same assumption for the learner, that only one underlying structure needs to be learned. Additionally, these studies stripped the input of any other cue to the underlying structure, ensuring that the statistical cue they intended to study was the only one available to learners to determine the underlying structure (Aslin, 2014). However, real world environments are not comprised of uniform sets of statistics, and the input to the learners has considerable amounts of variance. Learners in the real world must be able to account for all of the variance found in their environment to determine the accurate underlying structure. Thus, one of the primary challenges of learning in a real-world environment is determining whether variance encountered in the input signals that a change has occurred in the underlying structure, or indicates a need to expand the current model to accommodate the new information (see Qian, Jaeger, & Aslin, 2012). This notion accords with Piaget's (1964) description of assimilation and accommodation. The learner is faced with the task of determining the environment's underlying structure based on the input they receive, constantly having to update the model they have created to account for the information in their environment. This task is made more complicated by the fact the underlying structure could change implicitly, not providing the learner with any overt indication that a change has occurred (see Qian, Jaeger, & Aslin, 2012). In these cases, the learner is left to discern whether the variance they are experiencing in their environment is due to an unobservable change in context, or if it is information that should be accounted for in the current model.

Research on learning and memory has previously investigated learners' ability to adapt to change in their environment. However, this line of research has typically used an explicit error signal to notify participants that a change has occurred. Behrens et al., (2007) tested whether human learners could track the statistics of their reward structure, and adjust their learning rate accordingly. Participants were shown a green and blue rectangle and were told that the reward structure (the chance of a given color being the correct answer) depended only on the recent outcome history. They showed that participants could quickly learn in a stable environment (where the probability of a certain color being correct was 75%), but were also able to adjust their choices on the task depending on the reward probability of each of the choices (where every 30-40 trials the reward structure switched completely)(Behrens, Woolrich, Walton, & Rushworth, 2007). A more recent study tested participants' reaction to uncued changes to the reward structure of a Wisconsin Card Sorting task (Wilson & Niv, 2011). During this task, participants are instructed to sort the cards correctly in exchange for points, although the sorting rule is not disclosed.

After each card is sorted, participants are told whether they successfully earned points, which allows them to induce the reward structure over time. Wilson and colleagues were interested in how participants would react when the reward structure changed. They also found that participants were able to adapt to the change in the reward structure, but that it took them a couple trials after the switch to learn the novel reward structure. The results of this experiment demonstrate that, despite being presented with implicit changes in the rules, participants can continue to learn each correct reward structure; but that providing them with an error signal increases the speed of the process. Taken together, findings from this line of research provide evidence that learners are quickly able to react to change in their environment when presented with an error signal.

However, in many real-world settings, there is not always immediate feedback available to the learner. Imagine, for example, that you work the lunch shift at a restaurant and you typically serve 5 customers per hour. One day, 25 customers arrive within an hour and your manager is faced with the decision of whether to staff additional employees the following days, or to consider this instance an outlier and maintain the current schedule. It is possible that the 25 customers within the hour were due to conference attendees taking a lunch break. In this scenario, the manager could either consider this one instance an outlier and assimilate it to the current operational model, or create a different model based on the context change. For example, any time a conference is in town, the new model would apply and the manager should staff additional employees for the lunch shift. However, it could also be due to a positive review for the restaurant in the local newspaper, which would require the manager to accommodate their current model to include the novel information. In this case, the manager would need to update their current causal model to account for this new information, as it will continue to remain relevant in future staffing decisions (see Qian, Jaeger, & Aslin, 2012 for a related example used to generate this one).

Unlike in the experiments from the learning and memory literature, the manager in the example described above is not given explicit feedback about the reward associated with the choice they have made. In scenarios like the one above, as well as in the experience of most language learners, participants do not know whether the choice they have made (whether in a staffing decision, or in the segmentation strategy they have taken) is correct. How does statistical learning operate when a reward structure is not present? For example, learners are confronted with linguistic variance, and will receive slightly varied input depending on their geographical location or their current social situation, possibly receiving input from different speakers who may speak different dialects. Furthermore, since the majority of the world's population is multilingual, and are therefore exposed to multiple languages (Grosjean, 2010), this uniform variance approach may be an oversimplification of the challenge many learners face in the real world. While a monolingual learner inevitably encounters variance within their single language environment (such as speaker variability, context, and dialectical differences), a bilingual speaker additionally encounters structural variance that could signal the need for a second underlying causal model.

Variance and language acquisition

In the realm of language acquisition, learners can encounter variation that can be attributed to surface variance, or differences in the underlying structure. Surface variance can be attributed to differences in how the language is instantiated and presented, which can include (but is not limited to) dialectical differences and changes in speaker. Structural variance, however, is typically an indication of the presence of multiple models in the environment. Learners in bilingual environments, for example, are faced with the dilemma of determining where a structural switch happened and how to create models for the multiple languages in their environment. These types of variance are not mutually exclusive however, as a change in speaker can coincide with a change in the underlying structure. When multiple structures are present in the environment, feedback to the learner might come in the form of contextual changes that can notify a change has occurred.

Surface variance and statistical learning

Statistical learning research has thus far focused on investigating how learners account for surface variance in speaker voice (Houston & Jusczyk, 2000) and gestalt cues (Creel, Newport, & Aslin, 2004), among other manipulations. Previous research on statistical learning has used changes in voice as a contextual cue of change (Gebhart, Aslin, & Newport, 2009; Weiss, Gerfen, & Mitchel, 2009). Thus, this section will review research that has been conducted on variance in speaker voice, ultimately demonstrating that the presence of surface variance in the environment can aid in the acquisition of the underlying structure.

In order to investigate how young infants account for multiple speakers in their environment, Houston and Jusczyk (2000) conducted a follow up to the Juscyk & Aslin (1995) study, but presented participants with different speakers during familiarization and testing. During familiarization, they presented 7.5 month old infants with words produced by a female speaker. During the testing phase, they presented participants with passages recorded by a different female speaker that either did or did not contain the words they heard during familiarization. The findings replicate the original Jusczyk & Aslin, (1995) study, demonstrating that the infants listened longer to the passages that contained the familiar words, providing evidence that 7.5 month old infants can generalize token information to other speakers of the same sex. However, when the test passages were presented in a voice of the opposite sex, 7.5 month old infants did not listen longer to passages with familiar tokens when compared to passages with unfamiliar tokens. The same set of studies was then conducted with 10.5 month olds, and results show that the older infants are able to generalize not only across speakers of the same sex, but also across speakers of a different sex. This study provides suggestive evidence that, until about 10 months of age, children have not had enough experience with speakers of different genders, and therefore are not able to generalize speech across speaker gender. However, once a child has enough life experience with speakers of different genders, they become able to recognize differences in speaker voice as surface variance.

A recent study by Graf-Estes et al., (2014, under review) continued investigating this phenomenon by testing 8-and 10-month old infants on a statistical learning task with stimuli presented in 8 different female voices concatenated such that the speaker changed every 10 to 20 syllables. The test items were then presented in a novel female voice. During the testing phase, infants in both age groups listened longer to the words than to the part-words, demonstrating their ability to aggregate statistics across voices. Notably, infants were also able to generalize to male voices at test as well.

The aforementioned experiments exposed infants to high levels of variability, which may have facilitated learning. In a follow up study, Graf-Estes et al., (2014, under review) asked whether infants could generalize speech across only two speakers. The stimuli were the same as the previous experiments, but recorded by 2 female speakers alternating every 10-20 syllables. During the testing phase, neither of the age groups showed a difference in looking time between words and part-words, suggesting that the participants may not have been able to collapse the statistics from the speech heard from two separate speakers. These findings suggest that variance in a learner's input greatly impacts learning, which is consistent with work presented by Rost & McMurray (2009). The authors used a switch task with 14-month-old children (Werker et al., 1998) during which they manipulated the amount of speaker variability during familiarization. They showed that with greater variability, participants were better able to learn the minimal pairs presented during the experiment. This finding suggests that the presence of variability in the learner's input can facilitate the acquisition of the structure. Taken together, these studies converge on the fact that the amount of speakers in the environment can be indicative of how the learner should account for the variance. Having more speakers seems to suggest that the learner should assimilate the input produced by each speaker into a single underlying model.

Structural variance – simulating bilingual input

The previous studies demonstrate that learners can account for surface variance in their input; however, learners may also encounter variance that signals a change in the underlying structure. This structural variance may or may not be explicitly signaled, presenting the learner with an additional challenge: they must first realize a change in structure has occurred and then account for it. To investigate how multiple structures can be learned in different contexts, Conway and Christiansen (2006) conducted a study using an artificial grammar learning paradigm in which they exposed participants to two separate artificial grammars in different modalities. The first experiment exposed participants to one grammar in the visual modality and one in the auditory modality and found that participants were able to learn both. The second experiment tested whether participants could learn multiple grammars within the same modality, and found that with contextual cues of change participants could learn multiple grammars within the same modality. However, when both grammars were presented in the same modality, without any contextual cue that a change had occurred, participants were only able to master one of the two grammars.

In a recent study, Gonzales and colleagues (2015) tested how hearing multiple dialects impacts dialect-specific rule learning in infancy. They presented 12 month-old infants with streams of artificial languages that were structured to resemble dependencies in natural language between high and low frequency grammatical units. Results from a series of experiments reveal that infants actually track dialect specific statistics, and do not determine the grammatical structure of a given dialect by averaging statistics computed across dialects. Additionally, the researchers manipulated the presentation of the two artificial dialects, and found that infants successfully learned the two dialects when the stimuli were presented in 1 minute blocks, instead of when they were randomly interleaved (Gonzales, Gerken, & Gómez, 2015).

Gebhart, Aslin & Newport (2009), investigated learner's reaction to change in the underlying structure of the language in their environment by presenting participants with equal exposure to two languages in succession. They then tested whether participants had learned both languages by presenting them with words and part-words from both languages. When the same speaker recorded both languages, participants were only able to learn the first structure they heard, but performed at chance on the second structure. In order to see whether participants could learn both structures with the aid of a contextual cue, the researchers included a pitch shift at the point of the language change to provide participants with a cue that there might be a new structure. Even with the pitch shift cue though, participants were only able to perform significantly above chance on the first language, though the second language did trend toward significance. However, when a 30s pause was inserted between the two structures, or when exposure to the second structures was tripled, participants were able to learn both structures at above chance levels. These findings suggest that in order to learn multiple structures, participants need to overcome an expectation that the environment is not likely to undergo rapid change. In order to do so, participants require a correlated contextual cue of change, or additional exposure to the second structure. This, in turn, suggests that without explicit knowledge of a change in context, learners require extended exposure to a new structure containing novel statistics in order to recognize a change has occurred. This account is in line with the research on the role of attention in statistical learning; by providing contextual cues of change, the researchers are able to direct the participants' attention to the change in stimulus, which might have otherwise gone unnoticed.

Weiss, Gerfen & Mitchell, (2009) approached the issue of multiple structures from the perspective of a bilingual learner, and tested whether participants could form multiple representations for multiple inputs. They presented participants with 6, 2-minute blocks of two artificial languages. In one condition, the languages had congruent underlying statistics, meaning that both languages could be successfully segmented even if participants collapsed the statistics across languages. In the second conditions, the languages had incongruent underlying statistics, such that the collapsed statistics presented learners with statistical noise that would make it difficult to learn both. When either the congruent or incongruent languages were produced by two different speakers, participants were able to learn both languages. However, when the two languages were recorded by the same speaker, participants only learned the congruent languages at above chance levels, suggesting that they had not formed individual representations for each language but had collapsed statistics across the two. These findings are consistent with the findings of Conway & Christiansen (2006) using artificial grammars, providing further evidence that participants may need an explicit cue of change in order to form multiple representations and thus learn multiple structures.

A subsequent study by Zinszer & Weiss (2013) investigated whether additional variance in the input could aid participants in learning the structures. They hypothesized that multiple switches between the languages may help participants realize the existence of different structures. They first replicated the primacy effect found in Gebhart, Aslin & Newport (2009) and then revised the structure of the language exposure by switching between the structures three times while keeping the overall duration of each language constant. When participants heard the language change three separate times, they were able to learn both structures. Zinszer & Weiss concluded that the additional switches between languages can impact learning.

Statistical learning in bilinguals

While the aforementioned studies investigated how leaners react to change in their linguistic environment, they only included monolingual participants on their tasks (Weiss et al., 2009), or did not specify their participants' language background(Conway & Christiansen, 2006; Gebhart et al., 2009a). Given that these studies are relevant for multi-language acquisition, it would be natural to wonder whether bilingual learners might differ in their statistical learning abilities as a consequence of their experience with learning multiple languages. To date, there have been only a few studies focused on how bilinguals perform on statistical learning tasks. Thus far, there is no consensus on whether there are differences between bilinguals and monolinguals in this ability.

Yim and Rudoy, (2013) investigated this issue by testing monolingual and bilingual children between 5-13 years of age on two statistical learning tasks, one in the visual domain and one in the auditory domain. They found that while both groups performed above chance on the test in both modalities, there were no group differences in performance between the monolinguals and bilinguals. It is important to note, however, that both groups of participants were given the same length of exposure, and it is therefore possible that learning occurred at different rates for the monolinguals and bilinguals. In fact, a recent study comparing typically developing children to children with Specific Language Impairment found that while the outcome of a specific test was the same across groups, the time course of how the groups reached the end goal differed (Mainela-Arnold & Evans, 2014).

Kovacs & Mehler, (2009) tested the differences between monolingual and bilingual 7-month old infants on a rule-learning task. During this task, participants had to learn that a specific cue predicted the appearance of the reward in a specific location. During the first part of the study, infants learned that an auditory cue signaled a visual reward. However, in the second part, the infants were exposed to a different auditory cue, and had to learn that the visual reward changed as well. They found that the bilingual infants more readily learned two rules when compared to the monolingual infants, even when both groups were given an indexical cue of the change. A similar study tested whether 12 month-old bilingual infants are better at simultaneously learning multiple speech structures compared to monolinguals of the same age (Kovács & Mehler, 2009). They found that bilingual infants were able to learn multiple structures, while the monolingual infants were only able to do so when a contextual cue indicated the presence of multiple structures.

More recently, research from our lab has tested whether bilinguals are more willing to form multiple word mappings to a single object (Poepsel & Weiss, in review). The researchers found that adult bilinguals more readily form multiple word mappings in a cross situational statistical learning task than their monolingual counterparts. To our knowledge, only one study has directly compared bilingual and monolingual participants on a multi stream statistical learning task. Bogulski (2013) tested monolingual and bilingual adults on a task similar to the one used by Gebhart, Aslin & Newport, (2009) and found that bilingual populations show the same primacy effect reported by Gebhart et al., (2009) when presented with two consecutive structures.

As can be seen from the aforementioned studies, the literature testing bilingual participants on statistical learning studies has not come to a consensus about the differences in performance between monolingual and bilingual learners. While some studies find that bilinguals perform differently than their monolingual counterparts by more readily adapting to a change in the rules of their environment (Kovács & Mehler, 2009; Kovacs & Mehler, 2009) and learning multiple word mappings (Poepsel, & Weiss, in prep), others found no differences between the groups' performance on visual and auditory statistical learning tasks (Yim & Rudoy, 2013), or on their ability to learn two consecutive artificial languages (Bogulski, 2013). While bilingualism has been thought to shape cognitive function as a result of experience juggling multiple languages (e.g. Marian & Spivey, 2003, Kroll, Bobb, & Wodniecka, 2006), Bialystok (2015) recently proposed that the bilingual experience might actually change the way attention is directed to the environment. Following findings from studies that showed that bilingual infants paid more attention to talking faces, and were able to detect a change in the language spoken by the speaker even in the absence of sound (Sebastián-gallés, Albareda-castellot, Weikum, & Werker, 2012; Weikum et al., 2007), Bialystok hypothesizes that when an infant is in an environment that introduces two sets of sounds, cadences, structures, speakers and facial configurations, their attention is drawn to the contrasts between the two systems (Bialystok, 2015). Since contrasts create novelty, infants growing up in bilingual environments may learn to attend to subtle environmental differences. This hypothesis is supported by a recent study showing that bilingual infants pay attention to the mouth of a speaker longer over the course of development than monolingual speakers, who more quickly return to looking at the eyes when observing a speaker (Pons, Bosch, & Lewkowicz, 2015), as attending to the mouth may provide more information about contrasts present in the environment.

Thus, differences in how bilingual participants perform on statistical learning tasks may be dependent on the task itself. The studies described above tested different types of bilinguals on a variety of different tasks, and the lack of consistent results may have been masked my task differences and measures. It is possible that differences in the attentional mechanisms of bilinguals may impact their performance on tasks that create specific contrasts (as was seen in the Kovacs et al., (2009) studies), but that these differences in performance when compared to monolinguals may not be evident on the surface. As has been shown with other linguistic populations (see Mainela-Arnold et al., 2014 for a comparison of typically developing and SLI children), the outcome of a measure may be the same across groups, but the time course of achieving that outcome could differ. Previous tasks may not have been able to tap into these differences due to presenting both groups with set amounts of exposure and only relying on a post-test to determine differences across groups.

Chapter 2

Current Experiments

Previous research on multiple structures in statistical learning has found a robust primacy effect; such that only the first presented structure is learned unless there is a contextual cue of change that coincides with the switch the second structure, or exposure to the second language is tripled (Gebhart et al., 2009). Critically, however, all of the studies presented participants with exposure to both languages in succession, and then tested participants on their knowledge of both languages. Thus, researchers had no way of knowing when learning of either of the structures occurred.

The concept of stationarity (Aslin, 2014) could be used to explain the robust primacy effect. Stationarity implies that the learner has an a priori expectation that the environment is not likely to undergo rapid and frequent changes. This could lead participants to try to assimilate, or overgeneralize, the underlying structure of the first language presented during experiments. This suggestion is consistent with the notion of anchoring (Tversky & Kahneman, 1974). Learners may underadjust initial estimates when confronted with new information, as evidenced in the results of judgment and contingency learning experiments (such as Yates & Curley, 1986; Dennis & Ahn, 2001). Marsh and Ahn (2006) have argued that primacy effects occur because learners form a hypothesis about the relationship between events and underadjust their initial hypothesis as they gain additional evidence. Further evidence of primacy effects comes from conditioned learning; when a cue is reinforced to the point of near-perfection (such that the learner comes to expect only one outcome whenever the cue is presented), the learner fails to subsequently acquire an additional association (Blanco, Baeyens, & Beckers, 2014).

This tendency to underadjust may be related to the level of estimation uncertainty achieved by learners with respect to how well they can account for the statistical regularities in the environment (Qian et al., 2012). When learners lack confidence in their ability to accurately represent the causal model generating the statistics of their environment, they are less likely to view variance in the signal as indicative of a change in the underlying structure. In that vein, Qian and colleagues (2012) suggest that the results of Gebhart et al., (2009) might be attributable to the fact that learners never became certain in their representation of L1 prior to switching to L2.

I hypothesize that the primacy effect found by Gebhart and colleagues (2009) might actually be due to an alternative, but not necessarily conflicting, explanation. Participants may be learning early during exposure, and failure to detect a change in the underlying structure may be due to the composition of the learning environment. Learners may be acquiring the L1 relatively quickly and then receiving continued exposure to the already acquired set of statistics, which would in turn provide further reinforcement that they have learned the underlying structure, reinforcing their stationarity bias, and potentially lead to less consistent sampling of the environment. Therefore, the failure to detect that a change occurred in the stream may be related to the initial length of exposure. This hypothesis runs counter to that of Qian and colleagues (2012) as it assumes a relatively high degree of confidence in the participant's learning established early in familiarization. This hypothesis explains why participants may be able to overcome the primacy effect when they are given a cue of the change (allowing them to refocus attention and notice that a change in the statistics of the environment has occurred), or when exposure to the second language is tripled (if participants are sampling from the environment less frequently, extended exposure allows for the learner to still accrue enough information to demonstrate learning).

The scope of this project was to investigate the theory that participants may become entrenched in the statistics of the first language, leading them to not learn the second presented language. Additionally, this project took an individual differences approach to investigate whether differences in cognitive functioning can predict individuals' abilities to learn in a non-stationary environment. This project aimed at gaining insight into the specific time course of statistical sequence learning of multiple structures. Further, this project compares monolinguals and advanced second language learners on their ability to learn in a non-stationary environment, as I predict that experience with segmenting multiple structures in one's environment may enhance participants' abilities to learn multiple structures in an experimental setting.

I modified the paradigm used by Gebhart, Aslin & Newport, (2009) to be able to test participants' learning after every minute of exposure to determine both the length of exposure necessary to learn the statistical regularities underlying the artificial structures, and to account for entrenchment by advancing participants soon after learning the first set of statistical regularities has occurred. By testing for learning more frequently during familiarization, I may also be able to determine whether differences between monolingual and bilingual populations exist that were not apparent in previous research that relied exclusively on retrospective measures of learning that occurred after both artificial languages had been heard. Given their experience with learning multiple languages, and previous research showing that bilinguals may more readily learn multiple rules and word-to-world mappings in their environment, bilingual participants may perform differently than their monolingual counterparts. This difference could manifest itself in a speed advantage, wherein the bilingual participants might learn the underlying structures faster, or in the proportions of participants who exhibit the primacy effect, both differences that could not be determined by the designs of previous experiments.

Two sets of experiments during which participants are exposed to two consecutive languages follow. The first set (Experiments 1-3) investigates whether additional exposure to the first presented language after learning impacts participants' ability to learn the second language, and how contextual cues of change impact the ability to learn when participants have not become entrenched in the statistics of the first language. In Experiment 1, I familiarized participants to two artificial languages in five one-minute blocks, each followed by tests (and one retrospective test at the end). This provided a test of whether I could replicate the primacy effect observed by Gebhart and colleagues using this modified methodology. In Experiment 2, I advanced learners to the second language immediately after they achieved a high level of performance on one of the intermediate tests to determine whether performance was affected by additional exposure to the same language subsequent to learning. In Experiment 3, I used the same paradigm as Experiment 2 but added contextual cues such as a change in speaker voice and a 30 second pause to signal the change in structure to learners. Research has shown that experience juggling multiple languages may impact individual's cognitive control mechanisms, as constant competition between two languages requires inhibition and selection to produce the desired language (e.g. Marian & Spivey, 2003), and that speaking multiple languages may draw individuals' attention to contrasts in the environment (Bialystok, 2015). Thus, the second set of experiments (Experiments 4 and 5) compare monolinguals to bilinguals and advanced second language learners on the same paradigm used in Experiment 2 to determine whether experience with segmenting multiple languages in one's environment impact statistical learning abilities in non-stationary environments.

Investigating the primacy effect

Experiment 1

In this experiment I sought to replicate and extend the findings of Gebhart, Newport, & Aslin (2009) using a modified presentation method in which participants received an 8-item two alternative forced choice (2AFC) test following each one minute of language presentation. Learners encountered a total of five minutes of an L1 followed by 5 minutes of an L2 (each presented in 1-minute blocks). This paradigm allows not only for information regarding whether the first and second languages were learned, but also affords an estimate of how quickly they were acquired. In addition, participants received a flanker task in order to determine whether the ability to acquire L2 might be related to the ability to inhibit the statistics learned during L1 presentation.

Methods

Participants

Participants consisted of 54 Penn State monolingual undergraduate students (mean age = 18. years, sd=1.04; 38 females and 16 males) recruited from an Introduction to Psychology course. All participants received course credit for their participation in the study. An additional 25 participants were recruited but were excluded due to being proficient in a second language (n=19) or not following instructions (n=6).

Stimuli

The two artificial languages presented in this experiment were drawn from previous studies (Newport & Aslin, 2004; Gebhart, Aslin, & Newport, 2009). Each language consisted of 16 trisyllabic words constructed from consonant-vowel (CV) syllables. The languages were created using Speaker, a text-to speech application found in the speech synthesizer MacInTalk. The synthesizer was adjusted to remove any acoustic cues to the word boundaries and to produce equivalent levels of coarticulation across all syllables. Both speech streams were created using the same female voice (Victoria). Syllable duration was edited such that all syllables were .20-.22 seconds in length, thereby assuming that syllable duration could not cue learners about the position of the syllable within the word (see Newport & Aslin, 2004 for further details).

An inventory consisting of six consonants (d, k, b, p, g, and t) and six vowels (a, ae , e, i, o, u) was used in both languages. The vowels formed a consistent word frame while the consonants varied. There were two consonants that could occur in any consonant position (see Table 1). Consequently, the languages consisted of 16 possible words. These were concatenated to create a continuous speech stream in which each word was repeated the same number of times with the constraint that no word was ever reduplicated. The speech stream was looped to create a 67s stream, with a production rate of 284 syllables per minute. The transitional probability of the vowels within the word was 1.0, as each vowel was always followed by only one other vowel. Across the words, the vowel-vowel transitional probability was .5, as either frame could follow each word.

Table 2-1. Artificial language structure

	FRAMES	FILLERS
Language A	*_a_u_e*	[d_][k_][b_]
	_o_i_ae	[p_][g_][t_]
Language B	_ae_a_u	[t_][d_][k_]
	_e_o_i	[b_][p_][g_]

Each test item consisted of a word paired with a part word. Part words consisted of either the last syllable of one word followed by the first two syllables of the next (i.e., 3-1-2), or the last two syllables of one word followed by the first syllable of the next (i.e., 2-3-1). The stimuli were created in isolation with a falling intonation imposed at the end of every test item. Four words and part words were used in the test, with each item occurring twice and no pairings repeated. The ISI between test items was one second.

Procedure

Participants were seated in a sound attenuated chamber and instructed to attend to a recording played through headphones. They were informed that they would subsequently be asked questions about what they heard, but did not receive any further details regarding the experimental design.

Participants listened to one 67s block of the first language (L1). The ordering of the languages was counterbalanced across participants. After this familiarization period, participants received an 8-item 2AFC test. They were then presented with another 1-minute block of familiarization followed by another test. The cycle of familiarization and test repeated five times. After the fifth test participants received exposure to the second language (L2). The presentation of L2 was identical to L1 (i.e., five blocks of familiarization-test). Participants then received a 16-item post-test on both languages.

Each participant also completed a flanker task (Eriksen and Eriksen, 1974) that was modified according to Bunge, Dudukovic, Thomason, Vaidya, and Gabrieli (2002; see also Emmorey et al., 2009). The stimuli consisted of red arrows flanked by four black arrows. Participants were asked to respond to the red arrow by clicking the left or right mouse button depending on the direction the arrow was pointing. There were three types of blocks during the experiment: control, conflict and go/no-go. During the control blocks, participants saw a single arrow and were asked to respond with the direction of that arrow. During the conflict block, participants saw congruent and incongruent flanking arrows. During the go trials, the participants saw congruent or incongruent trials as in the second block. During the no-go trials, participants viewed the red arrow flanked by black Xs, which indicated that they should not respond to the stimuli.

Each trial began with a fixation cross in the middle of the screen for 250ms. Participants then viewed the target stimuli for 2s or until they produced a response. Prior to each block, participants received 12 practice trials. Each of the three block types was presented twice for a total of 84 trials per block type. In additional to these blocks, there were also two mixed blocks during which participants received intermixed congruent/incongruent and go/no go trials. Both response times and accuracy were recorded.

Results

Across all blocks of the experiment, the average accuracy for language L1 was 6.14 (sd=1.67) out of 8 correct on the test following familiarization. For L2, the average was 4.97 (sd=1.41). These scores were significantly different (paired t-test: $t(53) = 6.519$ $p<.001$).

A repeated measures ANOVA testing the effect of language and block showed a main effect of language ($F(1,52) = 42.49$, $p<.001$) such that L1 accuracy was significantly higher than L2 accuracy. When looking at blocks, there was no linear ($F(1,52) = 1.58$ $p= .214$) or quadratic trend ($F(1,52) = .028$

p<.867), suggesting that learning did not improve in a systematic way across the session. The interaction between language and block was also not significant ($F(1, 52) = .276$, $p=.602$). While participants were more accurate overall on L1, their accuracy remained consistent over the course of exposure ($\alpha= .758$). Accuracy on L2, however, was much less reliable over time ($\alpha= .541$).

In order to determine a threshold for accuracy that would provide an index of consistent learning, performance after reaching the peak accuracy was plotted (for subjects who scored above chance). Participants whose maximum performance was 5 or 6 out of 8, did not typically maintain that accuracy over continued exposure, sometimes regressing to chance performance (see Figure 2-1). However, for those who reached 7 or above on one of the tests, their above chance performance persisted for the duration of the familiarization period with that language. Consequently, some of the subsequent analyses focused on learners who achieved seven or above as an index of strong learning. Using this criterion (7 or above on any of the tests), 42 out of the 54 participants reached threshold on the first language, and, on average, it took them 1.59 (sd = .94) blocks. However, only 15 of the 42 reached this threshold on the second structure.

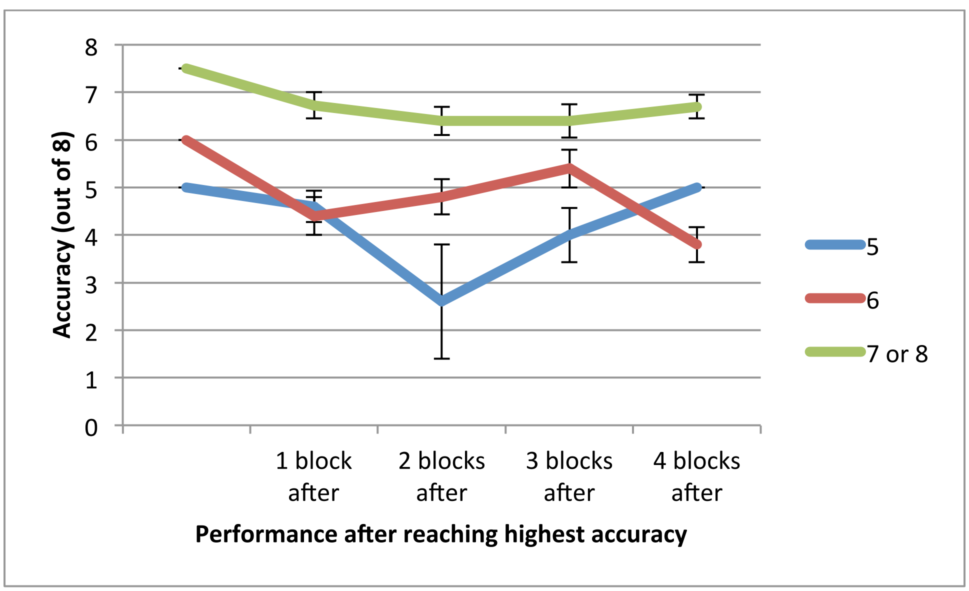

Figure 2-1. Average performance on blocks after reaching highest accuracy during exposure.

There was a significant difference in the number of blocks it took to learn L1 (to criteria) for the participants who also acquired L2 compared to those who only successfully learned L1 ($t(40) = -2.59$, $p=.013$). Participants who only learned L1 reached criterion after an average of 1.36 blocks (sd = .66) of exposure. For participants who learned both L1 and L2, they reached this criterion for L1 in 2.06 block (sd = 1.22) and 2.53 minutes for L2 (sd = 1.30). This difference in amount of exposure for L1 and L2 learning was not significant ($t(14) = -1.047$, $p=.313$).

A subset of participants (n=24) received post-tests for both of the languages to determine whether learning was retained. Participants who only learned L1 scored an average of 10.81 (sd = 3.15) out of 16 on the posttest for L1, which is significantly above chance ($t(10) = 2.961$, $p=.014$). Participants who learned both languages averaged 9.87 (sd = 2.23) out of 16 on the post-test for the first language and 9.13 (sd = 3.60) out of 16 on the post-test for the second language. The first post-test was significantly above chance ($t(7) = 2.376$, $p=.049$) while the post-test for the second language was not significantly above chance ($t(7) = .883$, $p=.406$).

There was no significant difference ($t(41) = .166$ $p=.869$) in reaction time for the Flanker task for participants who were able to learn one versus both languages. Participants who learned both structures exhibited reaction times that were on average 482.77 ms (sd = 71.79) while participants who only learned the first structure had reaction times of 487.20ms (sd = 86.63). Accuracy on this task was also calculated to determine whether differences in learning could be accounted for by an overall lack of effort during the experimental session. Participants who learned both languages performed the flanker task with an average accuracy of 88.71% (sd = 12.49) and participants who just learned the first language performed with an accuracy of 85.79% (sd = 13.48) which was not significantly different ($t(41) = -.681$, $p=.500$). Average reaction times and accuracy on the flanker task were also calculated for participants who did not reach criterion on either language (n=12). Their average reaction time was 556.52 ms (sd = 84.11) which was significantly different from those who learned both languages ($t(22) = 2.312$, $p=.031$), and those who learned only the first language ($t(37) = 2.197$, $p=.034$). Their accuracy was 79% (sd = 21.66) and while it was not significantly different from either of the other groups, it could be an indicator that these participants were less attentive to the task.

Discussion

Overall, the findings from Experiment 1 were consistent with the primacy effect reported in Gebhart et al., (2009). When I calculated accuracy for each language by collapsing across all tests, participants' accuracy was significantly higher for the first language relative to the second. Despite the changes that I made to the experimental procedure, I was able to replicate the results reported in the original study. Focusing on participants who scored seven or above on any of the tests, 36% of those who learned the first structure were also able to learn the second structure. While previous studies have not looked at the percentage of participants who successfully acquired each language, the percentage found here sheds novel insights into the primacy effect. While the primacy effect was replicated with these data, this design allows us to begin to look at how individuals may perform differently in these tasks.

Notably, participants who only reached criterion on the first structure were significantly faster at learning L1 than those who learned both structures. Those who learned L1 took 1.31 blocks on average to reach criterion on L1, while those who learned both required 2.06 blocks. Most participants in the former group therefore received 3 to 4 blocks of exposure to the same set of statistics before being advanced to the second. As noted above, this additional exposure may have impacted their ability to acquire the second stream. In Experiment 2, I more directly test whether this additional exposure after learning might impact the ability of learners to acquire L2.

Notably, the participants who learned both languages did not retain the second language significantly above chance when tested at the post-test. While they did reach the predetermined criteria on the second language, it is possible that the additional exposure to the first language after learning may have still had a blocking effect on L2, causing participants to not retain it. Additionally, by testing every minute and gaining insight into the progression of learning, these results provide some support that participants might never reach a low level of estimation uncertainty (Qian et al., 2012), while also showing that some participants might in fact reach a high level of confidence. The participants whose highest accuracy only reached 5 or 6 out of 8 did not maintain that accuracy throughout the course of exposure, supporting the theory that they might never have become confident in their interpretation of the environment, while those who reached 7 were able to maintain it (see Figure 2-1), demonstrating a low level of estimation uncertainty.

Experiment 2

In Experiment 2, I modified the design of Experiment 1 to allow participants to advance to the second language as soon as they reached the criterion of 7 out of 8 on the test of L1 that followed each block. Participants received up to 5 one-minute blocks of L1 followed by up to five blocks of L2. The goal of this experiment was to determine whether allowing participants to advance as soon as they learned L1 would reduce the primacy effect.

Methods

Participants

Participants consisted of 57 monolingual Penn State undergraduate students (45 females, 12 males; mean age=18.87; sd=.97) from an Introduction to Psychology course who participated in this experiment for credit. None had previously participated in a statistical learning experiment. An additional 24 participants were recruited but were excluded from analysis due to speaking another language proficiently (n=20), experimenter error (n=3) or sleeping during the study (n=1).

Stimuli

The stimuli were the same as those used in Experiment 1.

Procedure

There was one primary difference between Experiments 1 and 2. In Experiment 2, participants received exposure to L2 immediately after they scored a 7 or above out of 8 on one the tests following the 1-minute familiarization periods of L1. Participants received up to 5 familiarization-test blocks for the first language, and if they did not reach our criteria of 7 out of 8 correct during those 5 blocks, the experiment was terminated. If participants scored 7 out of 8, they received up to 5 familiarization-test blocks of L2. After completing this portion of the experiment, participants completed a 16-item post-test on the languages they learned (if they learned both, they received post-tests on both languages; if they only learned the first language, they received a post-test on just the first language), followed by the flanker task.

Results

22 participants did not reach the criterion of 7 out of 8 correct on L1. Of the 35 participants who learned L1, 23 of them (66%) also learned L2. The following set of analyses includes only the participants who learned L1. On average, it took 2.41 (sd = 1.62) blocks to learn L1 for learners who did not acquire L2. Participants who were able to learn both languages required 1.87 (sd = 1.22) blocks to learn L1 and 2.69 (sd = 1.36) blocks to learn L2 (the difference in time to learn the L1 and L2 was not significant; paired t-test: $t(22) = -1.716$, $p=.10$). There was no significant difference between groups in the length of time it took for participants to learn L1 (independent samples t-test: $t(33) = 1.125$ $p=.269$). (See Figure 2-2 for a comparison to Experiment 1).

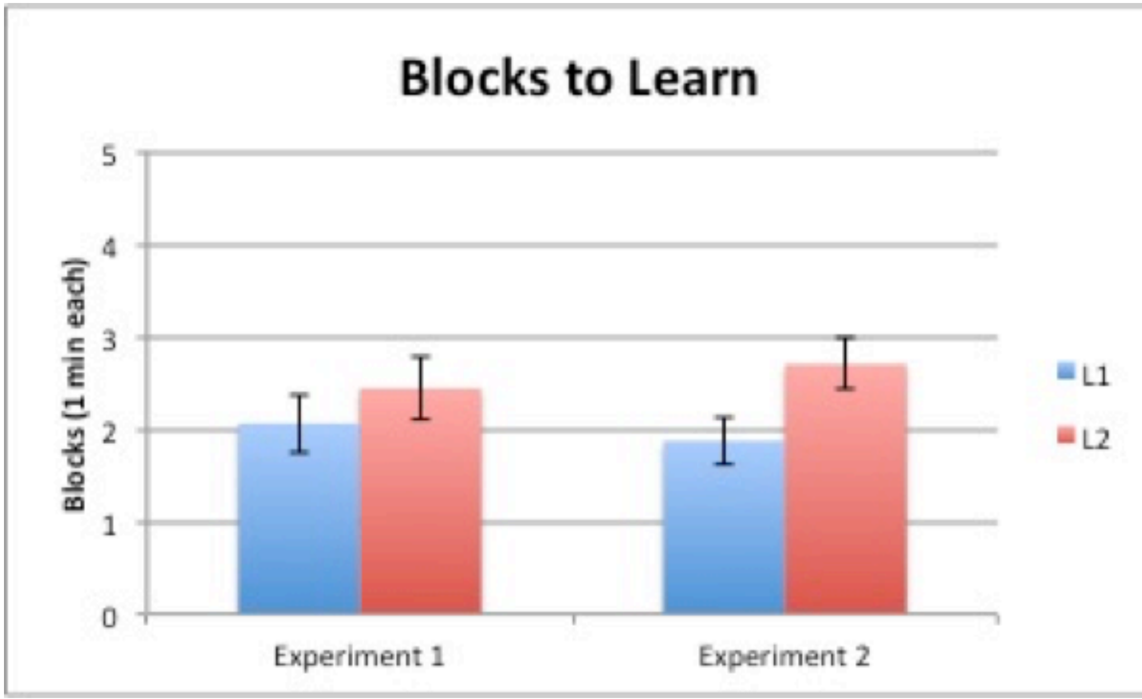

Figure 2-2. Blocks to learn each language for participants who learned both L1 and L2 across Experiments 1 and 2.

The post-test data was analyzed to determine whether participants who had learned both languages were able to retain what they had learned. The average accuracy for the L1 post-test was 10.22 (sd = 2.57) out of 16 ($t(22) = 4.129$, $p<.001$) and 11.22 (sd = 2.76) for L2 ($t(22) = 5.585$ $p<.001$). There was no significant difference between the posttests for the different languages (paired t-test: $t(22)= -1.151$ $p=.262$).

Reaction times on the flanker task for correct trials were computed for the different groups of learners. Those who learned both languages exhibited significantly faster reaction times on the flanker task (average = 437.80, sd = 49.27) than those who only learned L1 (average = 501.9, sd = 103.67; independent samples t-test $t(42) = -2.790$ $p=.008$). Participants who learned both structures performed the task with 81.5% accuracy (sd = 19.2%) and those who only learned the first structured performed with an accuracy of 81.00% (sd = 18.8%) which was not a significant difference (independent samples t-test t(42)

= .159 p=.874). The flanker data for participants who were not able to learn the first language were also analyzed. Their average reaction time was 594 ms long (sd = 96ms) and their average accuracy was 86.18 (sd = 17.85). Their reaction times were significantly slower than the group that learned both languages ($t(42) = 6.808$, $p<.001$) as well as the group that only learned one language ($t(29) = 2.677$ $p=.012$). However, their accuracy was not significantly different from either group ($t(42) = .799$ $p=.429$; $t(29) = 1.135$, $p=.265$).

In Experiment 1, only 34% (n=15 out of 42) reached criterion on L2 after learning L1. However, in Experiment 2, 66% of participants (n=23 out of 35) reached threshold on L2 after learning L1, representing a significant difference in performance (Barnard test Wald statistic = 2.62, p =.005).

Discussion

In Experiment 2, I allowed participants to advance to L2 as soon as they reached criterion on one of the tests of L1. I found that 66% of participants who learned L1 were also able to learn L2, which is a significant increase when compared to Experiment 1. Notably, even though most participants learned the first structure in under 2 minutes, they were able to retain learning at the post-test, as they did also for L2. The post-test scores exhibited in this experiment were also considerably higher than those exhibited by participants in Experiment 1. These differences could be due to the fact that participants in Experiment 2 had, on average, less time between each of the presented languages and its respective post-test. Additionally, those who learned the second language retained it to significantly above chance levels at the post-test, which was not the case in Experiment 1. This difference across Experiments could be due to the fact that the participants in Experiment 2 took the post-tests right after reaching their highest accuracy on a test for the second language. Interestingly, those who were able to learn both languages exhibited significantly faster reaction times on the flanker task relative to those who only learned L1 (or those who not learn either language). This may indicate that inhibitory control may be related to the ability to adapt in a constantly changing environment.

Experiment 3:

Previous research has shown that contextual cues can help learners overcome the primacy effect (Gebhart, Aslin, & Newport, 2009; see also Weiss, Gerfen & Mitchel, 2009). Gebhart et al., (2009) found that the second language was learned significantly above chance when participants were explicitly told they would be learning two languages and that the second language would start after a 30-second pause. Additionally, they found that learning of the second language trended towards significance when one of the languages was pitch shifted, while Weiss et al., (2009) showed that changing the speaker of one of the language allowed participants to create different representations for each presented language. In Experiment 3, I adopted the same method as Experiment 2 but added a series of contextual cues that marked the switch into L2 to investigate whether contextual cues could further increase the percentage of participants who could learn the second language.

Methods

Stimuli:

The stimuli for these conditions were the same as those used in Experiment 2, except for the following changes: In Conditions 1 and 3, we pitch shifted one of the languages by 60 Hz[1] (using Audacity), to sound as if it has been recorded by a male speaker, while the other language remained in the original female wvoice.

[1] The pitch-shifted language was tested in isolation, and was learned to the same accuracy level as the original language.

Procedure:

The procedure for the following three conditions was the same as Experiment 2, except for the following changes. In Condition 1 (Pitch-Shift Condition), one of the presented languages was the pitch-shifted language. The order of the languages was counterbalanced, such that half the participants received the original language first, while the others received the pitch-shifted language first. In Condition 2 (Explicit Cue Condition), the instructions given to the participants changed, such that they were told that during the experiment they would be learning two languages. In Condition 3 (Combined Condition), the participants were explicitly told that they would be learning two languages, that the second language would start after a 30-second pause in the experiment, and that the second language would be presented by a novel speaker (the pitch-shifted language was used as in Condition 1). The order of the languages was again counterbalanced.

Participants

One-hundred and forty-six three Penn State undergraduate students from an Introduction to Psychology course participated in this experiment for credit. 53 participants (40 female; mean age = 18.58; sd=.95) were assigned to Condition 1, 50 participants (40 female; mean age = 18.38; sd=.69) were assigned to Condition 2, and 43 participants (30 females; mean age = 19.43; sd=1.84) were assigned to Condition 3. None had previously participated in a statistical learning experiment**.** Across all conditions, an additional 75 participants were excluded due to speaking a second language proficiently (n=60),, or for not following experimenter instructions (n=15).

Results

In the Pitch-Shift condition, 46% of participants were able to learn the second language after learning the first, which was a significantly lower percentage when compared to Experiment 2 (Barnard test Wald statistic = 1.69, p=.048). In the Explicit-Cue condition, 57% were able to learn the second language, which was not significantly different from Experiment 2 (Barnard test Wald statistic = .77, p=.339). And in the Combination condition, 56% learned the second after learning the first, which was also not significantly different from Experiment 2 (Barnard test Wald statistic = .837, p=.241).

For each condition, participants who only learned L1 were not significantly different from those who learned both languages in the rate with which they acquired L1 (all *ps* > .05). Similarly, for those that learned both languages, there were no significant differences in the rate with which they learned each language in any of the three conditions (all *ps* > .05) (see Table 2-2).

Table 2-2. Time to learn in minute long blocks in Experiment 3

	Condition 1: Pitch Shift		**Condition 2: Explicit Knowledge**		**Condition 3: Pitch Shift + Explicit Knowledge + Pause**	
	L1	**L2**	**L1**	**L2**	**L1**	**L2**
L1 Only	1.45 (1.62)	-	1.67 (0.84)	-	1.60 (1.12)	-
L1 + L2	2.16 (1.01)	2.42 (1.46)	1.79 (1.28)	2.50 (1.56)	1.63 (0.68)	2.37 (1.42)

The post-test data was analyzed to determine whether participants who had learned both languages were able to retain learning. For those that learned both languages, they retained both L1 and L2 above chance (L1: Pitch-Shift Condition mean = 11.00 (1.45), t(18) = 9.00, p<.001, Contextual-Cue Condition mean = 9.87, t(23) = 3.353, p=.003, Combination Condition mean = 10.16, t(18) = 3.794, -=.001, L2: Pitch-Shift Condition mean = 10.05, t(18) = 5.318, p<.001, Contextual-Cue Condition mean = 11.13, t(23) = 6.281, p<.001, Combination Condition mean = 11.33, t(18) = 5.926, p<.001). Retention did not differ significantly between the languages across conditions (all ps > .05). Reaction times on the flanker task did not predict participants' abilities to learn both structures, as they had in Experiment 2 (see Appendix A).

Discussion:

While previous research showed that adding contextual cues, such as a 30-second pause that correlated with the change in language, helped participants overcome the primacy effect; contextual cues in this paradigm did not increase the percentage of learners who successfully acquired the second language after learning the first. When one of the languages was pitch-shifted to provide a cue that the underlying structure had changed, only 47% of participants were able to learn the second language. This percentage is actually significantly smaller when compared to the percentage of participants who were able to learn the second language in Experiment 2. Providing participants with explicit knowledge that they would be learning two languages, and in a separate condition providing them with the explicit knowledge combined with a 30-second pause and pitch shift, did not change the percentage of participants who learned the second language. Pitch shifting a language has previously been shown to improve learning, but not be sufficient to entirely overcome the primacy effect (Gebhart, Aslin, & Newport, 2009), and it is therefore possible that in this paradigm it introduced variance in the environment that is not always representative of a change in the underlying structure (such as surface variance described in the introduction).

Unlike in Experiment 2, performance on the Flanker Task did not predict participants' learning on the statistical learning task. It is possible that if participants are already noticing the change in the underlying structure (as in Experiment 2), contextual cues are not necessary and thus do not have an additive effect.

Comparing monolinguals and bilinguals

While some previous research has found no differences between monolinguals and bilinguals on visual and auditory statistical learning tasks (Yim & Rudoy, 2013), others have found that bilingual infants more readily learn two rules (Kovács & Mehler, 2009; Kovacs & Mehler, 2009) and that adult bilinguals may be more willing to form multiple word-to-world mappings (Poepsel & Weiss, in review). A study comparing monolinguals to late second language learners on a multi-stream statistical learning task found that both groups exhibit the same primacy effect (Bogulski, 2013) found in Gebhart et al., (2009). Experiments 4 and 5 compare bilinguals and advanced second language learners to monolingual participants to determine whether managing two languages (Experiment 4) and being actively engaged in learning a second language (Experiment 5) impacts statistical learning abilities in a non-stationary environment (e.g., Marian & Spivey, 2003; Kroll, Bobb, & Wodniecka, 2006).

Experiment 4

Experiment 4 served as a pilot study comparing bilingual's performance on Experiment 2. Since bilinguals have experience segmenting, forming representations for, and producing multiple underlying structures, they may perform differently on this paradigm than their monolingual peers. While a previous study found that bilinguals did not differ on the primacy effect found by Gebhart et al., (2009) (Bogulski dissertation, 2013), the paradigm of Experiment 2 might allow for finer grained differences to be observed.

Methods

Participants

Participants were 18 Spanish-English bilinguals (12 females, 7 males; mean age = 28.44, sd = 9.5) from the Penn State Community. All participants received $10 for participating in the experiment. None had previously participated in a statistical learning experiment.

All participants were asked to self-rate their proficiency in both languages. Their average proficiency on Spanish (L1) was 9.55 out of 10 (sd = .70) and on English (L2) was 9.11 out of (sd = 1.02). The age at which participants learned English ranged from birth (simultaneous bilingual) to age 25.

Stimuli

The stimuli were the same as those used in Experiments 1 and 2.

Procedure

The procedure was the same as that in Experiment 2.

Results

Eight participants did not reach the threshold on the first language. Of the 10 participants who showed significant learning on the first language, 3 of them (30%) also reached criteria on the second language.

The following analyses only include the 10 participants who were able to learn the first language, and are thus rather underpowered. On average, it took participants 1.9 blocks (sd = .35) to learn the first language. For those who were able to learn the second language, it took 1.33 blocks (sd = .33) to learn the first language, and 3.0 blocks (sd = 1) to learn the second language. This difference is not statistically significant because of a lack of power in the analysis. For participants who learned the first language but did not go on to learn the second, it took 2.14 blocks (sd = .45) to learn the first language. This is also not a significant difference when compared to the group who was able to learn the both languages.

The post-test data was analyzed to determine whether participants who were able to learn both languages were also able to retain both. The average accuracy on the post-test for L1 was 8 out of 16 (sd = .58), which is not significantly above chance. The average accuracy for post-test on L2 was 10.66 out of

16 (sd = .88) which is significantly above chance. Participants were significantly more accurate on the post-test for L2 than for the post-test for L1.

Reaction times for accurate trials during the mixed block on the flanker task were calculated for each of the groups. Participants who were able to learn both languages had average reaction times of 479.20 ms (sd = 103.69), those who only learned the first language had average reaction times of 503.53 ms (sd = 72.6); and those who did not meet criteria on the first language exhibited reaction times of 511.88 ms (sd = 164.62). The differences in reaction times between participants who learned both languages and those who only learned the first was not significant (independent samples t-test: ($t(8)$ = .433 p= .677). There was no significant difference on reaction times between participants who learned one language compared to those didn't learn the first (independent t-test: $t(13)$ = .124 p= .904. There was also no significant difference between the reaction times of participants who learned both languages compared to those who did not even learn the first language (independent t-test: $t(9)$ = .315 p= .760).

Accuracy on the flanker task was also calculated for each of the groups. Those who learned both languages performed with 82.87% accuracy (sd = 4.4%), those who just learned the first language performed with 83.93% accuracy (sd = 2.5%); and those who did not learn either language performed with 78.65% accuracy (sd = 13.5%). There was no significant difference on accuracy between participants who learned both languages and those who only learned the first (independent samples t-test: $t(8)$ = .491, p= .635), between those who learned the first and those who learned neither (independent samples t-test: $t(13)$ = -1.102, p = .330) or between those who learned both and those who learned neither (independent samples t-test: ($t(9)$ = -.514 p = .619).

Discussion

Based on this preliminary data, the bilingual group in this experiment performed differently than the monolinguals in Experiment 2. In Experiment 2, 67% of participants who learned the first language were also able to learn the second, however in this experiment, only 30% of participants were also able to learn the second language after learning the first. It is possible that this percentage would change with a larger group of participants. However, this difference may be attributable to the fact that they were not a truly homogenous group. While all participants were Spanish-English bilinguals with high proficiency in both languages, they learned English at varying times over the course of their lives. It is possible that there may be differences between early and late bilinguals on this type of task.

Experiment 5

In order to address the potential variance introduced by having non-homogenous bilinguals in Experiment 4, Experiment 5 compares a homogenous sample of advanced second language learners to a matched monolingual sample to further investigate the results found in the Experiment 4 (pilot bilingual data). A number of fluency and executive function tasks were also added to this experiment to ensure that the two groups were matched.

Methods

Participants

The advanced second language learners were recruited from upper level Spanish courses, with the requirement that they had completed a minimum of 5^{th} semester Spanish (Spanish 215). The monolinguals were recruited from an online research volunteer website at Penn State. All participants were compensated $10 for their participation. The bilingual participants were 29 English-Spanish learners aged between 19 and 22 (mean age = 20.69, sd = 1.17, 4 males). The monolingual participants were 25 native English speakers between 19 and 25 (mean age = 21.24, sd = 1.36, 5 males). None of the monolinguals spoke a second language fluently.

Stimuli

The stimuli were the same as the ones used in Experiment 2.

Procedure

The procedure for the statistical learning task was the same as Experiment 2 in which participants were advanced to the second language as soon as they reached criterion on the first language. Participants also performed an expanded range of fluency and executive function tasks. For the verbal fluency tasks, which is a widely used word retrieval task that assesses language representation and executive function (Luo, Luk, & Bialystok, 2010), participants were asked to list as many exemplars as they could think of in a given category in a minute (see (Gollan, Montoya, & Werner, 2002; Luo, Luk, & Bialystok, 2010 for other studies investigating verbal fluency in bilingual and monolingual participants).Bilingual participants performed this task in three categories in English and three categories in Spanish, while the monolingual participants only performed the task with the English categories. The categories in English were vegetables, animals, and furniture, while in Spanish participants were asked to name exemplars that belonged to the categories of fruits, body parts, and clothing. These categories were chosen because they all offer a large number of potential exemplars, which would allow for variance in the English condition, but also because they are relatively basic categories that Spanish learners should have access to in their second language. Exemplars were coded as correct if they conceivably fit the category. In the case that a subordinate exemplar was named along with a superordinate category ("pepper" was named as well as "serrano pepper"), only the subordinate items were credited (Sandoval et al., 2010). Since this group of participants are still actively learning their second language,, tokens that were mispronounced by only one phoneme were also included (such that participants would say "zapatas" instead of "zapatos" and it would still be counted).

For the executive function tasks, participants performed both the AX-CPT task and an Operational Span task. The AX-CPT task (Morales, Gómez-Ariza, & Bajo, 2013) measures inhibitory control, but unlike the Flanker task used in the previous experiments, measures both reactive and proactive inhibitory control(Braver, 2012).. Reactive inhibitory control is engaged when the need to inhibit a response only becomes evident at the end of the trial, while proactive inhibitory control occurs when even from the beginning, the participant knows they will need to inhibit a response. During the task, participants see a series of five letters, the first and last in a red font. Participants are asked to press the "no" button to the first four letters of the series regardless of letter identity and press the "yes" button to the last letter only if it is an "X" and the first letter is an "A". The task is comprised of 70 trials during which the first and last letters are "A" and "X" (hereafter AX trials), while an additional 10 trials begin with "A" but do not end with "X" (hereafter AY trials, testing reactive inhibitory control), 10 trials which begin with another letter but end in "X" (hereafter BX trials, testing proactive inhibitory control), and 10 trials which begin and end with letters that are not "A" and "X", respectively (hereafter BY trials, which eliminate the need for inhibitory control). This task provides both error rates and reaction times for participants in all 4 conditions of the experiment. The Operational Span Task is meant to test the participants' working memory abilities. Participants are presented with a math problem (e.g. 2*5 +1 = 10) and are asked to indicate whether it is correct or incorrect. Upon answering whether the problem is correct or incorrect, they are shown a word, and asked to remember that word. Over blocks of increasing difficulty, participants are asked to remember and recall more and more words (starting with 3 blocks of 2 math problems and 2 words, followed by a linear increase until 7 math problems and 7 words are reached). The score they receive on the task is the number of words correctly recalled after correctly answering a math problem.

Results

The two groups of participants were matched on their age, working memory abilities (through the OSpan task) and on their fluency in their native language (English verbal fluency). See Figure 2-3 for average values for each group.

Table 2-3. Comparison of monolinguals and bilinguals on control tasks.

	Monolinguals	Bilinguals
Age (*sd*)	21.24 (1.36)	20.69 (1.17)
English Verbal Fluency (*sd*)	16.37 (3.70)	16.79 (3.69)
OSPAN (*sd*)	43.72 (7.43)	47.32 (6.58)

Monolinguals

Seven participants did not reach the criterion of 7 out of 8 correct on L1. Of the 18 participants who learned L1, 6 of them (33%) also learned L2. The following set of analyses includes only the participants who learned at least L1. On average, it took participants who did not go on to learn L2 1.42 (sd = .67) minutes to learn the first language. Participants who were able to learn both languages required 2.17 minutes (sd = 1.83) to learn the first language and 2.67 (sd = 1.37) minutes to learn l2.. There was no significant difference between the length of time to learn L1 between participants who were able to learn both languages compared to those who only learned the first presented language (independent samples t-test: $t(16) = -1.287$, $p=.217$). There was also no significant difference in the length of time it took to learn participants to learn L1 and L2 when they were able to learn both (t-test: $t(5) = .473$, $p=.656$).

The post-test data was analyzed to determine whether participants were also able to retain the languages they learned during the exposure phase. The average accuracy on the L1 post-test for participants who just learned the first language was 10.75 (sd = 2.86) out of 16. For participants who learned both languages, the average accuracy on the post-test for L1 was 9.00 out of 16 (sd = 1.67) and 10.67 (sd = 3.25) for L2. All post-test averages are significantly above chance (all $p <.01$) except for the L1 post-test for participants who went on to learn the second language.

The numbers of exemplars named in each language were calculated for participants, and subdivided based on their performance on the statistical learning task. Participants who learned both languages named 15.87 (sd = 2.68) exemplars in English while participants who only learned the first presented language named 16.08 (sd = 3.35) exemplars in English. Participants who did not learn either language during the statistical learning task produced 17.30 (sd = 5.17) exemplars in English. The difference in number of exemplars produced in English and Spanish was not significantly different across all groups (all ps >.05). ,

For the AX-CPT task, participants' reaction times and error rates were calculated for each trial type. Reaction times on the different conditions were not significantly different for participants depending on whether they learned the first or both languages during the Statistical Learning task (all p>.05). Error rates in the AX, BX, and BY conditions were not significantly different across groups of learners (all p>.05), however participants who only learned the first language produced significantly more errors on the AY trials when compared to those who learned both languages (t(16) = 2.252, p = .039). Participants who learned both languages were significantly more accurate on BY trials when compared to those who did not learn either language t(11) = 3.411, p=.006). When compared to the group who only learned the first language, participants who did not learn either language exhibited significantly slower reaction times only on AX, AY and BX trials (all p<.05), but only showed significant differences in error rates on BY trials (p =.023). See Figure 7 for detailed reaction times and error rates for all conditions. Additionally, there were no significant correlations between the trial types and performance on the statistical learning task (all p > 0.05).

On the Operation Span Task, there were no significant differences in the scores between participants who learned both languages and those who are only able to learn the first. There were also no significant differences in the scores of participants who did not learn either language when compared to either of the previous groups (see Table 2-4).

Table 2-4. Mean proportions of errors and average reaction times for monolinguals depending on outcome during the Statistical Learning task.

	Monolinguals					
Learning	0		1		2	
Condition	Err (%)	RT (msec)	Err	RT (msec)	Err (%)	RT (msec)
AX	.13(.11)	305.46(27.48)	.07(.02)	376.83(68.5)	.15(.19)	346.79(96.79)
AY	.25(.11)	369.14(85.90)	.34(.23)	478.72(93.44)	.11(.09)	438.14(97.13)
BX	.16(.13)	245.26(62.48)	.10(.09)	317.61(65.57)	.15(.24)	312.93(113.89)
BY	.08(.05)	275.54(66.68)	.02(.04)	312.48(69.16)	.00(.00)	286.11(129.94)

Bilinguals

Nine participants did not reach the criterion of 7 out of 8 correct on L1. Of the 20 participants who learned L1, 9 of them (45%) also learned L2. The following set of analyses includes only the participants who learned at least L1. On average, it took participants who did not go on to learn L2 1.45 (sd = .69) minutes to learn the first language. Participants who were able to learn both languages required 2.11 minutes (sd = 1.36) to learn the first language and 2.33 (sd = 1.41) minutes to learn l2. There was no significant difference between the length of time to learn L1 between participants who were able to learn both languages compared to those who only learned the first presented language (independent samples t-test: $t(18) = -1.399$, $p=.179$). There was also no significant difference in the length of time it took participants to learn L1 and L2 when they were able to learn both (t-test: $t(8) = .339$, $p=.369$).

The post-test data was analyzed to determine whether participants were also able to retain the languages they learned during the exposure phase. The average accuracy on the L1 post-test for participants who just learned the first language was 12.09 (sd = 1.92) out of 16. For participants who learned both languages, the average accuracy on the post-test for L1 was 9.66 out of 16 (sd = 1.41) and 10.77 (sd = 1.85) for L2. All post-test averages are significantly above chance (all ps $<.01$).

The numbers of exemplars named in each language were calculated for participants, and subdivided based on their performance on the statistical learning task. Participants who learned both languages named 15.99 (sd = 1.95) exemplars in English and 9.59 (sd = 3.02) in Spanish. Participants who only learned the first presented language named 16.81(sd = 5.13) exemplars in English and 9.87 (sd = 4.98) in Spanish. Participants who did not learn either language during the Statistical Learning task produced 17.65 (sd = 3.27) exemplars in English and 10.25 (sd = 3.07) in English. The difference in number of exemplars produced in English and Spanish was not significantly different across all groups (all ps $>.05$). ,

For the AX-CPT task, participants' reaction times and error rates were calculated for each trial type. Reaction times on the different conditions were not significantly different for participants depending on whether they learned the first or both languages during the Statistical Learning task (all $p>.05$). Error rates in the AX, AY, BX, and BY conditions were not significantly different across groups of learners (all $p>.05$). Participants who learned both languages showed significantly faster reaction times on AX, BX and BY trial types when compared to those who did not learn either language (all $p <.05$), and marginally faster reaction times on AY trials ($p = .067$), but showed no significant differences in error rates. When compared to the group who only learned the first language, participants who did not learn either language exhibited significantly slower reaction times only on AY ($p=.033$) and BX trials ($p = .049$), and marginally faster reaction times on AX trials ($p = .053$), but did not show any significant differences in error rates. See Figure 7 for detailed reaction times and error rates for all conditions.

A series of Bonferroni corrected correlations were run ($p < .0125$) , revealing a significant correlation between reaction times on BX trials and the number of languages learned during the experiment ($r(28) = -.485$, $p = .009$) such that reaction times on BX trials became faster for participants who learned more languages during the experiment. While they did not reach significance after the Bonferroni correction, the correlation between reaction times on AX trials and the number of languages learned during the experiment ($r(29) = -.390$, $p = .036$), and AY trials and the number of language learned ($r(29) = -.363$, $p = .053$) trended towards significance. Lastly, reaction times on BY trials were not significantly correlated with learning ($r(27) = -.322$, $p = .101$).

On the Operation Span Task, there were no significant differences in the scores between participants who learn both languages and those who are only able to learn the first. There were additionally no significant differences in the scores of participants who did not learn either language when compared to either of the previous groups (see Table 2-5).

Table 2-5. Mean proportion of errors and average reaction times for bilinguals depending on outcome during the Statistical Learning task.

	Bilinguals					
Learning	0		1		2	
Condition	Err (%)	RT (msec)	Err	RT (msec)	Err (%)	RT (msec)
AX	.14(.29)	401.22(84.29)	.07(.02)	332.34(64.84)	.04(.02)	328.09(59.31)
AY	.20(.28)	522.50(117.14)	.34(.23)	417.29(86.72)	.29(.26)	428.30(83.21)
BX	.16(.31)	342.76(108.67)	.10(.09)	250.56(80.21)	.03(.06)	230.83(52.73)
BY	.12(.33)	347.54(94.32)	.02(.04)	282.24(152.06)	.02(.05)	256.53(62.40)

Comparing monolinguals and bilinguals

While the difference did not reach statistical significance (Barnard test Wald statistic = .485, p =.435), a higher percentage of bilingual participants (45%) successfully learned the second language when compared to the monolingual participants (33%). Monolingual participants who only learned the first language learned it in 1.42 minutes (sd = .69) and bilingual participants who only learned the first language learned it in 1.56 (sd = .72), this difference did not reach statistical significance. Monolinguals who learned both languages learned the first in 2.17 (sd = 1.83) minutes and the second language in 2.67 (sd = 1.37) minutes, while bilinguals learned the first language in 2.25 minutes (sd = 1.39) and the second language in 2.12 minutes (sd = 1.35). When comparing the length of time it took to learn across monolinguals and bilinguals, none of the differences were significant (all p>.05).

When comparing the post-test data, however, we find that the monolingual participants who successfully learned both languages during the experiment did not retain the first language,, while the bilingual group did. These monolingual participants had an average accuracy on the L1 post-test of 9.0 (sd = 1.67) out of 16, which is not significantly above chance, while the bilingual participants accurately answered 9.66 (sd = 1.41) out of 16 on the same post-test, which is significantly above chance. While both groups retained the second language to significantly above chance levels, the small sample size does not allow for firm conclusions to be drawn.

The test of inhibitory control further revealed that the monolingual group exhibited significantly slower reaction times on the BX trials when compared to the bilingual group, but no additional differences in reaction times or error rates were found. While there was a significant correlation between reaction times on BX trials and learning on the statistical learning task for the bilingual group, reaction times on the AX-CPT task did not follow the same pattern for the monolingual group. No further differences between the groups were found, as they had already been matched on their working memory abilities (based on their score on the Operational Span task) and their proficiency in their native language (the English Verbal Fluency task).

Discussion

Experiment 5 revealed differences between matched monolinguals and second language learners on a statistical learning task where participants are presented with multiple artificial languages. Unlike previous research that reported no differences in the primacy effect between monolinguals and bilinguals (Bogulski, 2013), this experiment found emerging differences between the groups which result thus far in a larger percentage of the bilingual participants recruited hacing learned the second language presented during the experiment relative to the monolingual group. Thus far, the monolingual participants were not able to retain the first language learned when they successfully acquired the second, revealing an important difference between the two groups, since the bilinguals retained both learned languages to above chance levels. While these differences did not reach significance, that could be a function of the small sample size, and a larger number of participants in both groups will need to be recruited to truly understand these trends.

These results provide preliminary support for the prediction that participants who are actively learning a second language, and are thus engaged in segmenting a second language in their environment on a regular basis, might outperform monolinguals on a task where the underlying structure of the environment quickly undergoes change. Since participants were matched on a working memory task and a language proficiency measure, their difference in performance on the statistical learning task is not likely to be due to differences in other cognitive skills or differences in effort on the task. Notably, proficiency in the second language was not related to their performance on the Statistical Learning task, and the number of exemplars produced by the bilinguals in Spanish did not predict whether participants were able to learn the second language after learning the first.

Chapter 3

General discussion

Previous studies have demonstrated that when learners are presented with two artificial languages in succession, they only acquire the first language unless exposure to the second is tripled in length, or the switch to the second language is signaled by a contextual cue (such as explicit instructions regarding the presence of two languages; Gebhart et al., 2009). Here, I sought to investigate the conditions under which this primacy effect occurs as a means of better understanding its cognitive underpinnings. In Experiment 1, I presented learners with two languages in succession (the same ones used by Gebhart et al., 2009), but modified the method of familiarization by presenting learners with successive 1-minute blocks each followed by a test. Thus, the overall exposure was similar to the Gebhart study (5 minutes of one language followed by 5 minutes of a second language), but the intermittent tests provided a means of assessing the progression of learning across the session. The results of Experiment 1 replicated the pattern reported by Gebhart and colleagues in that learners were better able to acquire L1 than L2. Of those who learned L1 (as indexed by 7 out of 8 on one of the intermittent tests, a level of performance that was indicative of robust learning), 35% of participants also acquired L2. In Experiment 2, learners advanced to L2 immediately after reaching this threshold. After learning L1, 67% went on to successfully learn L2, a significant increase relative to Experiment 1. Experiment 3 was identical to Experiment 2 except for the addition of contextual cues to signal a change in structure, as they have been demonstrated to impact learners' ability to acquire multiple inputs (Gebhart, Newport, & Aslin, 2009; Weiss, Gerfen, & Mitchel, 2009; Weiss & Mitchel, 2010; Poepsel, Gerfen, & Weiss, 2012). Experiment 3 utilized three contextual cues, including pitch-shifting the voice of one of the languages, explicitly notifying participants that they would be learning two languages, and a combination of these cues. In each of these conditions, the percentage of participants who successfully acquired the second language did not increase relative to Experiment 2 (47% for the Pitch-Shift Condition, 57% in the Explicit-Cue Condition, and 56% in the

Combined Condition), demonstrating that contextual cues may not be critically important for overcoming the primacy effect when participants advance to L2 immediately after they learn L1.

In a second series of experiments, I investigated whether bilinguals may differ on this task as a consequence of their experience maintaining multiple languages (e.g., Marian & Spivey, 2003; Kroll, Bobb, & Wodniecka, 2006). In Experiment 4, highly proficient Spanish-English bilinguals were tested using the same paradigm as Experiment 2 in which learners advanced immediately after reaching criteria on L1. While only 30% of the bilingual participants were able to acquire L2, these results were difficult to interpret in light of the small sample size and the heterogeneity of the participants' age of acquisition. Therefore, in Experiment 5, I recruited a homogenous group of advanced Spanish learners who were actively engaged in learning their second language and compared them with a group of matched monolinguals. Notably, a higher percentage of bilingual participants (45%) were able to learn the second language relative to the matched monolingual group (30%), a difference that trended toward significance. Thus, in contrast to a previous study which tested the sample population (Bogulski, 2013), I found that actively engaging in learning a second language might impact the ability of participants to segment consecutive languages, though a larger sample is necessary to draw a firmer conclusion (and note that the matched monolinguals also need a larger sample to determine the discrepancy between their performance and those recruited in Experiments 1-3 from the subject pool). Overall, the results of the experiments presented in this thesis were informative both with respect to better understanding the primacy effect, as well as providing some suggestive evidence that second language learning might interact with detecting changes in statistical learning.

One of the central findings in these experiments is that manipulating the amount of exposure to L1 after learning has occurred can attenuate the primacy effect reported in previous research (Gebhart et al., 2009) and in Experiment 1. By advancing learners to L2 immediately after learning L1, a significantly higher percentage of participants successfully acquired both languages. This insight was facilitated by testing participants after every one-minute block of exposure, thereby providing at least a rough estimate of the time course of learning. Surprisingly, in both Experiments 1 and 2, participants acquired L1 in less than 2 minutes (on average). Since participants in Experiment 1 received 5 minutes of exposure regardless of their performance on the test, they received as many as 4 additional blocks of exposure to the same set of statistics after learning. However, the critical difference between Experiments 1 and 2 is that in Experiment 2 participants were advanced to L2 at the end of the block during which learning reached a high threshold, and therefore they received minimal additional exposure after learning. Thus, primacy effects in this statistical learning task may be critically tied to the amount of exposure participants received after learning; additional exposure to the same set of statistics after learning appears to entrench learners in the statistics of L1. This finding is complementary to a recent study by Zinszer & Weiss (2013) who reported that the primacy effect could be overcome by providing participants with multiple switches between the languages. Participants were presented with the same amount of exposure as Gebhart et al., but broken down into alternating 2.5-minute blocks of each of the two languages. Unlike the Gebhart study, with this alternating presentation participants were able to learn both languages. The authors surmised that the ability to overcome the primacy effect was due to either the increased number of switches or to the initially shorter amount of exposure to the first language compared to the original experiment. Consequently, they lengthened the first block and found that switching alone could help learners overcome the primacy effect. The results of the present study are complementary in that they suggest that reducing the amount of exposure after learning can in fact impact the primacy effect as well. Taken together, these results indicate that the initial bias toward learning the first of two structures can be overcome by increasing the variability in the input, and that the timing of the variability also plays a role in the observed entrenchment.

In the initial study, Gebhart and colleagues (2009) suggested that the primacy effect might be due to a capacity limit in the number of structural representations an individual is able to learn. Thus, participants might choose to only learn the first language, since deleting the first representation to learn the second might not prove to be a useful strategy. If the first representation was deleted and then turned out to be necessary, relearning it might be more difficult. Additionally, the first presented structure might be the most important, so maintaining that representation may be more useful for the learner in the long term. The results of Experiments 1 and 2, however, find that both representations can be learned and retained. Despite participants learning L1 in under 2 minutes and being advanced to L2 immediately after learning, they were able to retain both languages when tested retrospectively at the post-test. The fact that they were able to retain the first language, despite having received as many as 5 minutes of exposure to the second language, signifies that they did not delete their representation of the first language in order to learn the second.

Qian and colleagues (2012) suggested that the primacy effect might have arisen as a consequence of participants having a high level of estimation uncertainty at the point of switching to L2. In other words, participants may never reach a high level of confidence in their representation of the input statistics, and therefore tend to underadjust as they update their model of the input. This is consistent with the notion of anchoring, the idea that a learner has a bias to rely on the information encountered first in order to make future decisions, as discussed by Kahneman & Tversky (1974; see below). The data presented here, however, suggest a plausible alternative explanation. Performance on the intermittent tests of L1 suggests that learners could achieve and sustain high levels of learning, which is more consistent with a low level of estimation uncertainty. Based on the results of Experiments 1 and 2, it is possible that once learners have formed a relatively robust representation of the statistical input, they may decrease their sampling of the input. According to this interpretation, in Experiment 1, the observed entrenchment effect arises directly as a consequence of the additional exposure to L1 after learning. In Experiment 2, participants still learn early during exposure, but then receive minimal additional exposure after learning.

This new interpretation is actually quite consistent with the data presented by Gebhart and colleagues (2009). In their experiment, all participants received five minutes of exposure to the first language and if they also learned the underlying structure of the language early during exposure, the result would be similar to that found in Experiment 1. However, when provided with explicit knowledge of a second structure along with a pause between languages (Gebhart et al., 2009), attention is drawn to the contrast between the two languages and attention may have been reengaged. Likewise, the notion that learners reduce their sampling is also consistent with the observation that when L2 was lengthened in duration, participants successfully acquired both structures. That is, given the reduction in sampling, learners therefore required more time to accrue enough information to notice the underlying structure of the statistics had undergone a change. Of course, it is also possible that sampling decays over time irrespective of the extent to which learning has occurred (i.e, that sampling decays regardless of performance). However, the results of Zinszer & Weiss (2013) temper this interpretation somewhat, as they demonstrated that successful learning of a new language could occur even after longer periods of exposure than presented here. Nonetheless, future experiments need to more precisely investigate how learners engage in sampling with methods that allow for more direct measurements.

Another finding from this series of experiments is that adding contextual cues did not further reduce the presence of the primacy effect when learners advanced immediately after acquiring L1. Several previous experiments have reported that contextual cues correlated with a change in structure facilitate learning. For example, a change in speaker voice (Weiss, Gerfen, & Mitchel, 2009, Poepsel, Gerfen, & Weiss, 2012), explicit knowledge that multiple structures are present in the environment (Gebhart et al., 2009; Poepsel, Gerfen, & Weiss, 2012), and a visual indicator of a change of identity (Mitchel & Weiss, 2010) facilitated learners in forming multiple representations to accommodate two input languages. In Experiment 3, the contextual cues provided for the learner did not result in a greater number of participants learning the second language. In fact, in the Pitch-Shift Condition, a significantly smaller percentage of participants learned the second language. In the Explicit-Cue Condition and the Combined Condition, the percentage of participants who learned the second language did not differ from Experiment 2. These results may suggest that the role of contextual cues in learning multiple structures is to overtly signal the change to learners, which, in turn, may trigger increased sampling of the input statistics. In the paradigm used in Experiments 2 and 3 in which participants did not receive additional exposure to the first language after learning, learners appear to detect the change in structure and consequently there may be no additive value to signaling the change has occurred. Notably, the Pitch-Shift Condition resulted in a significantly smaller percentage of participants learning the second language despite the fact that when the language was tested independently it was learned significantly above chance. Gebhart et al., (2009) also found that pitch shifting one of the languages was insufficient for participants to overcome the primacy effect. Thus, this cue, which is meant to approximate a change in voice, appears to function differently than actual changes in voice, as has been tested with other paradigms (e.g., Weiss, et al., 2009; Poepsel & Weiss, 2012).

Another question addressed in these experiments was whether differences in inhibitory control abilities could predict participants' success in learning in a non-stationary environment. The hypothesis was that learners may need to inhibit the first set of statistics in order to learn the second, and therefore it is possible that successful learning of both languages could correlate with performance on tasks that measure inhibitory control. In Experiments 1-3, participants performed the Flanker task in addition to the statistical learning task. Notably, inhibitory control abilities were only predictive of participants' ability to learn in Experiment 2; learners who acquired both languages exhibited significantly faster reaction times relative to participants who only learned L1. While the traditional Flanker effect was not found in this case, participants' ability to learn in a non-stationary environment may be related to their speed of processing as demonstrated by their faster reaction times when they had to juggle multiple trial types.

In the second series of experiments, I investigated how experience managing multiple languages (e.g., Marian & Spivey, 2003; Kroll, Bobb, & Wodniecka, 2006) impacts participants' ability to learn when presented with multiple artificial languages Previous research comparing monolinguals and bilinguals on statistical learning tasks has not reached consensus on how bilingualism impacts statistical learning. One study that compared monolingual and bilingual children on both visual and auditory statistical learning found no differences between the two groups (Yim & Rudoy, 2013). In contrast, some evidence suggests that bilingual infants more readily acquire multiple rules (Kovács & Mehler, 2009a,b) and that adult late second language learners more readily acquire multiple mappings in a cross-situational word learning task (Poepsel & Weiss, in review). Most similar to the studies conducted here, Bogulski (2013) compared monolinguals and a group of advanced English-Spanish learners similar to the one recruited in these experiments on an extended version of the Gebhart et al., (2009) task and found that both groups exhibited the same primacy effect. The findings from Experiment 4 suggested that bilingual participants might differ in their ability to learn the second language relative to the monolinguals recruited in Experiment 2 (adopting the same methodology as used in Experiment 2).

Consequently, Experiment 5 compared a homogenous group of advanced English-Spanish learners to a group of matched monolinguals on the same procedure from Experiment 2. Thus far, a higher percentage of bilingual participants learn the second language relative to the matched monolingual group, although this difference is not significant as the sample size is too small to draw any firm conclusions. Remarkably, the paid monolinguals recruited for this experiment performed very differently when compared to the monolinguals from Experiment 2. In Experiment 5, the paid monolinguals exhibited the original primacy effect found in the literature, with only 33% of them successfully learning the second language (contrary to the 67% who were able to acquire both languages in Experiment 2). While these paid participants were a little older than the participants recruited through the Subject Pool for the original studies, they should not have been inherently different in their statistical learning abilities, and thus these differences in performance when comparing the monolinguals in Experiments 2 and 3 to Experiment 5 are not easy to explain and require further investigation.

The monolingual and bilingual group had been previously matched on both working memory and verbal fluency in their native language so that any differences found between the groups would not be impacted by individual differences in these abilities. While reaction times and accuracy on the AX-CPT task did not reliably differ between the two groups (contrary to both Morales et al., (2012) and Bogulski (2013), there was a significant correlation between reaction times on BX (proactive control) trials and performance on the statistical learning task for bilinguals but not for monolinguals. Since reaction times on the other types of trials were also marginally correlated with performance on the task, these results indicate that the bilingual group relied on their inhibitory control mechanisms differently than their monolingual counterparts.

These findings accord with theories of bilingual advantages in cognitive control (Bialystok, Craik, Green, & Gollan, 2010), stating that actively speaking multiple languages has a systematic impact on the individual's cognitive functioning. Since both languages of a bilingual are active at all times (e.g. Marian & Spivey, 2003, see Bialystok et al., 2010 for a review), bilinguals constantly need to inhibit one language to speak the target language, which may result in increased executive function abilities over time. Further these results support a more recent theory stating that bilingualism may change the way an individual attends to their environment (Bialystok, 2015). Bialystok argues that the bilingual advantage may not be due solely to experience juggling multiple language, but more directly due to the individual having learned to attend to contrasts in their input to differentiate between two languages (Bialystok, 2015). Since the monolingual and advanced second language learners were matched on working memory and native language proficiency, and there was no consistent relation between performance on the AX-CPT and statistical learning tasks, their differences in performance on the statistical learning task are likely due to factors besides experience juggling multiple languages. The greater percentage of bilingual participants who successfully acquired the second language might be indicative of a change in the way attention is paid to contrasts in the input: causing the switch to the second language to appear more salient for the bilingual group compared to the monolingual group. Future research will need to individually test the components of bilingual differences in executive function to understand how each contributes to learning in an environment that quickly undergoes change.

While the small sample size is a limiting factor for interpreting these findings, the results trend toward differences between the two groups of participants, unlike the findings reported by Bogulski (2013) with a similar group of participants. Bogulski (2013) compared monolinguals to late second language learners on a similar task in which participants received 10 minutes of exposure to each of the same languages used in the experiments discussed above, but found that bilinguals and monolinguals did not differ on their performance. These results illustrate that retrospective tests of learning, such as those used by previous experiments, do not provide a sufficiently fine grained measure to understand differences between monolinguals and bilinguals on these tasks. By investigating the time course of learning, and allowing participants to advance as soon as they learned, this paradigm allows for more informative comparisons across the two groups.

Primacy effects are not specific to statistical learning tasks, but have also been observed in other domains. The anchoring effect, as described by Tversky & Kahneman (1974), is an individual's tendency to adjust estimates of the environment on the basis of the first available information. Anchoring has been shown to reliably influence decision making in a variety of domains (Tversky & Kahneman, 1974) and to be unaffected by motivational effects. In contingency learning, the acquisition of explicit or implicit relations between stimuli or responses, participants also exhibit a primacy effect. For example, Marsh & Ahn (2006) found that the primacy effect in contingency learning is reliant on an anchoring hypothesis of the relation between the stimuli, formed by participants early during exposure to the stimuli, that is then difficult to overcome even with additional information. However, they argue that the primacy effect could be overcome if learners are explicitly asked to state the relation between the stimuli more frequently throughout exposure, possibly as it allows them to reappraise the distribution of the input. Notably, the experiments presented here tested participants every minute, and while participants still exhibited a primacy effect in that L1 was learned significantly better than L2, performance on L2 was slightly above chance (which was not the case in the Gebhart study). In Experiment 1, when participants received additional exposure to the first language after learning, more consistent testing did not completely overcome the primacy effect, though it appears to have improved performance on the second language.

Sensitive periods may also serve as real life examples of these primacy effects. Knudsen and colleagues (1998) have shown that the barn own acquires specific relationships between sound and visual cues during an early sensitive period. During this specific period, these juvenile owls can also adjust to a switch in this relationship, and can recover the initial relationship at a later point in their life if they were exposed to it during the sensitive period. Knudsen (1998) has argued that the original exposure to a specific relationship left a representational trace that could then be relearned later in life. Similarly, studies have shown that traces for languages overheard during infancy and childhood can allow for easier learning of that language during adulthood (Au, Knightly, Jun, & Oh, 2001). Results from both Knudsen et al., (1998) and Au et al. (2001) demonstrate that initially learned representations can be retained, but that early exposure to these representations is critical, as they might not have been maintained had exposure to them not begun until after learning a different set of rules.

As presenting participants with multiple artificial languages in their input can reflect the type of environment encountered in the course of learning multiple languages, the following section addresses bilingual language acquisition. Bilingual language acquisition can happen simultaneously, where a child is exposed to multiple languages from as early as birth, or sequentially, where an individual masters their first language before becoming exposed to a second language through immigration to another country or in a school setting. Researchers have found that late second language learners are less likely to attain native-like proficiency in their L2 when compared to early second language learners or simultaneous bilinguals (Johnson and Newport, 1998), with transfer occurring from the first to the second language (MacWhinney, 2004). While age of acquisition has often been proposed as an explanation for this difference in learning outcomes, another theory that has been proposed is the entrenchment hypothesis(Monner, Vatz, Morini, Hwang, & DeKEYSER, 2012). During an entrenched state, previous knowledge (about the linguistic environment) is resistant to change, and may only be altered slowly, which might interfere with the acquisition of new information in the environment (see Monner, Vatz, Morini, Hwang, & Dekeyser, 2012 for further explanation). The results from this thesis seem to provide support for this hypothesis from a set of reductionist experiments. Without having the chance to become entrenched, participants are less resistant to the possibility that change has occurred, and thus can acquire both of the artificial languages presented to them during the experiment.

Limitations and future directions

One limitation of these studies is that the way the experiments were designed does not allow for a direct measure of how participants are sampling from their environment. While the results demonstrate that advancing participants to the second language as soon as learning occurs attenuates the presence of a primacy effect, the current studies cannot determine the exact reason for this. The attenuation of the primacy effect could be due to two related reasons: learners might continue to sample regularly because less time has passed before exposure to the second language begins, or because their expectation that the environment is stationary has not been reinforced by additional exposure to the first set of statistics.

With regards to the monolingual/bilingual comparison, the results so far cannot speak to definite differences between the two groups as the relatively small sample size does not allow for firm conclusions to be drawn. Another limitation of Experiment 5 in particular is that the paid monolinguals recruited for this experiment did not behave like the ones in Experiment 2. While the paid monolinguals should not have been inherently different in their statistical learning skills, they were recruited from a research volunteer website and may have been more invested in the research. Since statistical learning has been shown to be a passive, incidental task which can be achieved even when concurrently performing a secondary task (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), it is possible that paying too much attention to the speech stream and to the task itself may compromise one's ability to perform on the task. While research has shown that some amount of attention is necessary to successfully segment a speech stream (Toro, Sinnett, & Soto-Faraco, 2005), to my knowledge no research to date has investigated whether too much attention can actually hinder this process. Further research is necessary to understand why these differences in performance emerged when participants were recruited from a volunteer site compared to the Psychology Subject Pool. However, given that the bilingual participants were also recruited and compensated $10 for their participation, the differences found when comparing the matched groups may still be indicative of a change in the underlying mechanism as a result of learning a second language.

Besides recruiting more advanced second language learners and monolinguals, future directions also include testing a group of simultaneous or early bilinguals on the paradigm used in Experiments 2 and 5. Since acquiring a second language later in life seems to improve one's ability to learn in a non-stationary environment, having first hand experience learning multiple languages from an early age may also impact performance on a multiple-stream statistical learning task. Of future interest would also be to investigate how explicitly cuing bilingual participants (both late second language learners and early bilinguals) that they would be learning both languages might change their performance on the task. It is possible that learning a second language in an explicit environment (the classroom setting for example) would lead to better performance when provided with explicit cues, relative to a group of early bilinguals who only received implicit cues of a change in the underlying structure.

A final limitation that will also require future experiments, is the inability to determine which components of executive function predict successful learning of multiple structures during the statistical learning task. While inhibitory control, tested both using the Flanker task and the AX-CPT task was related to performance, the current data is not sufficient to fully understand the role of inhibitory control, and executive function more generally, in statistical learning.

References

Aslin, R. N. (2014). Infant Learning: Historical, Conceptual, and Methodological Challenges. *Infancy*, *19*(1), 2–27. doi:10.1111/infa.12036

Au, T. K., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2001). Overhearing a language during childhood. *Psychological Science*, *12*(5), 360–366. doi:10.1044/2014

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Bialystok, E. (2015). Bilingualism and the Development of Executive Function: The Role of Attention. *Child Development Perspectives*, *9*(2), 117–121. doi:10.1111/cdep.12116

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2010). Bilingual Minds, *10*(3), 89–129. doi:10.1177/1529100610387084

Blanco, F., Baeyens, F., & Beckers, T. (2014). Blocking in human causal learning is affected by outcome assumptions manipulated through causal structure. *Learning & Behavior*, *42*(2), 185–99. doi:10.3758/s13420-014-0137-y

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. doi:10.1016/j.tics.2011.12.010

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–32. doi:10.1016/j.cognition.2011.06.010

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science*, *17*(10), 905–12. doi:10.1111/j.1467-9280.2006.01801.x

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458–467. doi:10.1037//0278-7393.28.3.458

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*(7), 1243–52. doi:10.1177/0956797612472207

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009a). Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science*, *33*(6), 1087–1116. doi:10.1111/j.1551-6709.2009.01041.x

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009b). Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science*, *33*(6), 1087–1116. doi:10.1111/j.1551-6709.2009.01041.x

Gollan, T. H., Montoya, R. I., & Werner, G. a. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, *16*(4), 562–576. doi:10.1037/0894-4105.16.4.562

Gonzales, K., Gerken, L., & Gómez, R. L. (2015). Does hearing two dialects at different times help infants learn dialect-specific rules? *Cognition*, *140*, 60–71. doi:10.1016/j.cognition.2015.03.015

Houston, D. M., & Jusczyk, P. W. (2000). The Role of Talker-Specific Information in Word Segmentation by Infants, *26*(5), 1570–1582. doi:10.10371/0096-1523.26.5

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), B35–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11869728

Kovács, A. M., & Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science (New York, N.Y.)*, *325*(5940), 611–2. doi:10.1126/science.1173947

Kovacs, M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants.

Kroll, J. F., & Dussias, P. E. (2012). The Comprehension of Words and Sentences in Two Languages. *The Handbook of Bilingualism and Multilingualism: Second Edition*, 216–243. doi:10.1002/9781118332382.ch9

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2014). Bilingualism, Mind, and Brain. *Annual Review of Linguistics*, *1*(1), 140822192156002. doi:10.1146/annurev-linguist-030514-124937

Lany, J., & Saffran, J. R. (2013). *Statistical Learning Mechanisms in Infancy*. *Neural Circuit Development and Function in the Heathy and Diseased Brain* (Vol. #volume#). Elsevier. doi:10.1016/B978-0-12-397267-5.00034-0

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29–41. doi:10.1016/j.cognition.2009.08.014

MacWhinney, B. (2004). A unified model of language acquisition. *Linguistics and Second Language Acquisition.*

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI? *Journal of Child Language*, *41*(2), 327–51. doi:10.1017/S0305000912000736

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. doi:10.1017/S1366728903001068

Marsh, J. K., & Ahn, W. (2006). Order effects in contingency learning : The role of task complexity, *06511*(3), 568–576.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, *11*(1), 122–34. doi:10.1111/j.1467-7687.2007.00653.x

Mitchel, A. D., & Weiss, D. J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, *25*(4), 456–482. doi:10.1080/01690960903209888

Monner, D., Vatz, K., Morini, G., Hwang, S.-O., & DeKEYSER, R. (2012). A neural network model of the effects of entrenchment and memory development on grammatical gender learning. *Bilingualism: Language and Cognition*, *16*(02), 246–265. doi:10.1017/S1366728912000454

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, *25*(5), 531–546. doi:10.1080/20445911.2013.807812

Poepse, T. J., & Weiss, D. J. (2014). Context influences conscious appraisal of cross situational statistical learning. *Frontiers in Psychology*, *5*(JUL), 1–9. doi:10.3389/fpsyg.2014.00691

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism Modulates Infants' Selective Attention to the Mouth of a Talking Face. *Psychological Science*. doi:10.1177/0956797614568320

Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to represent a multi-context environment: more than detecting changes. *Frontiers in Psychology*, *3*(July), 228. doi:10.3389/fpsyg.2012.00228

Reeder, P. a, Newport, E. L., & Aslin, R. N. (2013). From shared contexts to syntactic categories: the role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, *66*(1), 30–54. doi:10.1016/j.cogpsych.2012.09.001

Saffran, J. R. (2003). Statistical language learning: mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110–114. doi:10.1111/1467-8721.01243

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults, *70*, 27–52.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. a., & Barrueco, S. (1997). INCIDENTAL LANGUAGE LEARNING:. Listening (and Learning) out of the Comer of Your Ear. *Psychological Science*, *8*(2), 101–105. doi:10.1111/j.1467-9280.1997.tb00690.x

Sebastián-gallés, N., Albareda-castellot, B., Weikum, W. M., & Werker, J. F. (2012). A Bilingual Advantage in Visual Language Discrimination in Infancy. doi:10.1177/0956797612436817

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–68. doi:10.1016/j.cognition.2007.06.010

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), 25–34. doi:10.1016/j.cognition.2005.01.006

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty : heuristics and biases. *Science*, *185*(4157), 1124–1131.

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science (New York, N.Y.)*, *316*(5828), 1159. doi:10.1126/science.1137686

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech Segmentation in a Simulated Bilingual Environment: a Challenge for Statistical Learning? *Language Learning and Development : The Official Journal of the Society for Language Development*, *5*(1), 30–49. doi:10.1080/15475440802340101

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7238150

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. doi:10.1016/S0163-6383(84)80022-3

Wilson, R. C., & Niv, Y. (2011). Inferring relevance in a changing world. *Frontiers in Human Neuroscience*, *5*(January), 189. doi:10.3389/fnhum.2011.00189

Yim, D., & Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech Language and Hearing Research JSLHR*, *56*(1), 310–22. doi:10.1044/1092-4388(2012/11-0243)b

Appendix

Flanker reaction times and accuracy in Experiment 3

	Condition 1: Pitch Shift			**Condition 2: Explicit Knowledge**			**Condition 3: Pitch Shift + Explicit Knowledge + Pause**		
	Neither	**L1**	**L1 + L2**	**Neither**	**L1**	**L1 + L2**	**Neither**	**L1**	**L1 + L2**
Reaction time (ms)	493.13 (108.25)	531.92 (144.3)	499.59 (86.51)	501.50 (73.39)	510.35 (49.58)	497.95 (52.02)	523.17 (67.57)	516.35 (79.20)	496.20 (99.32)
Accuracy (%)	90.74 (5.1)	81.76 (15.2)	88.2 (7.4)	73.34 (12.8)	78.23 (4.0)	73.31 (12.3)	68.91 (17.31)	72.19 (15.4)	75.26 (9.3)