How Accurately Do Infants Represent Lexical Stress Information in Recently Segmented Words?

A Thesis Presented for the

Master of Arts

Degree

The University of Tennessee, Knoxville

Sara Parvanezadeh Esfahani

May 2019

Abstract

Eight-month-old monolingual English learning infants are able to use co-occurrence statistics to find words in continuous artificial (e.g., Saffran, Aslin, & Newport, 1996) and natural languages (Pelucchi, Hay, & Saffran, 2009). Although these findings have been replicated numerous times, we still know very little about how these newly extracted words are represented. For example, if infants use TP information to segment a word with a trochaic (strong/weak) stress pattern in speech, will they recognize the same newly encountered word if it is presented with an iambic (weak/strong) stress pattern? Building on work by Pelucchi et al. (2009), infants were familiarized with Italian sentences that had two embedded high transitional probability (HTP; TP=1.0) trochaic target words (e.g., *FUga* & *MElo*) – their syllables never occurred anywhere else in the corpus. Following familiarization, infants were tested using the head-turn preference procedure on their ability to discriminate HTP words from two novel words (e.g., *PAne* & *TEma*) that had never occurred in the corpus. In a counterbalanced language the HTP and novel words were switched. In Control condition, the trochaic stress pattern of the target words was consistent across familiarization and test, while in the Experimental condition, the stress pattern of the words was changed between familiarization and test, such that if the HTP words in the corpus were trochaic (e.g., *FUga* and *MElo*), infants were tested on their ability to discriminate the iambic version of the target words (e.g. *fuGA* & *meLO*) from novel iambic words (e.g., *paNE* & *teMA*). Across conditions infants listened significantly longer to HTP words compared to Novel words, suggesting that infants' representation of stress pattern in newly encountered words is not robust yet. These findings suggest that segmental information may override suprasegmental information at this age.

Table of Contents

Chapter 1

Introduction

To use language as a means of communication, children need to develop robust lexical representations. Lexical representations refer to everything that a child knows about a word, from which sounds make up to a word (phonology), to what a word means (semantics) and how it is used in sentences (morpho-syntax). From the phonological perspective, words can be represented at three different levels: indexical (i.e., properties of the speaker's voice), segmental (i.e., phonotactic rules of language), and suprasegmental (i.e., prosody of language). At the indexical level, it is adaptive for infants to learn to ignore variation in words because who says a given word does not alter its meaning. Segmental changes are lexically contrastive across the world's languages (e.g., cat → bat) and thus, it is adaptive to learn to attend to variation at the segmental level. While suprasegmental information plays a vital role in word meaning in some languages (e.g., tone languages where varying the pitch contour can alter the word's meaning), not all languages use suprasegmental information this way. Indeed, English-learning infants sometimes need to pay attention to suprasegmental cues (e.g., to differentiate between the noun record /ˈrekərd/ and the verb record /rəˈkôrd/) and sometimes need to ignore them (e.g., when someone makes a pronunciation error). Although the fact that people rarely make stress errors during speech production (Cutler & Isard, 1980) suggests that suprasegmental information is represented on some level, it is unclear the extent to which it is adaptive for English-learning infants to represent suprasegmental information at the lexical level as stress does not typically differentiate word meanings.

Although a significant body of research suggests that infants home in on the relevant features of their native language relatively early on (see Werker & Gervain, 2013 for a review;

Bosch & Sebastián-Gallés, 2003; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984) and provide a basis for their lexical representations of familiar words, much less is known about what infants encode about words when they first encounter them in fluent speech. There is a great deal of evidence to suggest that infants are remarkably good at identifying structured information embedded in fluent speech – a process that has been referred to as statistical learning (Saffran, Aslin, & Newport, 1996).

Despite our knowledge of infants' statistical learning ability and the development of the lexical representation of familiar words, we know very little about how newly segmented novel words from continuous speech are represented. There is evidence that later in development, these recently extracted novel words that have no meaning or reference associations can make better object label candidates (Graf Estes, Evans, Alibali, & Saffran, 2007; Hay, Pelucchi, Graf-Estes, & Saffran, 2011), but it is not clear how they are represented at different phonological levels: indexical, segmental, and suprasegmental. In the present study, we aim to investigate how recently segmented words are represented at the suprasegmental level. Before going into details about the current study, we first review relevant findings on (1) the status of infants' lexical representations and (2) statistical learning in infancy.

Lexical Representation in Infancy

Phonologically, words can be represented at the indexical, segmental, and suprasegmental levels. Although 7.5-month-olds appear to have a difficult time recognizing target words when some features of the voice changed such as gender (Houston & Jusczyk, 2000), affect (happy vs neutral, Singh, Morgan, & White, 2004), and pitch (Singh, White, & Morgan, 2008b), as infants gain experience with their language they should come to ignore features that do not alter the meaning of a word (e.g., 'cat' produced by a male and 'cat' produced by a female voice mean the same thing). At two months, infants are sensitive to syllable changes across different speakers (Jusczyk, Pisoni, & Mullennix, 1992) whereas six-month-olds can ignore talker variability in vowel discrimination (Kuhl, 1979; 1983). Indeed, in a study investigating the robustness of statistically segmented words across acoustic variations at the indexical level (e.g., change of speaker from familiarization to testing) Graf Estes (2012) found that 11-month-old infants are able to recognize words recently segmented from an artificial language when the speaker changed from familiarization to test.

At the segmental level, however, which sounds make up a word is essential to word meaning – a change in either a consonant or a vowel can change the meaning of the entire word. From birth, infants show a remarkable ability to discriminate many, if not most, of the sounds from across the world's languages (see Werker & Tees, 1999 for a review; Aslin, Pisoni, Hennessy, & Perey, 1981; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Werker & Gervain, 2013). However, as they gain more experience, they attenuate this broad sensitivity to their native language sounds sometime between 6 to 10 months (for a review, see Maurer & Werker, 2014), and demonstrate improved ability to discriminate some difficult native contrasts (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005). Given this early perceptual narrowing, we might expect that even young infants would be sensitive to changes at the segmental level. Because 12 months is often thought of as the onset of word learning (although see Burgelson & Swingley, 2012), we do not know very much about the segmental specificity of young infants' lexical representation. Interestingly, work by Werker and colleagues suggest that at 14 months, infants ignore segmental changes in minimal pair novel word learning (Stager & Werker, 1997). However, if the words are familiar (Fennell & Werker, 2003), 14-month-olds appear to be sensitive to these segmental features of words. Future work is needed to determine the degree of segmental specificity of infants' early lexical representations.

In contrast to indexical features, which infants should learn to ignore, and segmental features, which infants should learn to attend, suprasegmental features fall in a grey zone for infants learning intonation languages, such as English. Suprasegmental features are prosodic cues such as lexical tone, pitch contour or stress pattern. In tonal languages (e.g., Mandarin Chinese) lexical tones or pitch contours are used to contrast meaning. However, in intonation languages (e.g., English) stress pattern plays an important role in the speech signal. Stressed syllables typically have a longer duration, and higher pitch and amplitude (Hayes, 1995). Although changing the stress pattern of words in English can sometimes change their meaning (e.g., record /ˈrekərd/ as a noun and record /rəˈkôrd/ as a verb), by and large, suprasegmental features are not relevant to word meaning in English.

Nonetheless, from birth, infants begin to respond to native language prosodic information (Mehler et al., 1998), and by two months, they are sensitive to stress pattern changes (Jusczyk & Thompson, 1978). While 9-month-olds prefer listening to words that are consistent with the dominant stress pattern in their native language, neither 6-month-olds (Jusczyk, Cutler, & Redanz, 1993), nor 7-month-olds (Echols, Crowhurst, & Childers, 1997) have shown such a preference, suggesting that perceptual narrowing of prosodic features as a result of more experience with language happens sometime between 6 to 9 months of age.

Research on whether stress information is encoded in lexical representations or whether it is stored separately during infancy does not yield conclusive findings. One camp of researchers (e.g., Chomsky & Halle, 1968; Halle, 1998) do not place stress as part of the lexical representation. They argue that lexical stress is stored separately and assigned according to a set of rules. In support of this account, Vihman and colleagues (Vihman, Nakai, DePaolis, & Halle, 2004) explored the role of prosodic and segmental features of lexical representation in infants. They tested both 9- and 11-month-olds on their ability to differentiate familiar words (e.g., baby) from more rare words (e.g., bridle). Across conditions, words were either produced correctly or they were mispronounced at the segmental level (e.g., bunny → vunny) or at the suprasegmental level (e.g., BAby → baBY). Nine-month-olds failed to differentiate familiar from novel words even when the words were correctly pronounced, suggesting rather weak lexical representations at this age. In contrast, 11-month-olds listened longer to familiar words than rare words both when the words were correctly pronounced and when there was a change in stress pattern. However, they failed to differentiate familiar words from rare words when there was a segmental change. These findings suggest that by 11 months, infants have a relatively robust mental representation of familiar words at the segmental level that may allow infants to disregard stress change (Vihman et al., 2004).

In contrast, Cutler and colleagues (Cutler, 1979; Cutler & Isard, 1980) argue that stress information is integrated into lexical representations for familiar words. They support their account by providing evidence from lexical stress errors. Cutler and Isard (1980) discuss that stress errors occur because same stem words are stored together, and when people access their lexicon, confusions cause pronunciation errors. However, they could not provide similar evidence for unfamiliar words (Cutler & Isard, 1980). Thus, there remains ambiguity about if and when in the process of learning words, stress information becomes integrated into the representation.

Curtin and colleagues (Curtin, Mintz, & Christiansen, 2005) found that English-learning 7- and 9-month-olds are sensitive to stress cues during speech segmentation. Specifically, after listening to a naturally produced artificial language that contained both stressed and unstressed versions of the same syllables, English-learning infants preferentially segmented words that had initial stress (i.e., SWW) over words that had medial (i.e., WSW) or final stress (i.e., WWS). After demonstrating that stress information can be used to posit word boundaries, Curtin and colleagues (Curtin et al., 2005) explored whether stress information is represented in recently encountered words. They familiarized a group of 7-month-olds with the same artificial familiarization language but tested them with English sentences that contained target words from previous experiments that were stressed either on the initial syllable or the middle syllable (e.g., '*I like your DObita option*') and control words that did not occur in the familiarization. Infants showed sensitivity and listened significantly longer to the sentences containing initially stressed words extracted from continuous speech, suggesting that infants preferred a certain parsing strategy (trochaic bias). Furthermore, infants maintained stress information in lexical representation after word segmentation. These findings suggest that stressed syllables are represented differently than their unstressed counterparts. And this difference facilitates infants' ability to use transitional probability information to segment words from a fluent speech stream by helping infants posit word boundaries. These findings suggest that infants may represent stress information even in newly encountered sound sequences (Curtin et al., 2005).

Statistical Learning

Findings from previous research show that infants are remarkable statistical learners (Saffran, Aslin, & Newport, 1996). By their first birthday, infants demonstrate sensitivity to language-general (e.g., Saffran et al., 1996; Aslin, Saffran, & Newport, 1998) as well as language-specific cues (e.g., Johnson & Seidl, 2009; Jusczyk et al., 1993; Jusczyk, Houston, Newsome, 1999; Hay & Saffran, 2012) during word segmentation. One language-general cue to word segmentation is transitional probability between syllables, computed as the frequency of XY (a syllable sequence) given the frequency of the syllable X (Swingley, 2005). Using an artificial language, that contained no other cues to word boundaries than differences in transitional probability within versus between words, Saffran and colleagues (e.g., Saffran et al., 1996; Aslin, Saffran, & Newport, 1998) found that infants can track transitional probabilities in order to locate words in continuous stream of speech. These findings have been replicated numerous times across many different labs (for a review, see Krogh, Vlach, & Johnson, 2013; Romberg & Saffran, 2010).

One critique of artificial language studies is that the languages themselves are artificially simplistic and do not represent the complexity of natural language input. In order to address this concern Pelucchi, Hay, and Saffran (2009a) set out to test statistical learning in complex natural language input. They familiarized 8-month-old English-learning infants with naturally produced Italian sentences that had four embedded trochaic target words. Two of the words were high TP (HTP) because their syllables never occurred anywhere else in the corpus (TP =1.0). Two of the words were low TP (LTP) because their first syllable occurred in many other words throughout the corpus (TP =.33). Following familiarization, infants were tested using the head-turn preference procedure on their ability to discriminate HTP words from LTP words. Infants preferred listening to the relatively familiar HTP words, even though both HTP and LTP words were heard an equal number of times in the speech stream, suggesting that 8-month-olds can also use co-occurrence statistics to extract words from natural language input (Pelucchi et al., 2009a).

Although Pelucchi and colleagues' findings (2009a) demonstrated that infants are able to track transitional probability information even in natural language input, it is important to note that TP was not the only word boundary information available in the familiarization corpus. Like

English, Italian has a predominantly trochaic (strong/weak) stress pattern – and thus, all of the target words used by Pelucchi and colleagues had that same stress pattern. According to a corpus analysis by Svartvik & Quirk (1980), stressed syllables occur at word onset in 75 percent of the words in English. Further, 7.5-month-old English learners can use syllable stress information to segment trochees from a fluent speech stream (Jusczyk et al., 1999). Thus, the predominant exposure to trochaic words makes English-learning infants attend to trochees more than non-trochees, and could have provided redundant cues to word boundaries that facilitated segmentation.

There has been a fair amount of debate in the literature about whether infants rely more heavily on statistical cues versus suprasegmental cues to word boundaries. In one study, Johnson and Jusczyk (2001) pitted stress cues against statistical cues (i.e., transitional probability between consecutive syllables and stress pattern of the words indicated conflicting cues to word boundaries). They showed that 8-month-old infants rely more heavily on stress cues to find the word boundaries, suggesting that stress cues carry more weight than statistical cues during word segmentation. In another study using artificial language materials, Thiessen and Saffran (2003) familiarized English-learning infants with either a trochaic or an iambic speech stream. In the trochaic language, the initial syllables of the statistically defined disyllabic words were stressed. Thus, for these English-learning infants, the stress pattern and statistical cues provided consistent cues to word boundaries. However, in the iambic language, the second (or final) syllables of statistically defined disyllabic words were stressed, leading to stress pattern and statistical cues that provided conflicting information about word boundaries. Both 7- and 9-month-old English learning monolingual infants successfully segmented words from trochaic language (i.e., distinguished words from part-words at test). However, in the iambic language condition, while

7-month-old infants performed similarly to previous groups and continued to rely on statistical cues to extract words, 9-month-olds paid greater attention to stress cues and mis-segmented words (i.e., they considered the stressed syllable as the onset of the words pulling out part-words instead of words). Thiessen and Saffran (2007) demonstrated that a brief exposure to iambic words results in the successful segmentation of iambs (i.e., iambic words) even in 7-month-old infants. Eleven-month-olds also demonstrated a preference for rhythmic compared to statistical information (Johnson and Seidl, 2009). These results indicate the significance of linguistic experience in learning language-specific cues such as stress in identifying word boundaries (Hay & Saffran, 2012).

Although Pelucchi and colleagues (Pelucchi et al., 2009a) have demonstrated that infants have powerful computational abilities, we still know very little about how these newly extracted words are represented. Follow up work with 17-month-olds using both artificial languages (Graf Estes et al., 2007) and natural languages (Hay et al., 2011) has found that words with strong internal TP make better object labels than those with weaker internal statistics, suggesting that statistical learning may lead to the extraction of candidate object labels. Nevertheless, how infants represent the indexical (i.e., talker information), suprasegmental (i.e., stress patterns), and segmental (i.e., individual sounds in words) features of these newly extracted statistically-defined words remains unknown. In the present study, we investigate the representation of statistically defined words at the suprasegmental level.

Current Study

The literature discussed here provides evidence that in the second half of their first year, infants use statistical cues to pull out words from continuous speech (Karaman & Hay, 2018; Pelucchi et al., 2009). Further, as infants' experience with language increases, their word segmentation ability may shift from relying primarily on statistical cues to using stress patterns as the more salient cue to rely on (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003). The literature on lexical representation also yields contradictory findings as to whether infants encode word stress in lexical representation (Curtin et al., 2005; Vihman et al., 2004). Thus, the main aim of the present study was to investigate how infants represent statistically defined words at the suprasegmental level.

There are few, if any, studies looking at the representation of the stress pattern in recently segmented words in young infants, and most studies have only examined infants' ability to segment over familiar or unfamiliar stress pattern (e.g., Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003; Nazzi, Iakimova, Bertoncini, Fredonie, & Alcantara, 2006). In the present study, however, using a natural foreign stream of speech, we explore whether 8-month-old infants' lexical representation of recently segmented words contains stress pattern information.

Sometime between 6 and 9 months, infants begin to perceptually orient toward only their native language prosodic pattern and lose sensitivity to non-native prosody (Echols et al., 1997). Also, since we aimed to look at the specificity of the stress pattern in the representation of newly segmented words using the natural language stimuli from Pelucchi et al. (2009), we decided to stick to the same age range that they tested in order to first establish a control based on the replication condition and then approach our experimental condition.

In a between-subject design, all infants were exposed with a series of Italian sentences in which two disyllabic target words with high transitional probability (HTP; TP = 1) were embedded (i.e., their syllables did not appear anywhere else in the corpus). In the control condition, infants were immediately tested on four words, two HTP words plus two novel words (Italian words that were not presented during familiarization, nor were their syllables). Here, we expected infants to listen longer to the HTP words than to the novel words, replicating the findings of Pelucchi and colleagues (2009a). In the experimental condition, infants were familiarized with the same language and then were immediately tested on the same four words produced as iambs instead of as trochees. If infants treat the stressed and unstressed syllable as functionally equivalent, then they should continue to prefer listening to modified HTP words at the test. However, if infants integrate stress into their lexical representation of newly segmented words, they should fail to recognize the HTP words at test, and thus fail to differentiate them from the novel words. This pattern of results would suggest that 8-month-olds have already formed a quite robust representation of the stress pattern in newly segmented words.

We chose to test infants in a between-subject design (i.e., HTP words vs Novel words with either trochaic or iambic stress pattern) rather than testing all infants in a single condition with trochaic vs iambic HTP words. If we tested all infants with trochaic vs iambic HTP words version, the results would not allow us to understand whether infants' preference to listen to trochaic HTP words were due to English learning infants' potential trochaic bias or lack of specificity of stress pattern in the representation of recently segmented words.

Chapter 2

Materials and Methods

Participants

Forty-eight 8-month-old infants (23 females) with a mean age of 8.6 months participated in this study. All infants were full-term with no record of hearing and/or vision problems and were recruited from monolingual English speaking families with no consistent exposure to another language especially Italian or Spanish. Participants were recruited from the Child Development Research Group database based in the Department of Psychology at the University of Tennessee, Knoxville.

In both conditions, infants were randomly assigned to one of the two counterbalanced languages (see appendix for the list of sentences). Thirty other infants participated in the study but were excluded from analysis due to experimental error (n = 3), experimenter error (n = 3), fussiness (n = 11), crying (n = 7), lack of attention to the stimuli (n = 2), and outliers (± 2 SD, n= 4). All parents gave informed consent before the experiment. All infants received a t-shirt as a gift for their participation.

Stimuli

Speech materials were taken from Pelucchi et al. (2009a, Experiment 1). A new female native speaker of Italian who was naïve to the purpose of the study recorded the stimuli with a lively manner. The familiarization phase language composed of 12 syntactically correct and semantically meaningful sentences that were repeated 3 times to create a speech stream of about 2 minutes and 20 seconds. All sentences were normalized using Praat software to an average of 77 dB_{SLP}.

Four disyllabic Italian words (*fuga, melo, pane,* and *tema*) were also recorded to use in testing phase. The syllables of the target words were all phonotactically legal in English. In each of the two counterbalanced languages, two of these words were embedded in a way that the internal TP of the words were 1.0 (HTP words) while the other two words were used as novel words. For example, in Language 1, *fuga* and *melo* were HTP words, and *pane* and *tema* were novel words. In Language 2, *pane* and *tema* were HTP words, and *fuga* and *melo* were novel words. All the test words were recorded with two different stress patterns. The trochaic stress pattern maintained the stress pattern of the HTP words in familiarization phase and was used for the control condition. For the experimental condition the target words were produced with an iambic stress pattern. So, for example if the target words in the corpus were *FUga* and *MElo* (see Table 1 for the acoustic characteristics of each token in the familiarization languages), the test words were pronounced as *fuGA* and *meLO* (see appendix for the acoustic characteristics of test tokens in control condition (Table 2), and experimental condition (Table 3)). During the test, the stress pattern of the novel words always matched the stress pattern of the target HTP words. All the words were edited in Praat to have an equal length of 500 ms and intensity of approximately 77 dB_{SLP}.

Procedure

There were two phases: a familiarization phase and a testing phase. Infants were exposed to one of the two counterbalanced languages in the familiarization phase and then were tested using Head Turn Preference Procedure (HTPP). The study was conducted in a soundproof booth that had a center screen, two side screens plus two side audio speakers. During the study, the caregiver held the infant on their lap while listening to a masking music to lower any possible bias. The experimenter observed the infant’s head turns in another room over a closed circuit camera. In Familiarization phase, a video of a spinning pinwheel appeared on the screens depending on the infant's looking behavior while the language was played continuously.

Right after familiarization, infants heard 12 test trials. All infants were presented with the same items (trochaic tokens in the control condition and iambic tokens in the experimental condition) regardless of the counterbalanced languages. Each of the test words (i.e., two familiar HTP words and two novel words) were repeated three times, randomized by block. Test trials began with the appearance of the spinning pinwheel on the center screen. When the infant oriented to the center screen, the experimenter initiated the trial which transferred the pinwheel from the center to one of the side screens. As soon as the infant made a head turn of at least 30° to the side screen showing the pinwheel, one of the test trials played constantly until the infant looked away for at least 2 seconds or 15 seconds had passed. The next trial began with the appearance of the pinwheel on the center screen and continued for all 12 trials. Therefore, in this procedure, the infant basically controlled the looking time or hearing the target words. Trials which lasted for less than one second were repeated at the end. The dependent variable was the infant's total looking time toward each test trial. After the experiment, the caregiver also filled out a demographic information survey and a Communicative Development Inventory (CDI).

Chapter 3

Results

A four-way mixed-model ANOVA was conducted to examine whether there were any main effects of familiarization language or participant sex on their listening time preferences for HTP versus Novel words. There were no main effects of interactions involving familiarization language, $F(1, 47) = .189$, $p = .666$, partial $\eta2=.004$, power=.071, or sex, $F(1, 47) = .090$, $p = .766$, partial $\eta2=.002$, power=.060, thus all subsequent analyses were collapsed across these variables.

A 2 Condition (control vs experimental) × 2 Word Type (HTP vs Novel) repeated measures ANOVA revealed a significant main effect of word type, $F(1,47) =14.681$, $p=.000$, partial $\eta2=.242$, power=.963, suggesting that overall, infants preferred listening to the HTP words relative to the Novel Words (see Figure 1). The main effect of condition was only marginally significant, $F(1, 47) = 3.644$, $P=.063$, partial $\eta2=.073$, power=.464, with infants showing a trend towards looking longer on the test trials where the words were trochees (i.e., control condition) than on trials where the words were iambs (i.e., experimental condition) (see Figure 2). These results are consistent with previous findings that suggest that infants prefer listening to trochaic words (in English; (Jusczyk et al., 1993; Echols et al., 1997) or words consistent with the predominant stress pattern (Polka, Sundara, & Blue, 2002). Importantly, there were no Word Type and Condition interactions, suggesting that infants did not perform differently in the control versus the experimental conditions.

Consistent with previous research (Karaman & Hay, 2018; Pelucchi et al., 2009a), planned comparisons revealed that infants in the control condition, $t(23)= 3.192$, $p= .004$, $d=$

.648, listened significantly longer on HTP (Mean= 10.08 s, SD =2.11) than on Novel word (Mean =9.52 s, SD = 2.02) test trials (see Figure 3). This suggests that the infants were successful at pulling the HTP words out of the speech stream and had a familiarity preference for listening to these recently segmented words at test. A second set of planned comparisons revealed that the same looking time pattern emerged in the experimental condition; infants listened significantly longer to modified iambic HTP words (Mean= 9.11 s, SD= 1.95) than to the iambic Novel words (Mean= 8.36 s, SD= 1.97), $t(23)= 2.561$, $p=.017$, $d=.522$. These results suggest that 8-month-old infants treat recently segmented trochaic words and their iambic version as functionally equivalent. Although they can segment an unfamiliar natural fluent speech, they may have not yet had a strong representation of stress pattern, and their lexical representations of newly encountered words are still fragile.

The difference scores between HTP and Novel words showed greater variations in the difference scores in the experimental condition compared to the control condition, suggesting a greater variation in the performance of infants in the experimental condition (see Figure 4).

Chapter 4

General Discussion and Conclusions

Discussion

In the present study, we assessed whether 8-month-old infants' immediate representation of words segmented from continuous speech contains stress information. My prediction was that if infants have already formed a robust representation of suprasegmental information and treat stressed vs unstressed same segment syllables differently, they should fail to discriminate between modified HTP words and Novel words in the experimental condition (iambic test words). However, if their lexical representation of suprasegmental information is still rather fragile, they should treat iambic test words similarly and continue to differentiate between target words and listen to modified HTP words longer than novel words (i.e., same results as control condition). Here, we provided evidence in support of the latter. Eight-month-old infants treated the modified HTP words functionally equal to the original HTP words and listened to them longer compared to the novel words.

The results reported here demonstrate that the stress pattern of newly encountered words does not have a strong lexical representation. These results are consistent with Vihman and colleagues' findings with familiar words that mispronunciation at the suprasegmental level do not block lexical representation at 11 months because segmental information overrides suprasegmental features, allowing infants to ignore changes in stress pattern (Vihman et al., 2004).

However, the results did not illustrate Cutler and colleagues' findings with familiar words that suggest that stress pattern of familiar words is included in lexical representation (Cutler, 1979; Cutler & Isard, 1980). But we also did not use familiar words. Further, Curtin and colleagues (Curtin et al., 2005) have demonstrated that at 7 and 9 months, infants use stress cues to segment trisyllabic words from a continuous speech in an artificial language, and concluded that infants encode stress information in their lexical representations.

Werker and Curtin's (2005) PRIMIR (Processing Rich Information from Multidimensional Interactive Representation) framework may shed light on the discrepancy in findings across studies on suprasegmental representations. Based on this account, lexical representations are accessed through three filters: initial biases, infants' developmental level, and task demands. The last two filters (developmental level of the infants and task requirements) can modulate access to linguistic information such as syllable stress or statistics in word segmentation. For example, in Thiessen and Saffran (2003), 7-month-olds relied more heavily on statistical cues whereas 9-month-olds used stress cues to segment language. Thus, in segmenting words recently encountered, since the representations are not as robust as those of familiar words, stress information might not be as important and become ignored.

Another similar explanation is supported by Word Recognition and Phonetic Structure Acquisition (WRAPSA) account proposed by Jusczyk (1993; 1997). According to this framework, auditory analyzers reweight representations based on the exemplar-based models. When infants encounter a familiarization language, they form an average representation across tokens, and they can recognize words when representations match traces of words heard in the familiarization (Jusczyk, 1993). In our study, infants listened to a novel language for only about 2 minutes, and we used a natural speech stream in which tokens' acoustic characteristics for each HTP word are not exactly identical (because it is a natural language). Therefore, infants may have not yet formed strong prosodic traces in their long term memory so they disregard the stress change in the test words.

Study Limitations

In the present study, we aimed to explore the specificity of lexical representation in newly segmented words at the suprasegmental level. We used a natural speech stream to increase the ecological validity compared to the previous studies that used artificial languages (e.g., Curtin, et al., 2005). However, as one potential limitation, testing infants in the lab setting where all the natural environment sounds are reduced would affect the extent to which our findings are comparable to real-life situations. Further, in the familiarization languages, HTP words had perfect TP which is not always the case in natural speech. Thus, there should be more research on probing how segmented words with lower TP are represented and if less strong statistics affects lexical representation in newly segmented words.

Future Directions

In order to have a better understanding of the development of the lexical representation of newly encountered words, I would be interested to see when infants start forming a robust representation of suprasegmental information. Thus, testing for example 11-month-olds may provide an understanding of the developmental changes of acoustic sensitivity to words extracted from continuous speech.

In the present study I only probed how suprasegmental information is represented in recently segmented words. In the next step, I will investigate whether the lexical representation of newly segmented words contains indexical (properties of a voice such as the speaker’s sex) and segmental (phonemic makeup of the words) information.

Conclusion

The findings of the present study provided evidence that 8-month-old infants can segment an unfamiliar natural speech stream, and a change in stress pattern of the target words does not block their word recognition. These findings demonstrate that infants at eight months probably have not yet formed a strong representation of stress pattern for the novel words they just encountered. The findings are also supported by language acquisition frameworks such as PRIMIR and WRAPSA.

List of References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321-324.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, *52*(4), 1135.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(2-3), 217-243.

Chomsky, N., & Halle, M. (1968). The sound pattern of English. New York: Harper and Row xiv, p. 470. Reprinted 1991, Boston: MIT Press.

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233-262.

Cutler, A. (1979). Errors of stress and intonation. In V. A. Fromkin (Ed.), Errors in linguistic performance: Slips of the Tongue, Ear, Pen, and Hand (pp. 167–180). New York: Academic Press.

Cutler, A., & Isard, S. D. (1980). The production of prosody. In *Language production* (pp. 245-269). Academic Press.

Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, *36*(2), 202-225.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*(2-3), 245-264.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, *4*, 1490.

Graf Estes, K. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology*, *3*, 447.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254-260.

Halle, M. (1998). The stress of English words 1968–1998. *Linguistic Inquiry*, *29*(4), 539-568.

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93-106.

Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(6), 610-641.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548-567.

Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental Science*, *12*(1), 131-141.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: The Massachusetts Institute of Technology Press.

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675-687.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3-4), 159-207.

Jusczyl, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, *43*(3), 253-291.

Jusczyk, P. W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics*, *23*(2), 105-109.

Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(2), 221.

Krogh, L., Vlach, H., & Johnson, S. P. (2013). Statistical learning across development: flexible yet constrained. *Frontiers in Psychology*, *3*, 598.

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, *6*(2-3), 263-285.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *The Journal of the Acoustical Society of America*, *66*(6), 1668-1679.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the" critical period". *Language Learning and Development*, *1*(3-4), 237-264.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, *56*(2), 154-178.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283-299.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674-685.

Polka, L., Sundara, M., & Blue, S. (2002). The role of language experience in word segmentation: A comparison of English, French, and bilingual infants. The 143rd Meeting of the Acoustical Society of America: Special Session in Memory of Peter Jusczyk, Pittsburgh, PA.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906-914.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*(2), 173-189.

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, *4*(2), 157-178.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381.

Svartvik, J., Quirk, R. (1980). A Corpus of English Conversation. Lund., CWK Gleerup.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, *8*(5), 432-443.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology*, *39*(4), 706.

Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, *3*(1), 73-100.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, *50*(3), 336-353.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197-234.

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.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, *50*(1), 509-535.

Appendices

Table 1. Acoustic characteristics of syllables of tokens in the familiarization languages.

Target Words in Familiarization Languages	1st syllable (Stressed)			2nd syllable		
	Duration (s)	Intensity (dB)	Pitch (Hz)	Duration (s)	Intensity (dB)	Pitch (Hz)
fuga	0.291	64.12	354.67	0.124	71.12	226.82
	0.173	60.36	334.14	0.118	67.99	370.12
	0.305	68.97	271.56	0.100	73.13	257.78
	0.208	62.58	354.55	0.099	73.21	354.68
	0.341	69.52	429.89	0.209	75.74	271.71
	0.285	67.43	633.66	0.139	73.54	261.12
Average	**0.267**	**65.49**	**396.41**	**0.131**	**72.45**	**290.37**
melo	0.167	68.55	214.84	0.096	68.39	212.17
	0.136	71.10	230.24	0.110	71.63	226.45
	0.202	68.14	205.08	0.134	67.30	237.93
	0.168	68.53	220.94	0.100	67.69	195.47
	0.126	65.05	191.78	0.112	63.86	193.69
	0.258	64.46	199.22	0.152	65.47	233.49
Average	**0.176**	**67.63**	**210.35**	**0.117**	**67.39**	**216.53**
pane	0.124	74.80	206.26	0.100	64.15	195.74
	0.117	74.62	325.32	0124	65.42	190.76
	0.210	72.86	158.80	0.149	61.60	208.70
	0.120	73.85	220.83	0.120	63.82	270.01
	0.136	74.07	462.12	0.136	64.14	236.30
	0.213	74.93	278.45	0.144	63.27	247.26
Average	**0.153**	**74.18**	**275.29**	**0.128**	**63.73**	**224.79**
tema	0.232	72.55	313.43	0.176	66.19	290.56
	0.157	73.16	219.55	0.113	70.45	196.81
	0.203	72.53	157.30	0.151	71.07	189.30
	0.091	70.59	301.27	0.132	71.64	322.32
	0.244	70.35	192.58	0.235	70.03	363.15
	0.222	74.44	272.95	0.195	69.90	227.13
Average	**0.191**	**72.27**	**242.84**	**0.167**	**69.88**	**264.87**

Table 2. Acoustic characteristics of syllables of test tokens in the control condition

Control Condition	1st syllable (stressed)			2nd syllable		
	Duration (s)	Intensity (dB)	Pitch (Hz)	Duration (s)	Intensity (dB)	Pitch (Hz)
FUga	0.316	77.46	334.98	0.186	78.09	178.94
MElo	0.272	78.77	368.66	0.230	75.73	196.22
PAne	0.271	79.41	318.75	0.225	67.63	201.15
TEma	0.253	78.93	304.09	0.246	74.74	205.40

Table 3: Acoustic characteristics of syllables of test tokens in the experimental condition

Experimental Condition	1st syllable			2nd syllable (stressed)		
	Duration (s)	Intensity (dB)	Pitch (Hz)	Duration (s)	Intensity (dB)	Pitch (Hz)
fuGA	0.224	64.05	211.51	0.275	80.21	256.64
meLO	0.232	69.98	216.34	0.266	80.00	260.55
paNE	0.166	73.92	238.66	0.334	78.37	275.20
teMA	0.174	73.07	239.31	0.323	78.79	262.62

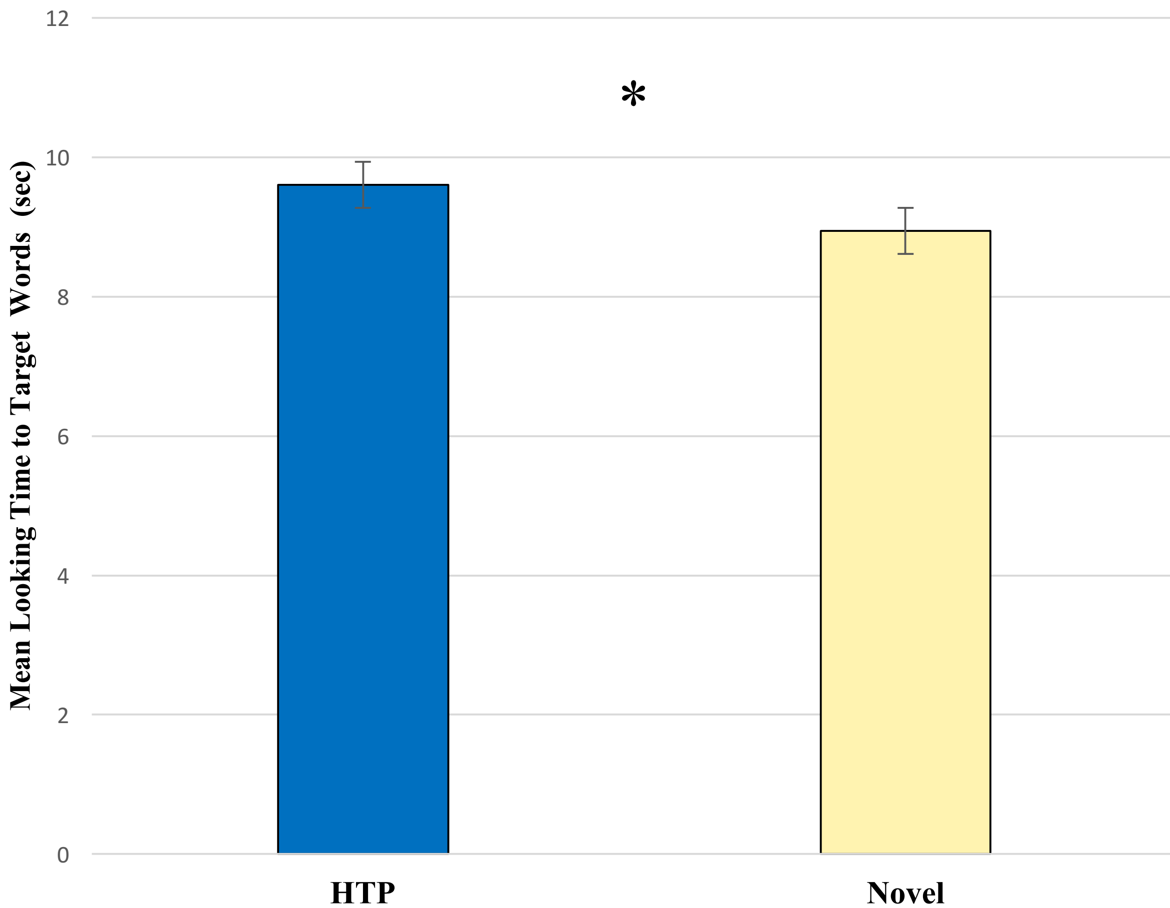

Figure 1. Mean looking time (sec) to HTP and Novel words across conditions. Error bars represented the standard error of the mean. The star (*) denotes statistically significant looking time difference to target words at p =.000.

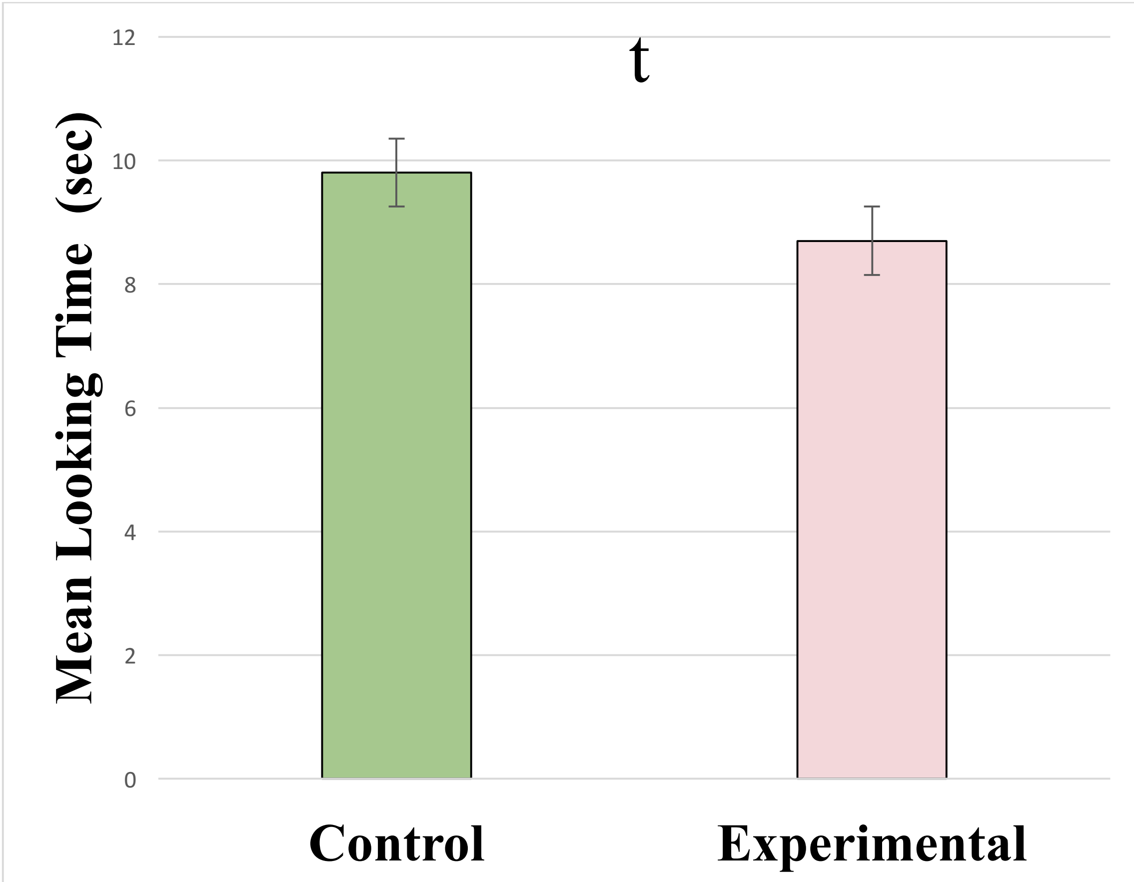

Figure 2. Mean looking time (sec) to target words in Control (trochaic tokens) and Experimental (iambic tokens) conditions. Error bars represented the standard error of the mean. (t) denotes marginally significant difference between the conditions at p =.063.

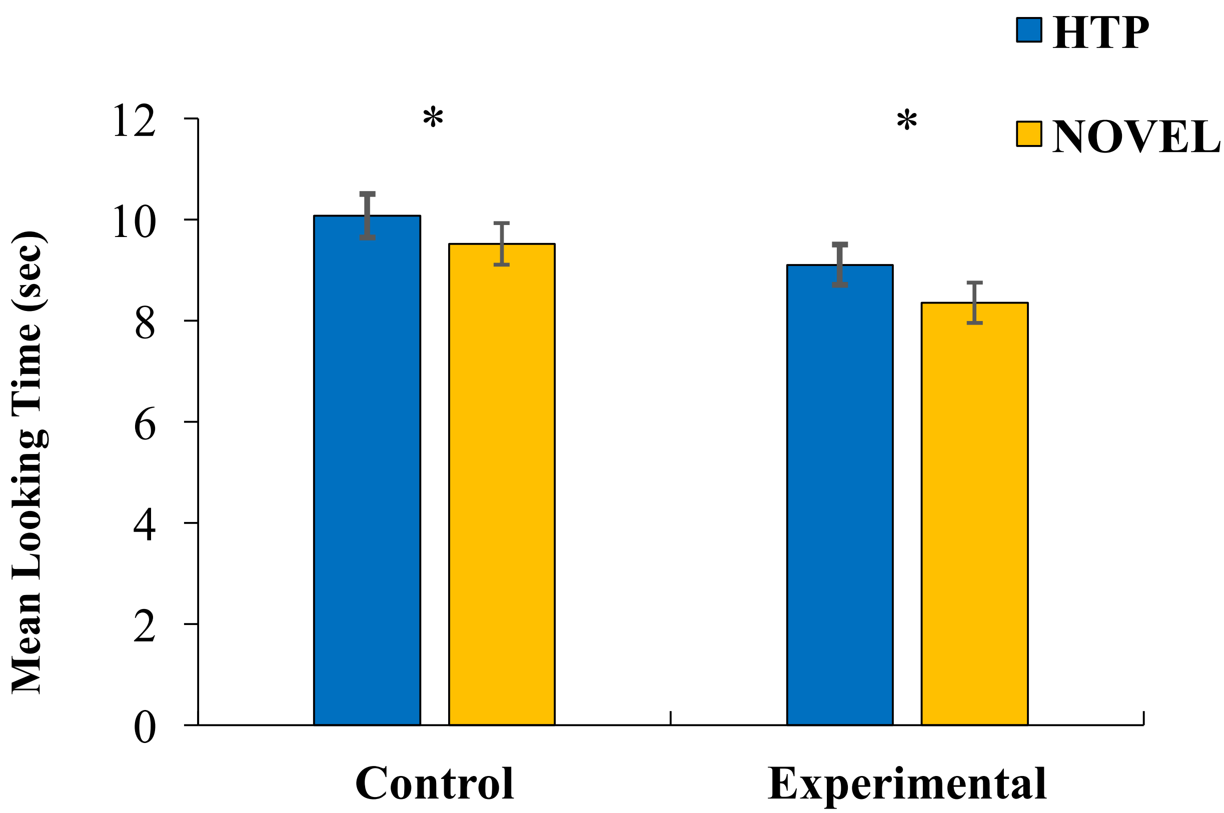

Figure 3. Mean looking time (sec) to HTP and Novel words in Control (trochaic tokens) and Experimental (iambic tokens) conditions. Error bars represented the standard error of the mean. The stars (*) denote significant pairwise comparisons at $p = .01$

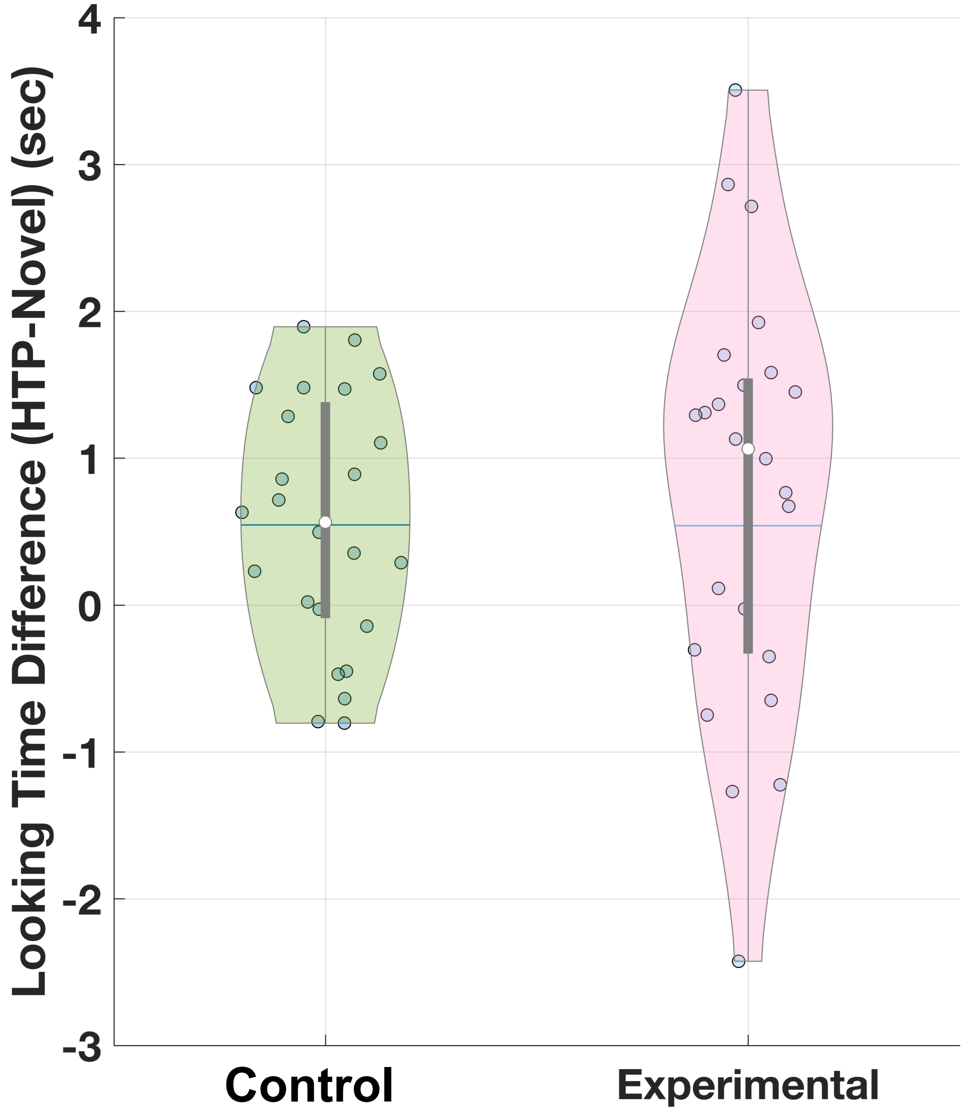

Figure 4. Looking time difference between HTP and Novel words by participant in Control and Experimental conditions.

Language 1

HTP words: fuga & melo Novel words: pane & tema

Torno a casa con le bici cariche di frutta in bilico sulla sella.

La zia Carola si e` esibita in una fuga colla bici verde.

Se porti il melo sulla bici forse cali un po' di chili.

La bici ha subito un danno dentro la casa del capo di Lara.

La cavia Bida e` in fuga da casa per aver giocato con le bilie blu.

La biscia in lenta fuga dal giardino capita in casa mia.

Il tuo melo arcano fuga l'afa che debilita la folla.

Arriviamo in bici fino al bivio del grande melo con un caro amico.

Il picchio si abitua a fare la sua casa in ogni melo cavo e alto.

Gusto i bigoli dentro casa o coricata all'ombra del melo verde.

Di rado una bici in rapida fuga rincorre la moto bigia e rossa.

Per ascoltare la fuga quasi cadi sul melo e inciampi sulla biro sull'erba.

Language 2

HTP words: pane & tema Novel words: fuga & melo

Torno a casa con le bici cariche di frutta in bilico sulla sella.

La zia Carola si e` esibita in una tema colla bici verde.

Se porti il pane sulla bici forse cali un po' di chili.

La bici ha subito un danno dentro la casa del capo di Lara.

La cavia Bida e` in tema da casa per aver giocato con le bilie blu.

La biscia in lenta tema dal giardino capita in casa mia.

Il tuo pane arcano tema l'afa che debilita la folla.

Arriviamo in bici fino al bivio del grande pane con un caro amico.

Il picchio si abitua a fare la sua casa in ogni pane cavo e alto.

Gusto i bigoli dentro casa o coricata all'ombra del pane verde.

Di rado una bici in rapida tema rincorre la moto bigia e rossa.

Per ascoltare la tema quasi cadi sul pane e inciampi sulla biro sull'erba.

Vita

Sara Parvanezadeh Esfahani is currently a fourth year doctoral student in Experimental Psychology at the University of Tennessee, Knoxville. Sara works in Infant Language and Perceptual Learning Lab under the supervision of Dr. Jessica Hay. Prior to joining Dr. Hay's lab, Sara received a Bachelor of Art degree in English Translation from Chamran University, Ahvaz, Iran and Master of Art degree in Teaching English as a Foreign Language from University of Tehran, Tehran, Iran. Sara will receive her second Master of Art degree shortly and continue to work with Dr. Hay through her PhD. Sara is currently a graduate teaching associate and research assistant. Her research interests focus on how specifically young infants represent sound sequences in natural languages.