Bilingual and Monolingual Infants' Recognition of Words Spoken in Foreign Accents

A Senior Honors Thesis for the Department of Psychology

Brendan R. Dillon

Tufts University

Author Note

Brendan R. Dillon, Department of Psychology, Tufts University, Class of 2013.

The author would like to thank Dr. Emily W. Bushnell for her incredible guidance and support as well as the Tufts University Department of Psychology for providing the resources necessary to conduct this study. I would also like to thank Shea Maloney and Elise Ressa for their support as research assistants as well as Dr. Katherine J. Midgley for sitting on this thesis committee.

Correspondence concerning this paper should be addressed to Brendan R. Dillon, Department of Psychology, Tufts University, 490 Boston Ave. Medford, MA 02155.
E-mail: Brendan.Dillon@tufts.edu

Abstract

Lexical acquisition is one of the most challenging tasks human infants face. Once they are able to separate and identify word forms from fluent speech and map those words onto their meanings, infants must also decipher how much perceptual change in speech is indicative of a phonetic and thus potential meaning change. The current study investigated the effect of foreign accents on recognition of familiar words in bilingual and monolingual 16 to 24-month-olds. Infants were presented with a pair of visual stimuli and asked to find one of them by a word spoken in either a familiar American English accent or a foreign Mandarin Chinese or Mexican Spanish accent. Successful performance on the task was indicated by a significantly greater persistence to the target picture if initially fixated when the word was said in comparison to persistence to the nontarget picture if initially fixated when the word was said. Within the monolingual group, infants displayed a strong understanding of the target words spoken in all three of the tested accents. Exposure to a foreign accent did not provide any additional benefit to later recognition of words spoken in that accent. Furthermore, age and word production both affected performance on the task, such that both older and more advanced talkers displayed stronger performance. Finally, bilingual infants also demonstrated a clear understanding of the target words, as measured through the persistence difference to the target versus nontarget pictures. As in the monolingual group, the bilingual infants understood all three accents equally well. However, there was some suggestion that the bilinguals showed weaker performance on the task overall. In summary, the early word learning infants in this study demonstrated the remarkable ability to understand words spoken in both familiar and foreign accents, even without previous exposure.

Keywords: lexical acquisition, word recognition, accented speech, bilingualism

Bilingual and Monolingual Infants' Recognition of Words Spoken in Foreign Accents

Language acquisition represents one of the most challenging tasks human infants face. On a daily basis, infants are confronted with a continuous stream of speech that they must decipher and connect to the separate continuous stream of perceptual-cognitive information around them. Traditionally, language acquisition involves the perception and production of speech, the acquisition of vocabulary and syntactic rules, and the development of certain pragmatic social skills required for successful communication. Out of all these fascinating components of language acquisition, the present study focuses on lexical acquisition and the processes behind which infants begin to create and then expand their set of word forms. Before they can even begin to understand a sentence as a whole and with meaning, infants must learn each word individually. Word learning involves at least three central challenges. The first challenge of segmentation requires that infants correctly separate and identify word forms out of the continuous stream of speech they hear. In the second challenge of mapping, infants must then associate each newly learned word form with its correct referent and learn the appropriate boundaries for extending it to more than one object or concept, such as the word dog referring to not only the first dog an infant may see and learn about but all newly encountered dogs thereafter. The third central challenge to word learning and the focus of this paper is specificity and how much perceptual change, such as a different accent, pitch, or volume, means a potential phonetic and meaning change for early word learning infants.

In order to learn a new word, infants must first connect the word to the relevant physical or even non-physical entity, separate the relevant sounds of the word from the stream of fluent speech, and develop a successful mapping between the two (Waxman & Lidz, 2006). Despite the inherent difficulties of this task, word production begins around 12 months of age. Furthermore, recent research has demonstrated that even 6-month-olds are able to understand some concrete words, such as “apple” or “shoe,” in gaze directed tasks (Bergelson & Swingley, 2012), and they soon acquire the ability to understand some abstract words, such as “eat” and “uh-oh,” by 10 months of age (Bergelson & Swingley, 2013). Such findings stand in contrast with previous consensus that infants must first learn the features of speech before they can understand words, and word learning is only possible when infants begin to understand a speaker’s communicative intent around 9 to 10.5 months of age (Waxman & Gelman, 2009; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998). Consequently, the beginning of language acquisition may be marked by the concurrent learning of both words as well as the unique sounds making up those words.

Nonetheless, the learning of different word forms is certainly associated with knowing the sound structure of the language. While newborns demonstrate a remarkable ability to discriminate a majority of sound contrasts found in the world’s languages, there is a decline in this ability for nonnative contrasts and a concurrent strengthening of this ability for native contrasts by 10 months of age (Gervain & Mehler, 2010). This perceptual narrowing of discrimination for universal phonemes to only those relevant to an infant’s native language appears to take place earlier for vowels than for consonants (Polka & Werker, 1994). Consequently, beginning at 6 months of age and by the end of their first year of life, infants display a working knowledge of the consonant and vowel phonemes of their native language based on their relevant experience (Werker & Tees, 1984; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992).

Language acquisition in infants depends on both learning the syntactic code of their native language in addition to the building of a lexicon (Gervain & Mehler, 2010). Given the continuous nature of speech, acquiring the knowledge of different word forms can be arduous, and this segmentation task can present a challenge to word learning. It isn't until about 6 to 8 months of age that infants begin to extract a set of learned word forms from speech (Gervain & Mehler, 2010). This segmentation is likely based on statistical information present in language, as 8-month-olds are able to segment a stream of syllables into word boundaries based on transition probabilities (the probability of one syllable predicting the latter or former one), in addition to cues specific to language, such as stress patterns, phonotactics, and allophonic variation (Saffran, Aslin, & Newport, 1996; Swingley, 2005). For example, 7.5-month-olds were found to display a trochaic bias and segment speech at strong-stress syllables. While these infants successfully segmented trochaic, multisyllabic words, they were not able to do so for iambic bisyllables, where the stress is on the last syllable, nor for strong monosyllables (Jusczyk, Houston, & Newsome, 1999). In addition, phonotactic trends, in which some phonemes tend to be located at the beginning while others in the middle of words, are employed by infants in the segmentation of presented sequences by 10.5 months of age (Mattys, Jusczyk, Luce, & Morgan, 1999). Allophonic cues, which describe variations of the same phoneme, can be successfully used alone and without distributional cues for segmentation at 10.5 months of age but not before (Jusczyk, Hohne, & Bauman, 1999).

Even when infants are successfully able to separate and recognize a word from fluent speech, they must then map the word onto the proper referent, which may be a concrete object or an abstract concept. In addition, identifying the correct referent can itself be a challenging task, given the complex environment surrounding infants and the many possible referents that exist when a word is heard. Mapping words onto their meanings is further complicated when infants must extend the same word, such as car or baby, onto other objects or individuals. In addition, an object can be associated with more than one word, as is often the case for both its noun and proper noun labels (Waxman & Lidz, 2006). The ease of an infant's accessibility to a concrete object, rather than an abstract concept, has been used to explain the appearance of nouns before verbs in an infant's early vocabulary (Gentner, 1982). An additional challenge is the connection between the different grammatical forms of a word and meaning. Thus, infants rely on a combination of the conceptual, perceptual, and linguistic systems as well as social pragmatic skills to successfully map words onto their meanings (Waxman & Lidz, 2006).

Although infants struggle to learn their first few words around 12 months of age, they soon display remarkable mastery in word learning within just one year after, and a rapid expansion in vocabulary ensues. In an attempt to explain this rapid development and incorporate the multiple cues for word learning, there is growing support for the emergentist coalitionist theory (Hollich et al., 2000). This theory is based on the wide range in approaches, including attentional, social, and linguistic, used by infants to successfully map words onto their proper referents. These word learning mechanisms are not fixed, but rather mature with time, and thus change in relative importance over the course of lexical acquisition.

Given the nearly infinite number of possible referents in an infant's immediate environment, the emergentist coalitionist theory suggests that infants rely on the interaction of multiple cues in attaching the label they hear to the correct object, action or event (Hollich et al., 2000). The properties of the object in question may play an important role, especially for novice word learners. These early word learners may demonstrate an innate bias in assuming that a label they hear refers to an object, rather an action or event, and the whole object for that matter, not a specific part or feature of it. Furthermore, during the first and second years of life, shape-based extensions of learned words are much more common than extensions based on other properties, such as size or color (Hollich et al., 2000). After two years of life, this shape bias seems to lessen, as older children are able to extend the same word to objects that are highly dissimilar in both shape and other factors. In addition, early word learning infants likely implement various attentional mechanisms, such as perceptual saliency, association, and frequency of co-occurrence between words and referents, in properly mapping words to their referents (Hollich et al., 2000). For example, temporal contiguity, which is the joint movement of an object coupled to the production of the word, may be important for early word learning, such that infants will attach labels to objects that actively recruit their attention.

As infants progress in their lexical development, these attentional cues seem to decline in importance and are replaced by a greater dependence on social and linguistic cues, such as eye gaze or grammar (Hollich et al., 2000). This shift becomes apparent around 18 to 19 months of age when infants display sensitivity to eye gaze, pointing, and speaker intention when mapping a word to its referent. Thus, social pragmatic skills become increasingly important, as children rely on adults as expert word learners and their guides to lexical acquisition. Furthermore, it has been suggested that adults facilitate word learning by talking about and identifying objects, actions, and events that infants are already focused on (Hollich et al., 2000). By doing so, adults may help remove some of the burden of identifying the proper referent out of the nearly infinite possibilities available. Thus, through their communication with infants, adults may limit the number of possible word-referent mappings that word learning infants must evaluate and consider when attaching a heard label to its object, action, or event. Not surprisingly, parents who employ such "joint attention" strategies have children with more advanced vocabularies compared to their peers (Hollich et al., 2000). As infants gain experience in the different cues discussed, the development of word learning heuristics likely takes place and may explain the rapid acquisition of vocabulary that is seen around two years of age. In summary, infants make use of a variety of cues, such as attentional, social, and linguistic, in the mapping of a word onto its proper referent. Successful mapping is likely not possible when only one of these cues is present, but rather their interaction is key for successful word learning.

This mapping problem certainly presents a specificity conflict for word learning infants. On the one hand, infants should encode fine phonetic detail as they learn new word forms. Indeed prior to the onset of mapping, this is the case for 7.5-month-olds who listened longer to familiarized word targets than those that differed in only one or two phonetic features, such as *feet* and *zeet* (Jusczyk & Aslin, 1995). Interestingly, Hallé and de Boysson-Bardies (1996) found that 11-month-old, French-learning infants displayed no preference for commonly occurring words and their phonetically similar counterparts, which differed only in the word-initial consonant. The researchers explained this apparent discrepancy by suggesting that the older infants employed a more generalized listening approach focusing on the semantic content, rather than the fine phonetic detail, of the words and thus were more tolerant of small changes in sound structure.

However, in both of these studies, the head-turn preference procedure was used, and infants were only required to recognize and demonstrate a listening preference for one word form over another. Thus, the head-turn preference procedure does not require the infant to link the word to meaning, but rather to just pay attention to the word's acoustical features. In comparison, preferential looking and word learning tasks more accurately test for the effect of the phonological encoding of a word on the recognition of its corresponding object. When applied to lexical acquisition research, preferential looking tasks compare relative looking times to the visual target object and to a non-target object presented simultaneously when the target word is spoken versus during baseline. Word learning tasks are advantageous in requiring an infant to map a novel word, either one that is nonsense or unfamiliar, onto an unknown object. Thus, looking preference and word learning tasks are superior to listening preference tasks in more accurately testing an infant's ability to map a word onto its meaning and not just the ability to discriminate the different phonetic features of a word without any semantic context.

The added challenge of object association can change infants' sensitivity to phonetic patterns of speech (Stager & Werker, 1997). Consequently, it is possible that while infants may become more tolerant of phonetic alterations with age in their performance on listening preference tasks, the supplemented challenge of object association may prolong their sensitivity to phonetic word changes in their performance on preferential looking and word learning tasks. This indeed appears to be the case, as infants must take both the phonetic features and meaning of a word into account. Using the preferential looking task, 12-month-olds were found to display sensitivity to mispronunciations of the vowels and consonants in familiar words, such that these mispronunciations impaired their ability to recognize the word's corresponding object (Mani & Plunkett, 2010). Furthermore, while word recognition was sensitive to all vocalic changes, this effect was heightened in infants with larger vocabularies. Fourteen-month-olds continue to display such sensitivity to close and distant mispronunciations, such as *baby* versus *vaby* or *raby* (Swingley & Aslin, 2002). This suggests that infants continue to utilize phonetic detail in word recognition tasks. In a test of whether infants can successfully learn and map phonetically dissimilar nonsense words, such as *lif* and *neef*, onto two novel objects and detect a "switch" in the word-label pairing, 14-month-olds were capable of making such associations and detecting a switch, while younger infants failed to (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). This ability seems to be enhanced with age. While 14-month-olds fail to successfully map phonetically similar nonsense words, such as *bih* and *dih*, onto two novel objects in this "switch" task, 20-month-olds demonstrated complete success, whereas 17-month-olds demonstrated marginal success (Werker, Fennell, Corcoran, & Stager, 2002). Based on these findings, the researchers advocated for a developmental explanation behind the increasing phonological specificity seen in early word learners, such that by 17 months of age, the infants had acquired a sufficiently large enough vocabulary to display such detailed phonological encoding.

In a second study, Swingley and Aslin (2000) found that 18 to 23-month-olds were significantly delayed in recognizing a familiar object when its label was mispronounced and differed in one or two phonetic features, such as *baby* and *vaby*. In this setup, infants heard these labels embedded in a carrier phrase, while two pictures were presented to them, one of which was the target referent. The infants were able to recognize the target object when either the correctly pronounced or mispronounced label was used, but they were significantly slower at recognizing the object with the mispronounced label, as assessed by the infants' looking times after the onset of the target word. Although a clear effect of mispronunciation on word recognition was observed, the researchers found no correlation between the effects of mispronunciation and age or vocabulary size, which stands in contrast with the developmental explanation advocated by Werker et al. (2002). Nonetheless, such findings also provide support for the increasingly important role of phonetic detail in recognition during early language acquisition. As infants become older and must learn additional word forms, their ability to distinguish between similar words becomes increasingly valuable. As such, phonological specificity appears to play a more important role when infants must evaluate two words in the context of their corresponding meanings, rather than just discriminating between their acoustic features. This may explain why older word learning infants no longer display sensitivity to phonetic changes in listening preference tasks but continue to do so in looking preference and word learning tasks, for which successful performance requires consideration of the word's meaning.

An additional explanation put forth for the increased encoding of phonological detail seen in older infants is based on the familiarity of words. Swingley (2007) found that 1.5-year-olds familiarized with a word's phonological form were sensitive to mispronunciations of that word and showed impaired recognition, while infants who did not undergo familiarization did not show such impairment. Based on the familiarity explanation, the delay in word recognition for mispronunciations of familiar words but not for those of recently learned words stems from prior detailed phonological encoding of familiar words and the lack thereof for recently learned words. Thus, an older child should display the same level of phonological encoding for familiar words as a younger child but should differ in the number of familiar words in their vocabulary and thus display more specificity in general (Bailey & Plunkett, 2002). This has been supported by the lack of a significant difference in the phonological specificity seen for familiar words in 14-month-olds versus 24-month-olds (Swingley & Aslin, 2000; Swingley & Aslin 2002). In an attempt to resolve the difference in the performance of 14-month-olds in the task employed by Werker et al. (2002) and in the task employed by Swingley & Aslin (2002), it is important to note the use of novel nonsense words and a switch task following habituation in the case of Werker et al. and the use of familiar words and a preferential looking task by Swingley and Aslin. In a modified study, 14-month-old infants were indeed able to associate two phonetically similar novel words with two different objects when a visual choice method, rather than the previous switch task, was implemented (Yoshida, Fennell, Swingley, & Werker, 2009).

Taken altogether, these studies demonstrate that 14 to 24-month-olds display sensitivity to mispronunciations of novel and familiar words. Increasing phonological sensitivity is certainly advantageous to infants during early lexical acquisition; however, focusing too much on specificity can potentially interfere with infants' ability to accommodate natural variation in speech, such as one speaker versus another. In addition to mispronunciations, it is expected that such modifications to speech, including speaker identity, would affect word perception in this age group. Different speaker identities certainly present a challenge to word learning infants, as they must not only understand and respond to words spoken by their primary caregiver but to words spoken by the increasingly large number of voices they are exposed to as they get older. In addition, word learning infants often face the challenge of having to preferentially listen to one stream of speech, while a simultaneous stream of fluent speech or background noise is occurring. Despite the expected difficulty of this task, 7.5-month-olds are indeed able to separate a stream of words spoken by a target voice from those spoken by a competing voice, but only when the target voice is sufficiently louder (Newman & Jusczyk, 1996).

Talker familiarity may be an important cue infants use when selectively streaming speech from one voice over that of another, as 7.5-month-olds infants preferentially listened to familiarized words when the words were familiarized with their mother's voice but did not show a preference when the words were familiarized with a stranger's voice (Barker & Newman, 2004). Furthermore, Houston and Jusczyk (2000) found that 7.5-month-olds succeeded at generalizing words across two speakers of the same sex but failed to do so for two speakers of the opposite sex; however, infants did show this ability by 10.5 months of age. 7.5-month-old infants also failed to recognize familiarized words when they differed in pitch or affect; however, these acoustical features seemed to decline in importance by the end of the first year of life and no longer impaired word recognition (Singh, Morgan, & White, 2004; Singh, White, & Morgan, 2008). Based on these findings, it is likely that both phonetic features of speech and talker-specific vocal characteristics are important for language processing in early word learning infants. However, while early word learners are able to generalize across a variety of acoustical features by their first birthday, 14 to 24-month-olds continue to display delayed recognition due to mispronunciations or modifications of the phonemes themselves (Swingley & Aslin, 2002; Swingley & Aslin, 2000; Mani & Plunkett, 2007; White & Morgan, 2008).

While speakers demonstrate variability in a wide range of characteristics, such as in tone, speed, or volume, one variable of particular interest is the different accents speakers have. Both changes in the actual phonemes as well as changes in the realization of those phonemes can occur when speakers of different dialects or languages pronounce words (White & Aslin, 2011). These changes typically require increased processing time for listeners and can ultimately lead to negative effects on speech perception. Although adult listeners are able to understand unfamiliar, foreign accents to some degree, accented speech is at times associated with low comprehensibility and intelligibility, while also increasing the time required to process the speech (Munro & Derwing, 1995a, b). As such, accents present a unique challenge to word learning infants. On the one hand, these early learners must encode an increasingly large number of word forms, and phonological specificity is a helpful strategy for doing so. Conversely, focusing too much on phonological detail would not permit or tolerate the small phonetic changes that often accompany unfamiliar accented speech.

Infants as young as 5 to 7 months old are able to discriminate between their own familiar accent and an unfamiliar regional accent of the same language in listening preference tasks, such as the head-turn preference procedure (Butler, Floccia, & Goslin, 2011). Furthermore, 9-month-olds were unable to recognize words that had been familiarized in their native accent when they were then later spoken in an unfamiliar dialectal or foreign accent (Schmale, Cristià, Seidl, & Johnson, 2010; Schmale & Seidl, 2009). However, the researchers found that 12 and 13-month-olds were indeed able to recognize the words spoken in the unfamiliar dialect or foreign accent. Similarly, both 15 and 19-month-olds were capable of recognizing familiar words said in their own native dialect, while only the older children were able to do so for a Jamaican nonnative dialect (Best, Tyler, Gooding, Orlando, & Quann, 2009). However, all of the aforementioned studies utilized listening preference tasks that only required infants to discriminate between the different phonemes and acoustical features of the accented words. As previously mentioned, preferential looking tasks are more suited for assessing infants' understanding of the meanings of words and have been used to study infants' perception of accented speech.

Using such a task known as the intermodal preferential looking procedure, Mulak, Best, Irwin, and Tyler (2008) reported that 19 to 20-month-olds were able to match a target word to one of two familiar pictures when the word was spoken in American English but failed when the word was spoken in Jamaican English. In this study, the target words were presented in isolation to infants, rather than embedded in a carrier phrase. The observed inability to normalize unfamiliar accented speech seems to continue with age, as older English-learning 24-month-olds were unable to recognize a newly learned nonsense word when it was spoken in an unfamiliar Spanish accent (Schmale, Hollich, & Seidl, 2011). Unlike Mulak et al. (2008), the infants in this study heard the target word embedded in a carrier phrase, both of which were spoken in the foreign accent. In addition, Schmale et al. (2011) employed the use of novel nonsense words in a word learning task, rather than testing the recognition of already familiar words as in Mulak et al. (2008). In summary, while infants as old as 24 months are able to discriminate between different words in a foreign accent, they are sometimes unable to generalize the meaning of a word to its foreign accented counterpart. Overall, hearing words in an unfamiliar accent has negative effects on processing and recognition in early word learning infants, although the exact cost may depend on whether the words are familiar or recently learned as well as on the unique properties of the accent itself. Novel accents can be detrimental to early word learners, as they can lead to an inability to understand speech in addition to potential mapping errors, such as associating an accented word with an incorrect referent (White & Morgan, 2008). Therefore, it is important to consider the conditions in which infants may be able or unable to tolerate the changes in speech processing brought about by an unfamiliar accent.

The present study sought to evaluate the effect two unfamiliar foreign accents would have on word recognition in early word learning infants. In particular, the effect of brief exposure to one of these unfamiliar accents on later recognition was investigated. Exposure to the foreign accents took place in two ways. Infants were exposed to one of the tested accents during a brief exposure story experience, which constituted the first form of exposure. In the second form of exposure, all of the infants were then exposed to each of the accents in the test trials, in which both the target word they were asked to locate and the carrier phrase it was embedded in were accented. Previous research in adults has demonstrated that just one minute of exposure to a non-native foreign accent is sufficient in reducing and even eliminating the increased processing time and reduced comprehension associated with hearing the accent (Clarke & Garrett, 2004). Such research has prompted additional investigation as to whether such an effect may take place in word learning infants. White and Aslin (2011) created an artificial “accent” by introducing a singular vowel shift to familiar words, such as the pair *dog* and *dag*. In a brief exposure phase consisting of ten labeling events, 18 to 20-month-olds were trained on this shift and later tested on their ability to generalize the shift to words they did not hear during the exposure phase. Both the standard and mispronounced versions of the words were asked for while embedded in a carrier phrase that was not spoken with the artificial "accent." While infants who had been exposed to the artificial "accent" accepted mispronunciations containing the same shift, infants who were exposed to standard pronunciations or another vowel shift did not. The researchers proposed a top-down explanation, whereby the infants experienced a broadening of lexical categories to accommodate for the newly exposed mispronunciations associated with the "accent."

However, it should be questioned as to how applicable these findings are to a natural accent that infants would likely encounter in everyday experience, rather than one artificially produced through only a vowel shift. Schmale, Cristià, and Seidl (2012) explored such an effect and investigated how two minutes of fluent speech in a Spanish foreign accent affected word learning in 24-month-olds. Infants were able to learn nonsense words, such as *neech* and *moof*, in both familiar and unfamiliar accents. However, their learning ability was enhanced for the nonsense words spoken in an unfamiliar accent if they had prior, albeit brief, exposure to that accent. During this approximately two minute exposure period, infants heard one or multiple speakers reading a passage either in the local familiar American English accent or the foreign unfamiliar Spanish accent. The ability of the infants to then learn nonsense words in both of these accents was tested through the preferential looking procedure and assessed via their visual fixation to the target object after the onset of its corresponding label versus their fixation to the nontarget object. In all training-test trials, infants were presented with two novel objects and learned to associate a nonsense word label spoken in the local familiar accent with one of the two objects. They were then asked to locate the trained object using this same label but in the foreign unfamiliar accent, after which they were then asked to locate the opposite novel object in the pair using a novel label they had not heard. If the infants had successfully learned the label for the trained object, then they were expected to fixate on the trained object when the previously learned label was heard and to not fixate on it when the novel label was heard. Although both groups of infants who were exposed or not exposed to the foreign accent prior to the training-test trials demonstrated this effect, it was heightened in the infants who had prior exposure. In addition, it is important to note that only the target word for each trial was spoken in the foreign accent, while the carrier phrase it was embedded in was spoken in the local familiar accent for all trials and all infants.

Similarly, the current study tested the effect of a brief two minute exposure to fluent speech spoken in one of two foreign accents. Unlike the word learning task utilized by Schmale et al. (2012), we investigated whether infants could recognize familiar objects labeled in three different accents: standard Northeastern American English, Northern Mandarin Chinese, and Mexican Spanish. Thus, we did not teach them new combinations of phonemes and their ability to map that label onto an object, but rather explored the effect of an unfamiliar accent on their recognition of familiar objects they likely knew from everyday experience. It was hypothesized that monolingual infants would show recognition for the objects when spoken in all three tested accents but might show a delay when the word was spoken in either of the two unfamiliar foreign accents. Prior exposure to one of the two foreign accents was expected to decrease or eliminate this delay if the infants displayed such quick adaptation as in White and Aslin (2011) or Schmale et al. (2012). However, it is likely familiar words are encoded with high phonological detail in the lexicons of early word learners (Mani & Plunkett, 2010; Swingley & Aslin, 2000; Swingley & Aslin, 2002). Thus, infants may not tolerate mispronunciations of familiar words when spoken in a foreign accent, as they seem to do for novel words, despite a brief but perhaps influential period of exposure.

There are several key components of this study to highlight. First, infants were exposed to foreign accents by two main means. The first was an exposure story period during which the infants heard a passage from a children's book read in one of three accents. The second was during the test trials, in which infants were asked to locate an object they were likely familiar with. Both the target word and the carrier phrase it was embedded in were spoken in the accent for each test trial. Unlike White and Aslin (2011), this study sought to expose word learning infants to natural accents they might encounter in everyday experience, rather than an artificial one created by introducing a consistent phonetic change. Although Schmale et al. (2012) pre-exposed infants to a natural accent in a similar manner, they tested the ability of infants to learn nonsense words in a foreign accent and only accented the newly learned words, not the carrier phrase, in the subsequent test trials. The present study investigated early word learning infants' recognition of words already presumed to be familiar spoken in foreign accents and perceptually set up the infants for the foreign accents by having both the target word and carrier phrase spoken in the accent during the test trials. An additional aim of this study was to examine speech perception and word recognition in early word learners of more than one language. To date, there is little research on the ability of bilingual infants to recognize objects labeled in foreign accents compared to the ability of their monolingual peers.

Unlike their monolingual peers, bilingual infants face the challenge of learning two sets of regularities simultaneously, yet show a similar pattern of development (Werker, Byers-Heinlein, & Fennell, 2009). For example, despite the expected challenge of mapping two labels onto the same object, bilingual infants can produce translation equivalents, which are words in two languages that have the same meaning, early in the development of language production (Pearson, Fernandez, & Oller, 1995). Furthermore, bilingual infants seem to recognize familiar words as early and as accurately as monolinguals, although there may be an increased processing time required to do so (Werker et al., 2009). One study even found that 10 to 12-month-old bilingual infants outperformed their monolinguals peers in discriminating native phonetic contrasts (Sundara, Polka, & Molnar, 2008). In addition, bilingual infants are certainly exposed to a more varied phonetic environment on a daily basis than their monolingual peers and are more likely to hear words in one or both of their languages spoken with more than one accent. For example, a Spanish-English bilingual infant growing up in the United States will likely hear English words in both an American English and Spanish accent during the course of their development. As such, it was hypothesized that bilingual infants may be more tolerant of accent-induced variations than monolingual infants due to their more frequent experience with accented speech.

Method

Word recognition was tested using a modified version of the preferential looking paradigm developed by Golinkoff, Hirsh-Pasek, Cauley, and Gordon (1987). In this setup, infants are presented with two visual stimuli simultaneously and then asked for either of the two via an auditory utterance. If infants understand the target word and are already fixated on the corresponding target picture when the word is said, then they would be expected to continue to look at the target picture. However, if infants understand the target word and are fixated on the nontarget picture when the word is said, then they would be expected to rapidly switch their gaze to the target picture. The present study used this procedure with the target word requested in three different accents, after different groups of infants had been exposed to one of these accents.

Two of the accents were expected to be foreign to the infant, while one was expected to be familiar.

Participants

Twenty-six 16-24-month-olds (10 females and 16 males) constituted the final group of monolingual, English-learning infants (mean age = 19.3 months, SD = 2.4 months). Two additional monolingual infants were also tested, but their data were not included in the analyses, as they did not complete a sufficient number of trials for analysis (explained in the Coding of Videos section below) due to fussiness or disinterest in the task. In addition, nine 16-23-month-old infants (5 females and 4 males) constituted a separate group of bilingual infants (mean age = 19.0 months, SD = 2.8 months). The bilingual infants came from a variety of language environments. All of these infants were being raised bilingual in English and also in another language, which varied across infants. These additional languages included two infants each of Spanish and Portuguese and one infant each of Japanese, Laotian, Mandarin Chinese, German, and Hungarian. Based on parental report, the bilingual infants' exposure to English on a daily basis ranged from 25% to 75% of the time. Four additional bilingual infants were tested, but their data were excluded, as they did not complete a sufficient number of trials for analysis due to fussiness or disinterest in the task. All infants were recruited by mail or email solicitation from the local community, and their ethnic distribution reflected the local community's demographics. The monolingual infants were mostly Caucasian with the exception of one Asian infant. The bilingual babies were a more diverse group made up of varying ethnicities, including Caucasian (3), Asian (1), Hispanic (4) and biracial (Asian-African American) (1).

Exposure story experience

Prior to the sequence of test trials, each infant heard approximately two minutes of an excerpt from "The Gingerbread Boy" by Paul Galdone read by one of three speakers. Each speaker was recorded as described in the Auditory Stimuli section below. Scanned pictures taken from the storybook were assembled into a Power Point show, which was displayed on both screens simultaneously to the infant while hearing the auditory story. This visual display was used simply to keep the infant engaged and fixated on the monitors during this exposure period. The purpose of this exposure story experience was to immerse each infant in one of the three accents to be used in the test trials. It is important to note that the story excerpt did not contain any of the words that were then later used in the test trials.

Visual stimuli and setup

Following the exposure story, 36 pairs of pictures (4 repeated blocks of 9 picture pairs) were presented one at a time to the infants. All pictures were of concrete objects and represented count nouns that infants were likely familiar with from their everyday experience, such as baby, banana, and stroller. These target words were chosen from the infant and toddler versions of the MacArthur Communicative Development Inventories (CDI) and reports of infants' early vocabulary (Dale & Fenson, 1996). Most of these target words were two syllables, with the exception of some that contained one or three syllables. Target word pictures were selected from Google Images and scaled to match each other roughly in size, complexity, and brightness. Although some flexibility in size was permitted between target picture pairs, the two pictures within each presented pair were adjusted to match one another exactly in size. All visual stimuli filled the screens from top to bottom with black or white side edges depending on the particular background of the picture. As such, the width of each picture pair varied from 18 to 28 cm, while the height remained constant at 27 cm. See Figure 1a on the following page for pictures of the nine test stimuli pairs used and their corresponding target words.

The target pictures described above were displayed on two Dell flat panel monitors (36.5 cm x 27 cm) whose inner edges were separated by a distance of 24 cm. Equidistant between the two monitors was a Panasonic SD/Hard Disk Video Camera used to record each infant's looking behavior for later coding. An Altec Lansing Multimedia speaker was also located equidistant between the monitors and was used to play all auditory stimuli. Video and auditory output were routed to the monitors and speaker from computers located behind the setup. Black poster board and curtains were used to conceal all elements of the setup, with the exception of the monitor screens, from the infant's view. Infants were seated approximately 50 cm to 72 cm away from the center of the setup. See Figure 1b on the following page for a picture of the overall visual setup.

Auditory stimuli

For each pair of pictures, the infant was directed to look at one of the two visual stimuli through an auditory utterance. The target words were inserted into three carrier phases. These included (1) "Where's the ____, can you find the ____?", (2) "I see a ____, do you see the ____?", and (3) "Look at the ____, it is a nice ____.", which were adjusted if need be for plural target words, such as shoes or fingers. The carrier phrases were rotated through the words, such that each target word was said in only one corresponding carrier phrase (See *Appendix A.1* for the complete auditory script). The exposure story excerpt (described above) and the complete set of auditory utterances asking for the target words were recorded by each of three speakers, one in each of the three tested accents. These three accents were Northeastern American English, Northern Mandarin Chinese, and Mexican Spanish. The three speakers were all female and spoke in an infant-directed register. All auditory stimuli were recorded using a Panasonic SD/Hard Disk Video Camera (Model No. SDR-H60 P/PC), and the sound track portion of the video was played back to the infant during the visual exposure story experience and test trials. All auditory stimuli were played at the same comfortable volume setting for all infants tested.

(a) (b)

Figure 1 (a) Test stimuli pairs, (b) visual setup and display.

Experimental design

The 26 monolingual infants were randomly assigned to one of three story groups. The first group (n=9) heard the exposure story read in the American English accent and served as the control group, since it was expected that these infants were familiar with this accent based on their everyday experience. The two other monolingual groups were exposed to one of two foreign accents. The second group (n=9) heard the exposure story read in the Mandarin Chinese accent, while the third group (n=8) heard the story read in the Mexican Spanish accent. After the exposure story experience, the nine pairs of pictures were presented to the infants in blocks of four, thus a total of 36 test trials were presented. The order of the nine picture pairs was standard throughout the four blocks. Within each block of nine trials, the infants were asked for the target word in the three different test trial accents in sets of three. The order of these three sets of accent type was counterbalanced across infants. In the second block of trials, the opposite picture within each pair was asked for from what was asked for in the first block. Thus, all 18 visual stimuli were asked for once within the first 18 trials. The third and fourth blocks were identical in setup to the first and second blocks, except the left/right position of the pictures was reversed. The design was the same for the bilingual group, except all of these infants heard the exposure story read in the American English accent. See *Appendix A.2* for a sample order of stimuli presented.

Procedure

Infants arrived at the lab with their parents and were greeted in the waiting room, where the parents filled out an informed consent form and a questionnaire about their child's experience with language (See *Appendix B* for the informed consent form and *Appendix C* for the parent questionnaire). The infant and parent were then brought into the testing room and seated at a table in front of the experimental setup with the infant seated on the parent's lap. Parents were instructed to look at the center of the setup and not shift their head or body for the duration of the study. In addition, they were instructed not to label any of the visual stimuli or direct their infant's looking at any point, although they could comfort their child if need be. The lights were then turned off, so that the only source of light came from the monitor screens. The exposure story was then played in the accent designated for that infant with the visual display accompanying it for approximately two minutes. Following the exposure story experience, the test trials began. Each test trial began with the pair of visual stimuli presented for three seconds with no auditory input to familiarize the infant with both pictures and reduce the effect of any inherent preference for either. The carrier phrase containing the target word was then played and lasted for approximately five seconds followed by three to four more seconds of silence, after which the pictures were removed. The test trials were separated by one to two seconds of darkness. Thus, neither the start nor end of the trial was infant controlled. The test phase proceeded this way until all 36 trials were presented or the child became too fussy to continue. The infants and parents were then given a gift and certificate of appreciation as well as thanked for their participation.

Coding of looking behavior

The camera videos were uploaded and the infants' eye movements were coded off-line frame-by-frame (1 frame = 33 msec) using the Avidemux 2.5.6 multi-platform video editor. For each trial, the time frame of interest was the three seconds directly after when the first syllable of the target word was spoken, as used in previous and similar studies (Swingley & Aslin, 2002; White & Aslin, 2011). For the time preceding the utterance of the first syllable, it was noted whether or not the child looked at each of the two visual stimuli, as failure to do this was grounds for trial elimination as described below. In addition, the video frame number just after the first syllable was said was recorded, as was whether the child was looking at the target, nontarget, or neither at that moment. Next, the video frame number when the child looked away from the picture fixated as the first syllable was said (often towards the other picture of the pair) was recorded. The two video frame numbers were identified with the sound off and thus the observer blind to which side the target was on. Using the difference between these two frame numbers, the time “persisting” to look at the target or the nontarget, whichever was initially fixated, within the first three seconds of interest was calculated. These persistence times for either the target or nontarget constituted the main variable used in the analyses. If the infants understood the target word and were able to recognize it, then it was expected that overall persistence to the nontarget would be shorter than overall persistence to the target. If infants were initially fixated on the target picture and understood the word, then they were expected to continue fixating on it during and shortly after the onset of the target word. However, if they were initially fixated on the nontarget picture and understood the word, then they were expected to rapidly switch their gaze to the target picture during and shortly after the onset of the target word.

Data reduction

Data elimination

Data from infants who failed to complete at least the first two blocks of test trials (1 monolingual, 2 bilingual) or who did not have at least four usable trials for each of the three test trial accents (1 monolingual, 2 bilingual) were excluded. For each infant, individual trials were also excluded from analysis if the infant did not look at both the left and right stimuli before the target word was said (approximately 3% of test trials) or if the infant could not be scored as looking at either the target or nontarget when the word was said (approximately 6% of test trials). In addition, is it important to note that there was no systematic preference across the infants tested for either the left or right side. Of all the trials used in the analyses, the infants were looking at the left monitor when the first syllable of the target word was said in 49% of the trials and were looking at the right monitor in 51% of the trials.

Data construction

For each infant, the number of usable trials where he or she was looking at the target or nontarget within each accent type was identified within the first two blocks of trials. Only the first two blocks of trials were initially included for data compilation, as all infants whose data were included in the analyses completed these two blocks but not necessarily the third and fourth blocks.

Overall, the infants were looking at the target picture when the word was said in 48% of the trials by chance and were looking at the nontarget picture when the word was said in 52% of the trials by chance. Within the first two blocks, there were six test trials for each of the three accents. Ideally, each infant would have at least two trials for persisting to the target and two for persisting to the nontarget for each of the three accents tested within the two initial blocks. If the infant didn't have at least two trials within any one of these six categories, we "dipped" into the infant's third and fourth blocks to replace the missing trial(s) for that category. The first one or two usable trials which matched the missing type were then used when this was the case. However, on average, each infant had only one trial within the first two blocks that was excluded from analysis, so this "dipping" procedure was more of an exception than standard procedure. Based on at least two trials for each category, an average persistence time for continuing to look at the target and likewise an average persistence time for continuing to look at the nontarget were calculated for each of the three accents tested.

Results

Preliminary analyses

In order to look across several variables, such as exposure story accent and test trial accent, we initially determined a persistence difference score for each type of test trial (the three accents) for each infant. This score was calculated by subtracting the average persistence time continuing to look at the nontarget picture from the average persistence time continuing to look at the target picture. As such, if the infants understood the words and recognized them, then this persistence difference was expected to be greater than zero. Additionally, a higher persistence difference could be interpreted as an on average faster recognition of the target word when spoken in that particular accent. Thus, each infant had a persistence score for the American English accent, the Mandarin Chinese accent, and the Mexican Spanish accent. Using the persistence scores for the monolingual infants only, a 3 (exposure story accent – English, Chinese, Spanish) x 3 (test trial accent – English, Chinese, Spanish) mixed Analysis of Variance (ANOVA) was performed. There was no main effect of exposure story accent ($F(2, 23) = 2.52$, $p = .1024$) nor was there a main effect of test trial accent ($F(2, 46) = .79$, $p = .4599$). There was also no interaction between exposure story accent and test trial accent ($F(4, 46) = .79$, $p = .4599$). Overall, the persistence difference means did not differ significantly across the monolingual group (Figure 2). Given there was no effect of exposure story accent, in all further analyses the three exposure groups of monolingual infants were collapsed. This enabled conducting additional analyses using the separate average persistence time to the target and average persistence time to the nontarget for each test trial accent, rather than the persistence difference score that was used to initially consolidate the two times.

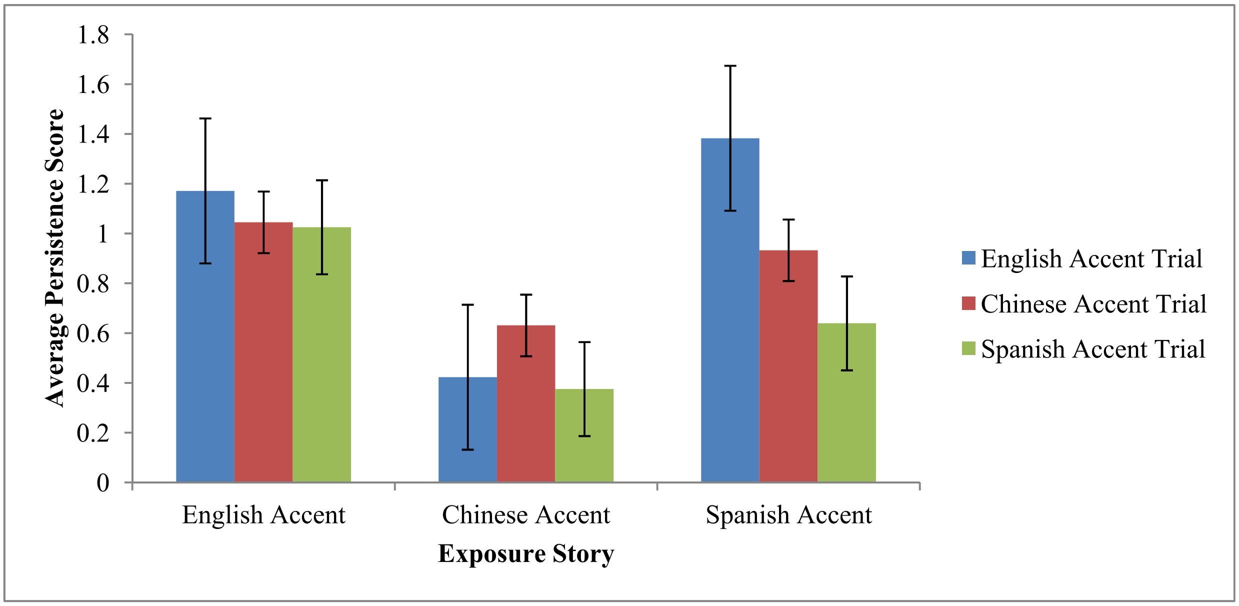

Figure 2 Average persistence difference scores across monolingual infants. English Accent Exposure Story Group n=9, Chinese Accent Exposure Story Group n=9, Spanish Accent Exposure Story Group n=8. Error bars represent +/- standard error from the mean.

In order to test for any gender effects within the monolingual infants, a 2 x 2 mixed ANOVA was performed, with gender as a between-subjects factor and average persistence to the target versus average persistence to the nontarget (across all three test trial accents) as the within-subject factor. While there was no main effect of gender ($F(1, 24) = .21$, $p = .6509$) nor was there a significant interaction ($F(1, 24) = .95$, $p = .3394$), a main effect of persistence to the target versus the nontarget was observed ($F(1, 24) = 41.86$, $p < .0001$). This key effect will be further examined below. Based on these findings, all further analyses were done treating the monolingual infants as a single mixed sex group.

Monolingual infants

If the infants indeed understood and recognized the target words, then it was expected that the time persisting to the target picture would be significantly greater than the time persisting to the nontarget picture. In order to test for this understanding within all monolingual infants, a 3 x 2 repeated measures ANOVA was performed, with both test trial accent and persistence time to the target versus nontarget as within-subject factors. There was no main effect of test trial accent as predicted based on preliminary analyses ($F(2, 50) = .63$, $p = .5353$), nor was there a significant interaction ($F(2, 50) = .83$, $p = .4418$). However, there was a strong main effect of persistence to the target versus nontarget ($F(1, 25) = 42.81$, $p<.0001$). Overall, the monolingual infants displayed a higher average persistence time to the target ($M = 2.17$, $SD = .78$) than average persistence time to the nontarget ($M = 1.33$, $SD = .77$) across all three test trial accents. See Figure 3 for category means.

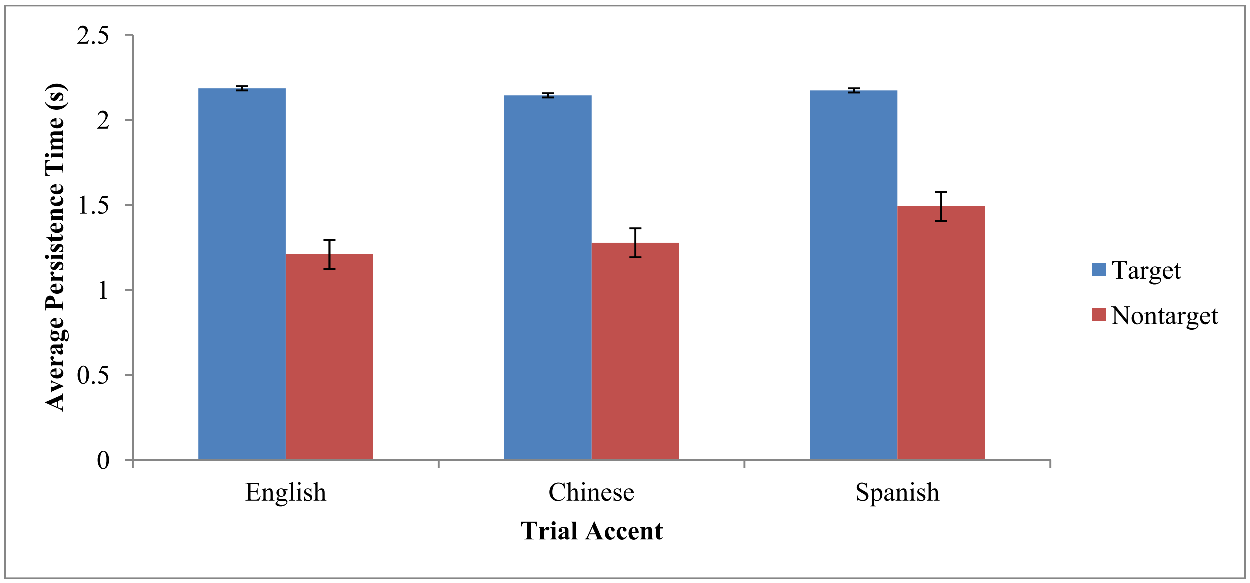

Figure 3 Target and nontarget persistence times across test trial accent in monolingual infants. Monolingual infants n = 26. Error bars represent +/- standard error from the mean.

Age effects

Although preliminary analyses revealed no significant effect of exposure story accent, the three means observed for the Chinese accent exposure story group were noted as lower when compared to the means of the English or Spanish accent exposure story groups (Figure 2). To investigate this and explore other potential variables, a one-way ANOVA for independent samples was conducted comparing the three monolingual story groups in terms of age. Indeed, a significant difference in the age of the three groups was found ($F(2, 23) = 5.18$, $p = .0139$).

Further post hoc comparisons using the Tukey HSD test indicated that there was no significant difference between the mean ages of the English and Spanish exposure story groups (*M* = 20.26, *SD* = 1.81; *M* = 20.19, *SD* = 2.83, respectively); however, the mean age of the Chinese exposure story group (*M* = 17.49, *SD* = 1.39) was significantly lower than the mean ages of the English and Spanish exposure story groups. Although this finding was unexpected due to the random assignment of participants to the three story groups, it suggests that the lower persistence difference scores evident in Figure 2 are perhaps age related and not story related. Thus, we proceeded to explore the possible effects of age on persistence time to the target and nontarget in all the monolingual infants. A median split was performed, which separated all the monolingual infants into a 16-19 month old group (n=13) and a 20-24 month old group (n=13). A 2 x 3 mixed ANOVA was run using the initial persistence difference scores calculated for each test trial accent. The between-subjects factor was age, while the within-subject factor was test trial accent. Consistent with the initial findings, there was no main effect of test trial accent on persistence score ($F(2, 48) = 1.11$, $p = .3379$) nor was there an interaction ($F(2, 48) = .17$, $p = .8442$). However, a main effect of age was found ($F(1, 24) = 15.27$, $p = .0007$), whereby the mean persistence difference score of the 20-24-month-olds (*M* = 1.21, *SD* = .86) was higher than the mean persistence difference score of the 16-19-month-olds (*M* = .44, *SD* = .83).

Given the lack of a test trial accent effect, the separate average persistence times to the target versus nontarget, rather than the persistence difference scores, were used to further investigate this age effect. A 2 x 2 mixed ANOVA was performed, with age again as a between-subjects factor and persistence to the target versus persistence to the nontarget as a within-subject factor. No main effect of age on persistence was indicated ($F(1, 24) = 1.46$, $p = .2387$), while the expected main effect of persistence time to the target versus nontarget was again observed

($F(1, 24) = 57.56$, $p < .0001$). The mean persistence time to the target across both age groups ($M = 2.17$, $SD = .56$) was significantly greater than the mean persistence time to the nontarget across both age groups ($M = 1.33$, $SD = .61$). In addition, persistence to the target was significantly greater than persistence to the nontarget within both the 16 to 19-month-old and 20 to 24-month-old age groups, as independently assessed by correlated *t*-tests ($t(12) = 3.04$, $p = .0103$; $t(12) = 7.94$, $p < .0001$, respectively). Interestingly, a strong interaction was found ($F(1, 24) = 9.81$, $p = .0045$). While the younger and older age groups did not differ in their mean persistence times to the target ($M = 2.11$, $SD = .56$; $M = 2.23$, $SD = .57$, respectively), the 16 to 19-month-olds persisted significantly longer to the nontarget ($M = 1.61$, $SD = .58$) than the 20 to 24-month-olds ($M = 1.04$, $SD = .50$) (Figure 4). Thus, in comparison to the older infants, the younger infants appeared to fixate longer on the nontarget before changing their fixation to the target picture.

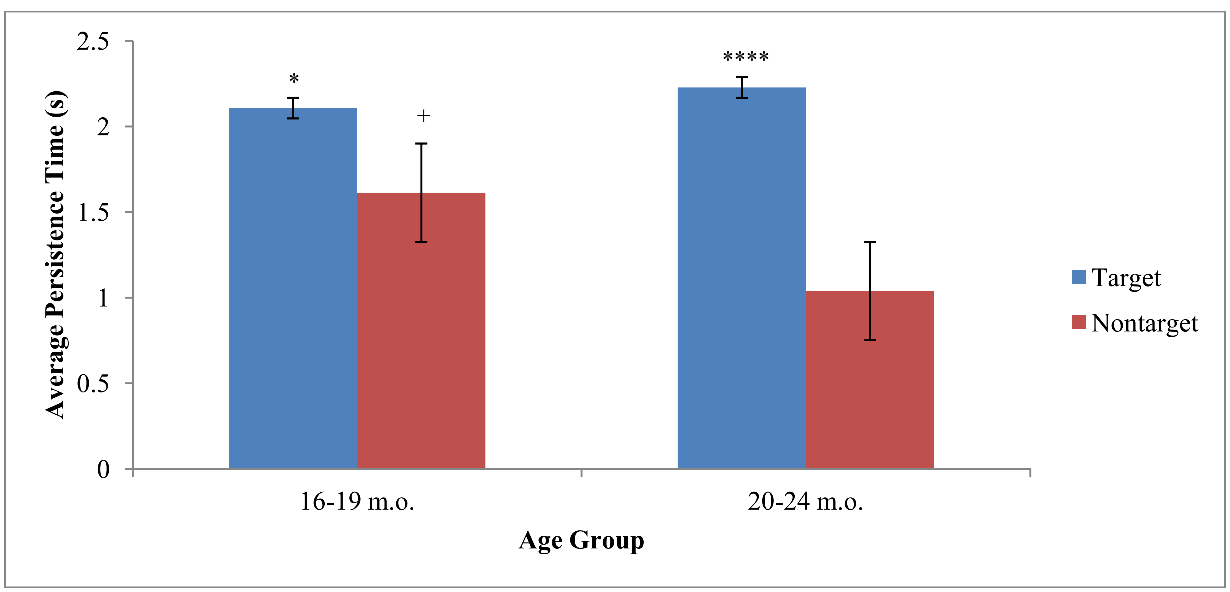

Figure 4 Target vs. nontarget persistence times in 16-19 m.o. vs. 20-24 m.o. monolingual infants. 16-19 m.o. group n = 13, 20-24 m.o. group n = 13. Error bars represent +/- standard error from the mean.
* $p < .05$ vs. 16-19 m.o. nontarget persistence
**** $p < .0001$ vs. 20-24 m.o. nontarget persistence
+ $p < .05$ vs. 20-24 m.o. nontarget persistence

To further investigate this differential persistence to the target versus nontarget in the two monolingual age groups, the shortest persistence time to the nontarget for each infant was identified. If the younger infants understood fewer words than the older infants, then using the average persistence time to the target versus nontarget would take into account those trials for which the younger infants might not have known the target word when the older infants did. Thus, it was hypothesized that the younger infants may have had longer average nontarget persistence times than the older infants because they knew fewer of the test words. Conversely, the younger infants may have had a slower reaction time and increased processing delay than older infants in the word recognition task, despite understanding the tested words as equally well as the older infants.

In order to explore these competing hypotheses, the shortest persistence times to the nontarget were analyzed, as these trials were interpreted as ones in which the infants confidently understood the word and were able to recognize the target object. An independent samples *t*-test was performed comparing the infants in each of the two monolingual age groups in terms of their shortest persistence time to the nontarget across all the test trial accents, given the lack of a test trial accent effect in the previous monolingual infant analyses. No significant difference was found between the average shortest nontarget persistence time for the 16 to 19-month-olds ($M = .17$, $SD = .21$) and that of the 20 to 24-month-olds ($M = .25$, $SD = .18$), ($t(24) = 1.00$, $p = .3273$). As such, it appeared that the younger infants displayed the same reaction time as the older infants in trials in which they confidently knew the word and recognized the target object. This finding supported the previously discussed hypothesis that the younger infants were just as quick as older infants to switch to the target picture when they understood the word but knew fewer of the tested words and thus had a longer average persistence time to the nontarget, since the average time took into account both trials in which they understood or did not understand the word. Therefore, the number of words an infant knows may be an important factor in their performance on word recognition tasks.

Word production effects

To further investigate the effect that word comprehension and production may have, the language questionnaires parents filled out were used to determine any clear differences between infants in such variables. Parental response to the language questionnaire did not yield any clear patterns of distinction between infants in terms of the words they comprehend, as most reported that their child knew most of the 56 words asked about, but rather parental response revealed a clear distinction in terms of the number of words the infants could say. The monolingual infants were divided into two groups based on parental report of language production, with the exception of one infant whose parent did not completely fill out the questionnaire. One group consisted of beginner talkers who were able to say 15 or less words (n=7), while the second group consisted of more advanced talkers who were able to say 20 or more words (n=18). An independent samples *t*-test indicated no significant difference between the age of the beginner talkers and the age of the more advanced talkers ($t(23) = 1.31$, $p = .2031$), which were respectively 18.17 months ($SD = 2.77$) and 19.48 months ($SD = 2.02$).

To investigate the possible effects of word production on test trial performance, a 2 x 3 mixed ANOVA was performed using the initial persistence difference scores calculated for each test trial accent. The between-subjects factor was word production, while the within-subject factor was test trial accent. As expected based on earlier analyses, there was no main effect of test trial accent ($F(2, 46) = 1.07$, $p = .3514$), nor was there an interaction ($F(2, 46) = .8$, $p = .4555$). However, a main effect of word production was found ($F(1, 23) = 10.86$, $p = .0032$).

The mean persistence difference score of the more advanced talkers (*M* = 1.01, *SD* = .82) was significantly higher than the mean persistence difference score of the beginner talkers (*M* = .21, *SD* = 1.01).

Given the lack of a test trial accent effect, the average persistence times to the target versus nontarget, each collapsed across the three test trial accents, were used to further explore the word production effect. A 2 x 2 mixed ANOVA was performed, with word production again as a between-subjects factor and persistence to the target versus persistence to the nontarget as a within-subject factor. No main effect of word production was indicated ($F(1, 23) = .89$, $p = .3553$), while the expected main effect of persistence time to the target versus nontarget was observed ($F(1, 23) = 60.21$, $p < .0001$), as in earlier analyses. The mean persistence time to the target of all the infants (*M* = 2.18, *SD* = .57) was significantly greater than the mean persistence time to the nontarget of all the infants (*M* = 1.36, *SD* = .60). Similar to the age effect, a strong interaction was found ($F(1, 23) = 14.79$, $p = .0008$). Unlike the two age groups, the beginner and more advanced talkers did not differ in their mean persistence times to the nontarget (*M* = 1.53, *SD* = .73; *M* = 1.29, *SD* = .55, respectively), but rather differed in their mean persistence times to the target. The more advanced talkers persisted significantly longer to the target (*M* = 2.36, *SD* = .49) than the beginner talkers (*M* = 1.70, *SD* = .49) (Figure 5).

In order to test for whether the beginner and more advanced talkers even understood and recognized the words, a paired *t*-test comparing persistence time to the target versus persistence time to the nontarget was conducted both within the beginner talkers as well as within the more advanced talkers. While the advanced talkers persisted significantly longer to the target than the nontarget ($t(17) = 8.7$, $p<.0001$), the beginner talkers did not ($t(6) = .84$, $p = .43$). Thus, the beginner talkers may have not only been able to say fewer words than the advanced talkers but may have understood fewer words as well.

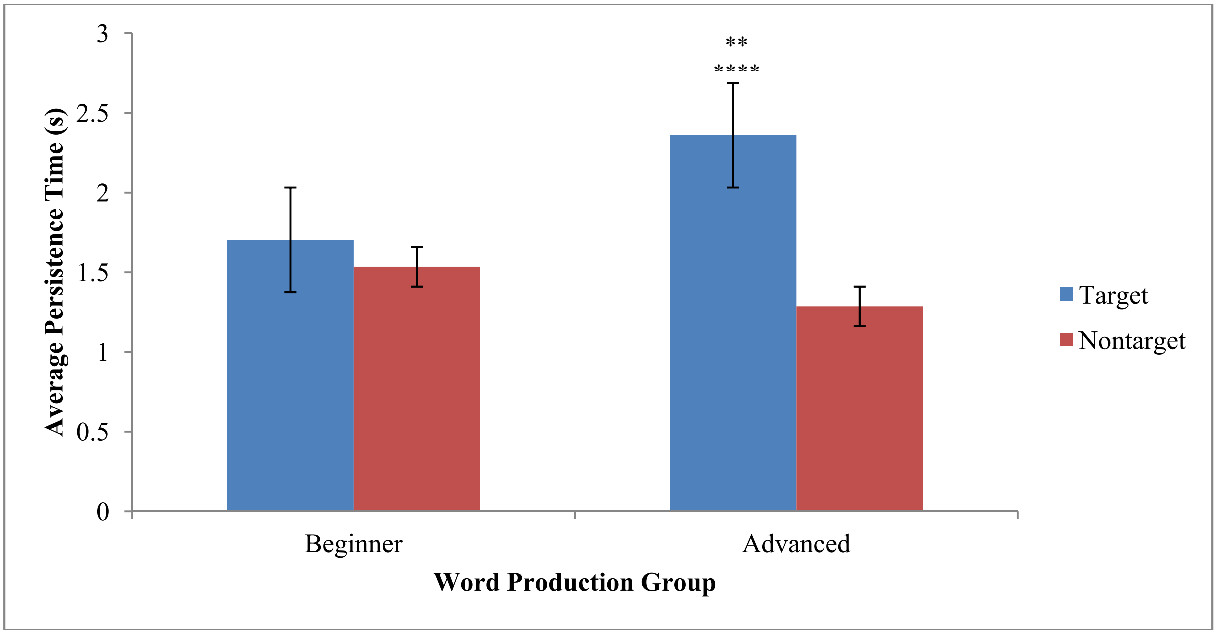

Figure 5 Effect of word production on persistence to target versus nontarget. Beginner talkers group n = 7, Advanced talkers group n = 18. Error bars represent +/- standard error from the mean. ** *p*<.01 vs. beginner target persistence, **** *p*<.0001 vs. advanced nontarget persistence

Bilingual infants

As previously stated, it was hypothesized that if the infants indeed understood and recognized the target words, then the time persisting to the target picture would be significantly greater than the time persisting to the nontarget picture. In order to test for this understanding and any potential test trial accent effect within the bilingual infants, a 3 x 2 repeated measures ANOVA was performed. Test trial accent and persistence time to the target versus nontarget both served as within-subject factors. Similar to the monolingual infants, there was no main effect of test trial accent ($F(2, 16) = .52$, $p = .6034$) and no interaction ($F(2, 16) = .70$, $p = .5115$). Also as within the monolingual infants, there was a main effect of persistence to the target versus nontarget ($F(1, 8) = 11.02$, $p = .0105$). Overall, the bilingual infants persisted significantly longer to the target picture ($M = 1.86$, $SD = .84$) than they persisted to the nontarget picture (*M* = 1.36, *SD* = .73). See Figure 6 for category means.

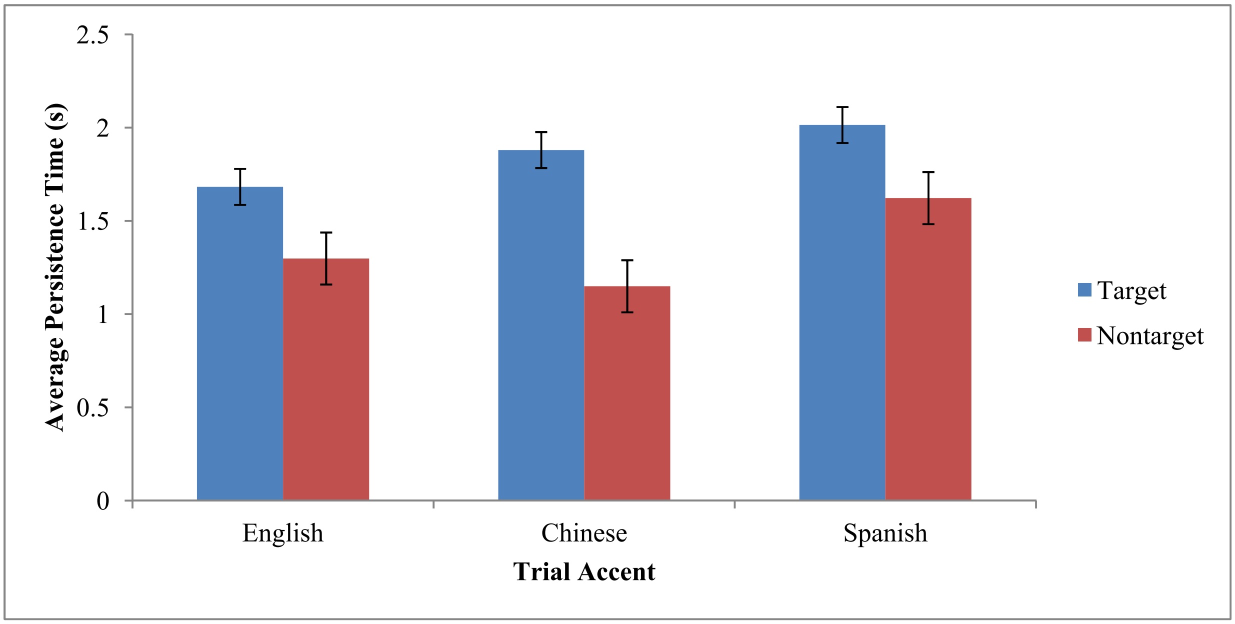

Figure 6 Target and nontarget persistence times across test trial accent in bilinguals. Bilingual infants n = 9. Error bars represent +/- standard error from the mean.

Monolingual versus bilingual infants

Given there was no effect of story exposure accent or test trial accent in the monolingual infants, the exposure story groups were collapsed and the persistence times were collapsed across the three test trial accents. Similarly, the persistence times for the bilingual infants were collapsed across the three test trial accents based on the lack of a test trial accent effect within this group. A 2 x 2 mixed ANOVA was done, with monolingual versus bilingual as a between-subjects factor and average persistence to the target versus average persistence to the nontarget as a within-subject factor. While there was no significant difference in the persistence times of the monolingual and bilingual infants ($F(1, 33) = .63$, $p = .4330$) and no interaction ($F(1, 33) = 2.05$, $p = .1616$), there was a significant effect of persistence to the target versus persistence to the nontarget ($F(1, 33) = 52.42$, $p < .0001$). Across the three test trial accents and for all infants, the average time persisting to the target picture (*M* = 2.09, *SD* = .53) was significantly greater than the average time persisting to the nontarget picture (M = 1.33, SD = .56) See Figure 7 for category means. Paired t-tests revealed a significant difference between persistence time to the target and persistence time to the nontarget within both the monolingual and bilingual groups ($t(25) = 6.54$, $p < .0001$; $t(16) = 2.58$, $p = .0201$, respectively).

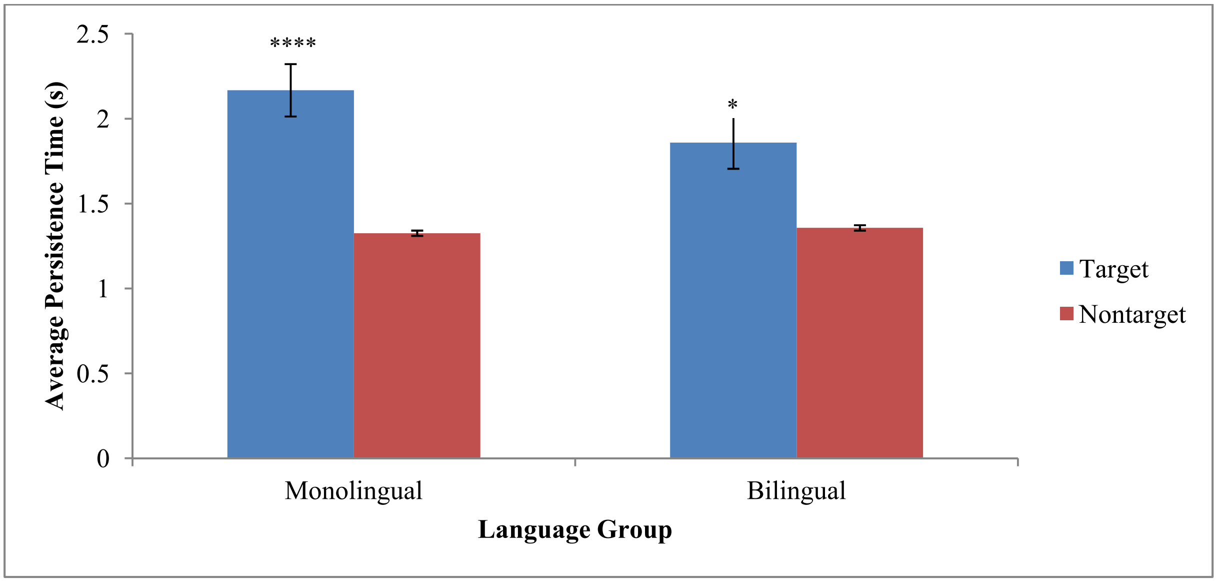

Figure 7 Average persistence times to the target and nontarget collapsing across exposure and test trial accent. Monolingual group n = 26, Bilingual group n = 9. Error bars represent +/- standard error from the mean.
**** $p < .0001$ vs. monolingual persistence to the nontarget
* $p < .05$ vs. bilingual persistence to the nontarget

To further compare the bilingual and monolingual infants, we used the group of monolingual infants who heard the English exposure story and experienced identical experimental conditions. An independent samples t-test revealed no significant difference ($t(16) = 1.16$, $p = .2631$) in the mean age of the monolingual English exposure story group (M = 20.26, SD = 1.81) and the mean age of the bilingual group (M = 18.98, SD = 2.77).

In order to investigate whether the bilingual and monolingual infants differentially understood the three test trial accents, a 2 x 3 mixed ANOVA was performed using the persistence difference scores initially calculated for each of the three test trial accents. The between-subjects factor was bilingual versus monolingual, while the within-subject factor was the three test trial accents. As in earlier analyses within the monolinguals, no main effect of test trial accent was found ($F(2, 32) = .28$, $p = .7576$), nor was there an interaction ($F(2, 32) = .49$, $p = .6171$). However, a main effect was indicated for the bilingual versus monolingual infants ($F(1, 16) = 4.71$, $p = .0454$), such that the average persistence difference score for the monolinguals ($M = 1.08$, $SD = .87$) was greater than that for the bilinguals ($M = .50$, $SD = .73$) (Figure 8). This main effect suggests there was a stronger understanding of the words within the monolingual infants than within the bilingual infants. Interestingly, the persistence difference scores of the bilingual group (Figure 8) were strikingly similar to those of the monolingual Chinese accent exposure story group (Figure 2), which was significantly younger than the two other monolingual groups. In addition, earlier age analyses supported that younger monolinguals showed weaker performance on the task due to understanding fewer of the target words on average than older monolinguals. Thus, bilinguals may have demonstrated persistence difference scores similar to younger monolinguals due to knowing fewer of the target words in English, as they may know some of their words in only one of either of their two languages.

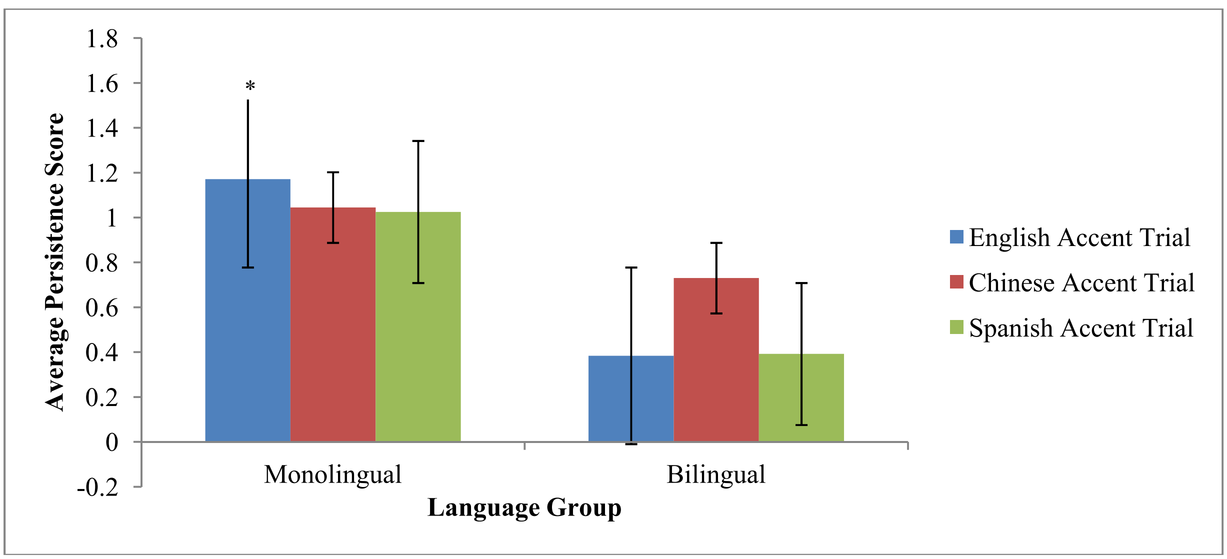

Figure 8 Average persistence difference scores of matched monolingual infants versus bilingual infants. Monolingual group n = 9, Bilingual group n = 9. Error bars represent +/- standard error from the mean.
* $p<.05$ vs. bilingual English accent trial persistence difference score

Given the absence of a test trial accent effect in the monolinguals versus bilinguals, a further 2 x 2 mixed ANOVA was performed with monolingual versus bilingual as a between-subjects factor and average persistence to the target versus average persistence to the nontarget across all three test trial accents as a within-subject factor. A main effect of persistence to the target versus nontarget was indicated ($F(1, 16) = 30.06$, $p<.0001$). In concordance with earlier findings, the overall persistence to the target ($M = 2.18$, $SD = .56$) was significantly greater than the overall persistence to the nontarget ($M = 1.40$, $SD = .54$). This effect was also observed within both groups as independently assessed by correlated *t*-tests (see Figure 9 for means and *t*-test statistics). Furthermore, there was a non-significant but suggestive main effect of bilingual versus monolingual ($F(1, 16) = 3.56$, $p = .0775$). Similarly, a non-significant but suggestive interaction was found ($F(1, 16) = 3.72$, $p = .0717$), whereby the difference in the average target and nontarget persistence times for the monolingual infants ($M = 2.50$, $SD = .53$; $M = 1.45$, $SD = .64$, respectively) was greater than the difference in the average target and nontarget persistence times for the bilingual infants ($M = 1.86$, $SD = .39$; $M = 1.36$, $SD = .44$, respectively). As Figure 9 shows, this potential interaction appeared to be primarily due to the fact that the monolingual infants persisted longer to the target than the bilingual infants did. When these measures were compared using an independent samples *t*-test, the difference was indeed significant ($t(16) = 2.94$, $p = .0096$). This effect is reminiscent of the results comparing beginner talkers to advanced talkers (see Figure 5). However, unlike the beginner talkers, the bilingual infants showed evidence of understanding the words.

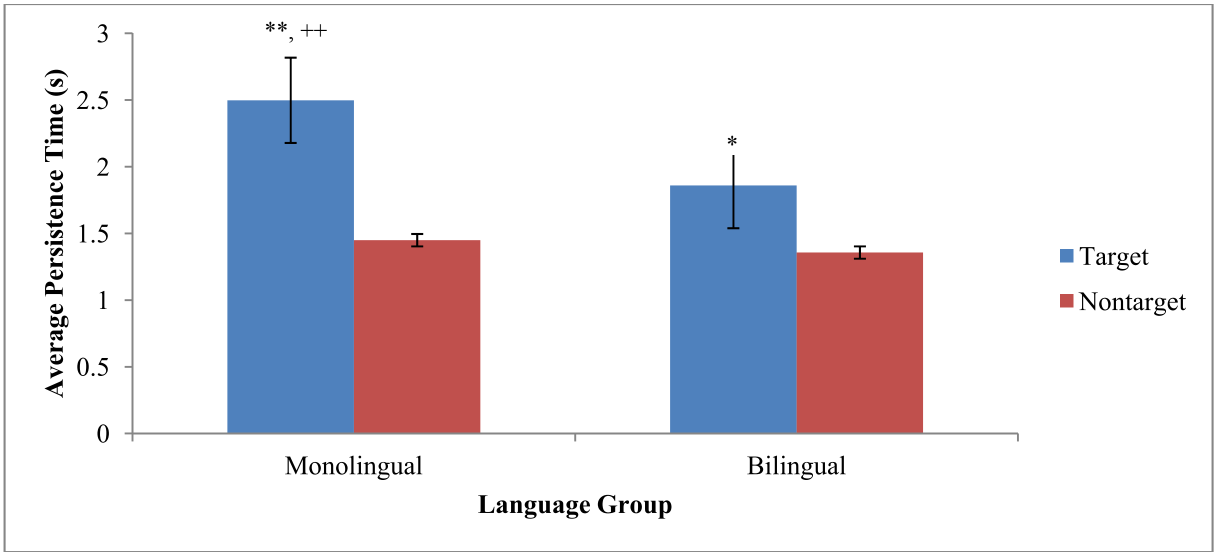

Figure 9 Average target and nontarget persistence times of matched monolingual infants versus bilingual infants. Monolingual group n = 9, Bilingual group n = 9. Error bars represent +/- standard error from the mean.
** $t(8) = 4.46$, $p < .01$ vs. monolingual nontarget persistence, ++ $p < .01$ vs. bilingual target persistence
* $t(8) = 3.32$, $p < .05$ vs. bilingual nontarget persistence

Discussion

Lexical acquisition is a complicated and arduous task for early word learning infants. Not only must they recognize novel word forms and correctly map these word forms onto their corresponding meanings, but infants must also determine how much perceptual change in speech means a phonetic and meaning change. This specificity problem is certainly evident in the wide range of accents infants may encounter in their everyday experience. Novel accents can induce changes in both the actual phonemes of words and the realization of those phonemes (White & Aslin, 2011). As such, they can hinder word comprehension and cause potential mapping errors in early word learning infants (White & Morgan, 2008). However, brief exposure to a novel accent, including either a natural foreign accent or an artificial "accent" created by a phonetic change, seems to improve the performance of 18 to 24-month-old infants on word learning and recognition tasks (Schmale et al., 2012; White & Aslin, 2011). This study investigated the effect of foreign accents on word recognition in 16 to 24-month-olds. Using a modified version of the preferential looking procedure, infants were presented with a pair of visual stimuli and asked to find the target word while embedded in a carrier phrase, both of which were spoken in one of three accents. While one of these accents (American English) was expected to be familiar to the infants, the other two were expected to be foreign (Mandarin Chinese and Mexican Spanish).

Interestingly, there was no indication of an additional benefit of exposure to a foreign accent on monolingual infants' ability to later understand a word spoken in that accent. Across all analyses performed, there were no significant effects or interactions involving the accent infants were exposed to before the test trials. There are two key explanations for this observation. First, all of the infants in this study understood the words presented to them while spoken in either a familiar or foreign accent. This understanding was assessed by comparing the time infants persisted in looking at the target picture if initially fixated when the target word was said versus the time infants persisted in looking at the nontarget picture if initially fixated when the target word was said. As hypothesized, persistence time to the target was significantly greater than persistence time to the nontarget. Furthermore, this was true across all three of the test trial accents. Thus, prior exposure to the accents may have had little facilitating effect, since all the monolingual infants understood the words spoken in the foreign accents anyway – a "ceiling effect." Unlike the infants in this study who understood a natural foreign accent even without pre-exposure, the infants in White and Aslin (2011) did not spontaneously understand an artificial "accent" created by introducing a single phonetic change. This may explain why the researchers found an added benefit of exposure beforehand, while this study did not.

A second explanation for the lack of an effect of exposure on later performance in this study is the nature of the task itself. Infants faced a two-alternative forced-choice test, in which they were asked to locate an object they were likely familiar with through everyday experience, such as *banana* or *stroller*, out of two possible options. It is likely that these familiar words are firmly encoded in the lexicons of early word learners and have already been heard spoken by a variety of speakers. As such, infants may be more tolerant of slight perceptual changes in speech when performing in preferential looking tasks testing their recognition of highly familiar words, such as the one utilized in this study. However, in more demanding tasks, such as the word learning task utilized by Schmale et al. (2012) testing the ability of infants to learn nonsense words in both familiar and foreign accents, infants may be less tolerant of such changes in speech. This may account for why Schmale et al. (2012) observed an added benefit of brief exposure to a natural foreign accent on later performance, while this study did not.

In addition to the exposure story having no effect on later recognition, there was no evidence of an effect of test trial accent on performance on the word recognition task within the monolingual group. All the monolingual infants understood words spoken in the three accents equally well. This finding conflicts with previous research that foreign accented speech impairs word comprehension in early word learning infants, although it may be partially explained by key differences in the method utilized by these studies. For example, Mulak et al. (2008) found that infants were able to match a target word to one of two familiar pictures when the word was spoken in American English but failed to do so when the word was spoken in Jamaican English. However, the target word was presented in singular isolation in Mulak et al. (2008), whereas in the present study, the target word was embedded in a carrier phrase also spoken in the foreign accent. Thus, placing the target word in the context of a carrier phrase may aid in the understanding of the accented word, particularly when the whole phrase is spoken in that same foreign accent. In keeping with this same idea, White and Aslin (2011) embedded the target word in a carrier phrase as well, but only the target word was spoken with the artificial “accent.”

However, in the present study, infants were perceptually “set up” to understand the foreign accent by accenting both the target word and the carrier phrase it was contained in. This may partly explain why novels accents hindered word comprehension in White and Aslin (2011) but not in this study. On this point, Schmale et al. (2011) found that infants could not recognize a target word when both the target word and its carrier phrase were spoken in a foreign accent. However, in this case as discussed before, infants’ recognition was tested with recently learned nonsense words, while in the present study, infants’ recognition was tested with words they were likely already familiar with.

It is also possible that the nature of the foreign accents used in this study may also explain why infants were so readily able to understand the accented target words. Whereas in this study natural Mandarin Chinese and Mexican Spanish accents were used, White and Aslin (2011) created an artificial “accent” by introducing a singular phonetic change to each target word but did not do the same to the carrier phrase each target word was contained in. While they found that infants who did not undergo pre-exposure to the artificial “accent” failed to later recognize words spoken in that “accent,” it is questionable how relatable this finding is to the natural accents infants are more likely to encounter in their everyday experience. Since foreign accented speech is often accompanied by multiple perceptual changes, it more likely that infants perceived this artificial “accent” as an isolated incident of mispronunciation or even a different word altogether, rather than a generalized shift in the pattern of speech production. Thus, early word learning infants may be more tolerant of a natural foreign accented word when it perceptually “makes sense” within a set of changes both within individual words and also applied across the surrounding words as well. Finally, the foreign accents used in this study (Mandarin Chinese and Mexican Spanish) may have been not so perceptually different from the familiar accent as to cause a change in word recognition, while the other foreign accents used in previous studies, such as the Jamaican English accent used in Mulak et al. (2008), were sufficiently different from the familiar accent to impair word recognition.

Despite the random assignment of the monolingual infants to each of the three exposure story groups, the Chinese exposure story group proved to be significantly younger than the English and Spanish exposure story groups. Although this was unexpected, it allowed for an investigation of the effect of age on persistence to the target and nontarget in 16 to 19-month-olds versus 20 to 24-month-olds within the monolingual group. The average persistence to the target was significantly greater than the average persistence to the nontarget within both age groups, supporting that both the younger and older infants understood the target words. However, using the initial persistence difference scores, there was evidence that the older infants performed better on the word recognition task than the younger infants, as there was a greater mean difference in persistence to the target versus nontarget in the older infants than in the younger infants. When further investigated, this greater persistence difference in the 20 to 24-month-olds appeared to be primarily due to a shorter persistence to the nontarget in the older group than in the younger 16 to 19-month-olds. Thus, the younger infants were slower on average to switch to the target picture when initially fixated on the nontarget picture than the older infants were to switch.

This observation presents two competing hypotheses as to why the younger infants may have been slower to switch from the nontarget to target picture. The first was that there was an inherent longer processing time in the younger infants to understand the target word and locate the corresponding picture within each visual stimuli pair. The second was that the younger infants on average knew fewer of the words tested than the older infants. As a result, the average persistence time to the nontarget took into account both trials in which infants knew the target word and the ones in which they did not. In order to explore these two explanations further, the single shortest persistence time to the nontarget was identified for each infant, as these times were interpreted as trials in which the infants confidently knew the target word. If the first hypothesis was correct that younger infants displayed a longer processing time even when they knew the target word, then the significant difference in the nontarget persistence times of the younger and older infants previously noted should have been observed again. However, this was not the case and the average shortest persistence time to the nontarget of the 16 to 19-month-olds was statistically equivalent to that of the 20 to 24-month-olds. This finding supported the second explanation that the younger infants on average knew fewer of the target words tested than the older infants. As such, the number of words an infant knows may be an important factor in their performance on word recognition tasks.

While parental response to the language questionnaire failed to differentiate the monolingual infants in terms of the words they comprehend, it revealed a clear distinction between the infants in terms of the words they produce. The monolingual infants were separated into two groups, one consisting of beginner talkers who could say 15 or less words and the other consisting of more advanced talkers who could say 20 or more words. The mean ages of these two groups were statistically equivalent. Based on the initial persistence difference scores, it appeared that the advanced talkers performed better on the task than the beginner talkers, as they showed a statistically greater average difference between their persistence times to the target versus the nontarget. Additional analysis revealed this greater difference was primarily due to a longer persistence time to the target in the advanced talkers than in the beginner talkers. Further investigation revealed that while the persistence time to the target was significantly greater than persistence time to the nontarget within the advanced talkers, it was not significantly greater within the beginner talkers. Thus, there was strong evidence that the advanced talkers understood the target words, while the beginner talkers did not. This made intuitive sense but certainly conflicted with parental report that the infants understood most of the words tested. As such, it seems that the parents were better at gauging the language production of their children than they were at gauging their language comprehension; parents may have overestimated the word comprehension of their children on the questionnaire. As previously discussed, verbal communication is accomplished through both linguistic and nonlinguistic cues, such as certain social pragmatic skills and physical cues in the environment. Communication with infants is certainly not an exception to this, and such cues may even be accentuated when parents communicate with their children. With such nonlinguistic cues absent, as in our forced-choice pictorial task, early word learning infants may display a lower comprehension of certain words. Thus, in this study, parents may have overestimated the word comprehension of their child on the questionnaire based on the nonlinguistic cues that they often provide during communication with their child. Therefore, this study’s findings in regards to the effect of language production support the potential role both linguistic and nonlinguistic cues play in the word comprehension of early word learning infants.

An additional goal of this study was to investigate word recognition in early learners of more than one language. Despite the added challenge of learning an additional language, bilingual infants reach key development milestones in their language acquisition at a similar pace as their monolingual peers (Pearson, Fernandez, & Oller, 1995). For example, bilingual infants are able to recognize familiar words as early and as accurately as monolingual infants (Werker et al., 2009). Thus, as expected, the bilingual infants in this study indeed understood and recognized the target words, as evidenced by their significantly greater overall persistence time to the target than to the nontarget. As in the monolingual infants, there was no effect of test trial accent on performance on the word recognition task within the bilingual group. Thus, the bilingual infants understood words spoken in the three accents equally well.

Although both the monolingual and bilingual infants understood words spoken in the foreign accents, there was a question of whether or not bilingual infants may demonstrate a strengthened or reduced ability to recognize words spoken in foreign accents when compared to their monolingual peers. On the one hand, bilingual infants may require an increased processing time to recognize a target word due to the higher language demands and multiple labels they face on a daily basis. As such, they may be able to recognize familiar words but with a delay when compared to their monolingual peers. On the other hand, bilingual infants are exposed to a more varied phonetic environment on a daily basis than monolinguals are, as they are more likely to experience hearing words in one or both of their languages spoken with more than one accent. Therefore, bilingual infants may be more tolerant of accent-induced variations in speech and thus demonstrate better performance on word recognition tasks than monolingual infants.

In order to explore these competing ideas, the bilingual group was compared to the monolingual English exposure story group that experienced identical experimental conditions. Upon analysis, the bilingual infants did not show a greater tolerance for foreign accents. On the contrary, there was some suggestion that they actually showed weaker performance on the recognition task in general, as the mean difference in persistence to the target versus nontarget of the bilinguals was significantly lower than that of the monolinguals. This difference is akin to that found between younger and older monolingual infants, suggesting that language development may be “delayed” in bilingual infants. When further investigated though, this smaller difference in bilinguals was primarily due to a shorter persistence time to the target, rather than a longer persistence time to the nontarget as in younger monolinguals. This may be partly explained by increased processing demands in bilingual infants, as they attempt to locate the target picture with certainty. In particular, word access may be more limited in bilingual infants, as they know some words only in English, others only in their second language, and still others with multiple labels across both languages. Consequently, even though the monolingual and bilingual infants persisted to the nontarget equivalently before switching to the target, the bilingual infants may have persisted for a shorter time to the target if initially fixated due to a greater need to “check” their choice against the opposite object. Thus, bilingual infants may be just as capable and quick as their monolingual peers to reject an object as not matching a given label; however, they may require an increased processing time to settle on a positive decision with certainty.

There are important limitations to the present study that are worth mentioning and call for future investigation. The first of such limitations involves the three accents tested in this study. Unfortunately, the Mandarin Chinese accent exposure story group was significantly younger in comparison to the other two story groups. Thus, a completely valid analysis of the exposure story effect was only possible for the Mexican Spanish foreign accent, as this effect in the Mandarin Chinese exposure story group was confounded by the effect of age. Future research should continue to explore a potential exposure effect using additional foreign accents and fully matched groups of infants. This could include both exposure involving passive listening to a foreign accent, as in this study, in addition to exposure involving listening to a foreign accent in the context of active social interaction. The effect of this more demanding exposure may be worth investigating, given the effects of word production noted in this study. As such, the level of exposure may be an important variable in the degree of adaptation to novel accents in early word learning infants. Additionally, the present study did not include any measures to qualitatively or quantitatively assess how perceptually different the foreign accents were in comparison to the American English accent. It may be that the foreign accents used in this study were not sufficiently different to affect performance on the word recognition task. Further investigation may want to explore the potential threshold in which accent-induced variations in speech begin to impair word comprehension or even to investigate understanding accents in more “degraded” listening tasks, such as listening to someone on the phone, processing speech in the presence of background noise, or performing in tasks consisting of multiple choices an infant must pick from.

Additionally, this study also tested the ability of word learning infants to recognize a target word contained within a carrier phrase that was spoken in the same accent as the target word. Thus, we investigated the ability of infants to recognize words presented in the context of the task, rather than in singular isolated events. This was done since infants are much more likely to hear a novel accent in a fluent string of speech, rather than as a single word, in their everyday experience. However, by perceptually setting the infants up for the task in the tested accent, the potential of the foreign accents to impair comprehension of the target word alone may have been diminished or eliminated. Future investigations may want to explore whether the foreign accents used in this study impair word recognition when the target word is presented alone and without any additional semantic context, as was the case in Mulak et al. (2008).

Additional research is certainly needed to further explore word recognition in bilingual versus monolingual infants. Although there was some suggestion that bilingual infants displayed slightly weaker performance on the task than their monolingual peers, this effect may either amplify or diminish when the group size is increased and the variation within the bilinguals is reduced. Due to constraints in recruiting subjects, there was a rather large variation in the second language the bilinguals infants were learning in addition to English as well as the degree of dominance of one language over the other for each infant. The implications of bilingualism may stem not from the mere fact that two languages are being learned and practiced simultaneously, but rather from the unique properties conferred by the two respective languages themselves (Werker et al., 2009). Thus, controlling for both the second language of English learning bilingual infants and the degree of dominance for either English or the second language may shed further insight on the effects of learning more than one language on word recognition. Finally, future studies may want to utilize additional measures of language development, such as word learning tasks, to compare monolingual versus bilingual infants. Due to the more varied phonetic environment they encounter on a daily basis and their greater exposure to different accents, bilingual infants may demonstrate an advantage over their monolingual peers in learning novel words spoken in foreign accents, although they showed no advantage for recognizing familiar words in our study.

Overall, the 16 to 24 month-old infants in this study were remarkably capable of recognizing familiar words spoken in both familiar and foreign accents. This was true regardless of whether they had prior exposure to these accents as well as whether they were monolingual or bilingual. Such flexibility is an important skill as infants move from the home, where they experience a rather limited phonetic environment, out into the world, where they hear a variety of different speakers from different cultures in a wide range of listening situations. Thus, the findings of this study support the impressive ability of early word learners to accommodate for slight variations in speech and adjust accordingly during the early phases of language acquisition.

References

Bailey, T.M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*(2), 1265-1282.

Barker, B.A., & Newman, R.S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*(2), B45-53.

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.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391-397.

Best, C.T., Tyler, M.D., Gooding, T.N., Orlando, C.B., & Quann, C.A. (2009). Development of phonological constancy: Toddler's perception of native- and Jamaican-accented words. *Psychological Science*, *20*(5), 539-542.

Butler, J., Floccia, C., & Goslin, J. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*(4), 392-417.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4): i-vi, 1-143.

Clarke, C.M., & Garrett, M.F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*(6), 3647-3658.

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125-127.

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191-218.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuozaj, II (Ed.), *Language Development*, *Volume 2: Language Thought and Culture* (pp. 301-334). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-45.

Hallé, P.A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, *19*, 463-481.

Hollich, G.J., Hirsh-Pasek, K., Golinkoff, R.M., Brand, R.J., Brown, E., Chung, H.L., Hennon, E., Rocroi, C., & Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), 1-135.

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), 1570-1582.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1-23.

Jusczyk, P.W., Hohne, E.A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues to word segmentation. *Perception & Psychophysics*, *61*, 1465-1476.

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

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*, 606-608.

Mattys, S.L., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465-494.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-272.

Mani, N., & Plunkett, K. (2010). Twelve-month-olds know their cups from their keps and tups. *Infancy*, *15*(5), 445-470.

Mulak, K.E., Best, C.T., Irwin, J.R., & Tyler, M.D. (2008). The effect of dialect on toddler identification of depicted words. International Conference on Infant Studies, Vancouver, BC, Canada.

Munro, M.J., & Derwing, T.M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97.

Munro, M.J., & Derwing, T.M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language Learning*, *38*(3), 289-306.

Newman, R.S., & Jusczyk, P.W. (1996). The cocktail party effect in infants. *Perception and Psychophysics*, *58*, 1145-1156.

Pearson, B.Z., Fernandez, S., & Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language*, *22*, 345-368.

Polka, L., & Werker, J.F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421-435.

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

Schmale, R., Cristià, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, *15*(6), 732-738.

Schmale, R., Cristià, A., Seidl, A., & Johnson, E.K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy*, *15*(6), 650-662.

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, *38*(5), 1096-1108.

Schmale, R., & Seidl, A. (2009). Accomodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, *12*(4), 583-601.

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 phonological lexicon in the face of variable input: Effects of pitch and amplitude variation on early 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), 381-382.

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, *108*, 232-242.

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

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*(2), 454-464.

Swingley, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147-166.

Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484.

Waxman, S.R., & Gelman, S.A. (2009). Early word-learning entails reference, not merely association. *Trends in Cognitive Sciences*, *13*(6), 258-263.

Waxman, S.R., & Lidz, J. (2006). Early word learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology*, *6th Edition*, *Volume 2* (pp. 299-335). Hoboken, N.J.: Wiley.

Werker, J.F., Byers-Heinlein, K., & Fennell, C.T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of The Royal Society B: Biological Sciences*, *364*, 3649-3663.

Werker, J.F., Cohen, L.B., Lloyd, V.L., Casasola, M., & Stager, C.L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289-1309.

Werker, J.F., Fennell, C.T., Corcoran, K.M., & Stager, C.L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1-30.

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*, 49-63.

White, K.S., & Aslin, R.N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372-384.

White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

Yoshida, K.A., Fennell, C.T., Swingley, D., & Werker, J.F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, *12*(3), 412-418.

Appendix A. Stimuli Preparation and Presentation

A.1. Script for Auditory Stimuli

1. Where's the **sailboat**, can you find the **sailboat**?
2. I see **bubbles**, do you see the **bubbles**?
3. Look at the **glasses**, they are nice **glasses**.
4. Where are the **fingers**, can you find the **fingers**?
5. I see a **flower**, do you see a **flower**?
6. Look at the **star**, it is a nice **star**.
7. Where's the **banana**, can you find the **banana**?
8. Look at the **baby**, it is a nice **baby**.
9. Where's the **window**, can you find the **window**?
10. Look at the **kitty cat**, it is a nice **kitty cat**.
11. I see **shoes**, do you see the **shoes**?
12. Look at the **teddy bear**, it is a nice **teddy bear**.
13. Where's the **stroller**, can you find the **stroller**?
14. I see an **elephant**, do you see the **elephant**?
15. Where's the **ducky**, can you find the **ducky**?
16. I see an **apple**, do you see the **apple**?
17. Look at the **butterfly**, it is a nice **butterfly**.
18. Where's the **sippy cup**, can you find the **sippy cup**?

A.2. **Sample Order of Stimuli Presented**

	ACCENT	BABY'S LEFT	LK TIME	BABY'S RIGHT	LK TIME	LATENCY	COMMENTS
1	English	***Shoes***		Baby			
2	English	Banana		***Stroller***			
3	English	Teddy Bear		***Window***			
4	Spanish	***Sailboat***		Apple			
5	Spanish	***Butterfly***		Fingers			
6	Spanish	Sippy Cup		***Kitty Cat***			
7	Chinese	***Flower***		Elephant			
8	Chinese	Ducky		***Bubbles***			
9	Chinese	Glasses		***Star***			
10	English	Shoes		***Baby***			
11	English	***Banana***		Stroller			
12	English	***Teddy Bear***		Window			
13	Spanish	Sailboat		***Apple***			
14	Spanish	Butterfly		***Fingers***			
15	Spanish	***Sippy Cup***		Kitty Cat			
16	Chinese	Flower		***Elephant***			
17	Chinese	***Ducky***		Bubbles			
18	Chinese	***Glasses***		Star			
19	English	Baby		***Shoes***			
20	English	***Stroller***		Banana			
21	English	***Window***		Teddy Bear			
22	Spanish	Apple		***Sailboat***			
23	Spanish	Fingers		***Butterfly***			
24	Spanish	***Kitty Cat***		Sippy Cup			
25	Chinese	Elephant		***Flower***			
26	Chinese	***Bubbles***		Ducky			
27	Chinese	***Star***		Glasses			
28	English	***Baby***		Shoes			
29	English	Stroller		***Banana***			
30	English	Window		***Teddy Bear***			
31	Spanish	***Apple***		Sailboat			
32	Spanish	***Fingers***		Butterfly			
33	Spanish	Kitty Cat		***Sippy Cup***			
34	Chinese	***Elephant***		Flower			
35	Chinese	Bubbles		***Ducky***			
36	Chinese	Star		***Glasses***			

Appendix B. Informed consent form

WORD LEARNING STUDY 2012-13 CONSENT FORM

Dear Parent:

You and your child have been invited to participate in our research on infant language development, and we thank you so much for coming! In today's project, we will be using the "preferential looking procedure." With this technique, infants are essentially given a choice between two pictures to look at, and their relative attention to each of the choices is recorded. Variations of this procedure have been used with infants from newborns to 2 year-olds to study a wide range of abilities. In our project, we will use the looking procedure to examine how infants just learning to talk respond to familiar words spoken in an unfamiliar accent. We are also interested in whether hearing the unfamiliar accent for a short period of time first helps infants to better understand familiar words spoken with that accent.

Your child will be seated on your lap in front of a "puppet theater" arrangement, and you will be with your child throughout the procedure. Your child will first hear a few minutes of a children's story (the Gingerbread Boy) being read by someone speaking English with either an American accent, an Hispanic accent, or a Chinese accent. After the story, we will then show your baby a series of pictures two-at-a-time on computer screens. These pictures will contrast different objects and animals that infants will probably be familiar with and may know the names of, such as a banana, teddy bear, or stroller. For each pair of pictures shown, a voice will play from a speaker directing your child to look at one of the two objects on the screens by using its verbal name. The voice will sometimes call attention to the chosen object with an American accent, sometimes with an Hispanic accent, and sometimes with a Chinese accent. As the pictures are shown, we will be observing your infant from behind the puppet stage and video recording their eye movements, so that we can later determine how quickly and for how long he or she looked at the named object.

The whole series of pictures should take less than 15 minutes to present to your baby, and babies often enjoy the show! Nevertheless, if at any time during the procedure, your baby becomes tired, overly bored, or uncomfortable, we will shorten or terminate the session. Likewise, if you decide at any time that you would like to end the session, you may certainly do so. The results of your infant's session will be combined with the results of other infants tested with the same procedure to examine our research questions; no information will ever be given in classes, research presentations, or anywhere else which could identify you or your child individually. The records of your child's behavior will be kept in a locked room and handled only by persons connected with our research. Please feel free to ask us any questions you may have about the goals of this study and the procedures involved. And, if at any time in the future you have questions about the research, please feel free to call the Behavior Developmental Lab at (617) 627-3057 or email us at tuftsbabylab@tufts.edu. Finally, if you have any questions about your rights as a research participant, you may contact Tufts' Institutional Review Board at (617) 627-3417.

Thank you again for allowing your child to participate with us today. You and your baby have helped us explore how infants learn and understand words, particularly when the words may sound different depending on who is saying them. We hope you will both enjoy this research experience!

Sincerely,

Emily W. Bushnell, Ph.D.
Director, Behavioral Development Labs

I have read the above description informing me of the purpose of this project. I understand that I may discontinue my child's participation at any time for any reason, and that all individual information used in this study will be kept confidential.

With that understanding, I give permission for my child ________________________ to participate.

Signature: ______________________________ Today's Date: ________________

Name and Address: ____

WORD LEARNING STUDY 2012-13 CONSENT FORM

ADDITIONAL PERMISSION TO USE VIDEO FOR TEACHING PURPOSES

On some occasions, we may wish to use selected portions of the video of your child's behavior to illustrate research findings to students in our classes and to colleagues in professional groups such as the Society for Research in Child Development and the International Society for Infant Studies. If you would be willing for us to use the video of your child for this purpose, please indicate below. Please be assured that in all such cases, your child will be identified only by age in order to preserve his or her anonymity.

I am so willing _______________ I am not willing _______________

Signature: ____ Date: ______________

Appendix C. Parent Questionnaire

Participant: Date:
Word Learning 2012-13, Tufts University Behavioral Development Lab

As part of our study of infants' comprehension of accented speech, we are interested in how familiar your child may be with hearing the words both used and not used in this study. Please answer the following to the best of your ability. If you have any questions or concerns, feel free to ask. If you are unsure or would prefer to leave any particular item blank, feel free to do so.

PART 1: Please fill in the blank next to each question or circle the most appropriate response; if a question does not apply, please write "N/A"

1. How old is your child (in months)? ____________________

2. Does you child have any older siblings? If so, how old are they? __________________________

3. Does your child hear more than one language on a daily basis? If so, what are the languages?

   Language 1 _________________________ Language 2 __________________________
   Others ______________________

4. Of the time your child spends hearing speech, about what portion of the time would you say they hear Language 1 indicated above?

   0% 25% 50% 75% 100%

5. Of the time your child spends hearing speech, about what portion of the time would you say they hear Language 2 indicated above?

   0% 25% 50% 75% 100%

6. Of the time your child spends hearing speech, about what portion of the time would you say they hear the other languages indicated above?

   0% 25% 50% 75% 100%

7. Does your child say any words in either of the languages mentioned above? If so, which languages?

   ____

8. How many words does your child know how to say in Language 1?

   0-4 5-9 10-14 15-19 20+

9. How many words does your child know how to say in Language 2?

   0-4 5-9 10-14 15-19 20+

Please see next page for additional questions

PART 2: In this section, we are interested in how familiar your child is with various vocabulary words, some of which are used in our study. By "knows", here we mean understands or recognizes a picture or example of – that is, your child doesn't have to be able to say or pronounce the word in question. Please indicate your response by checking the appropriate box.

Word	Not familiar/ Does Not Know	Somewhat Familiar/May Know	Very Familiar/ Knows	**Word**	Not Familiar/ Does Not Know	Somewhat Familiar/ May Know	Very Familiar/ Knows
Animal				Street			
Butterfly				Swing			
Monkey				Tree			
Bug				Fingers			
Bird				Elephant			
Duck				House			
Shoes				Hair			
Banana				Eye			
Airplane				Milk			
Water				Apple			
Baby				Crayon			
Train				Cookie			
Star				Ball			
Bubbles				Bottle			
Bread				Bicycle			
Flower				Cheerios			
Jacket				Teddy Bear			
Hat				Car			
Mouth				Sippy Cup			
Feet				Pajamas			
Hand				Window			
Table				Horse			
High chair				Stroller			
Sailboat				Puppy			
Chair				Balloon			
Glasses				Cell phone			
Rock				Kitty Cat			
Slide				Blanket			