8-2014

How Acoustic Salience Influences Infants' Word Mapping

Qian Zhao
University of Tennessee - Knoxville, qzhao2@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Qian Zhao entitled "How Acoustic Salience Influences Infants' Word Mapping." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Experimental Psychology.

Jessica S. Hay, Major Professor

We have read this thesis and recommend its acceptance:

Daniela M. Corbetta, Gregory D. Reynolds

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

How Acoustic Salience Influences Infants' Word Mapping

A Thesis Presented for the

Master of Arts

Degree

The University of Tennessee, Knoxville

Qian Zhao

August 2014

Abstract

Young language learners have the challenge of discovering which sounds in their complex auditory environment form acceptable object labels. During early word learning infants demonstrate both flexibility and constraint regarding what sounds form meaningful distinctions. Through language experience they hone in on the sounds and sound patterns that are meaningfully relevant in their native language. In the current study, I investigated the role that acoustic salience plays in early word learning. Using the Switch paradigm, 14-month-old infants were taught to associate two novel labels that differed only in pitch contour to two novel objects. Results from previous discrimination studies were used to select two pairs of pitch contours. For half of the infants the two pitch contours were highly discriminable (Tone 1, level vs. Tone 3, dipping; Salient Condition). For the other half of the infants the labels were less discriminable (Tone 2, rising vs. Tone 3, dipping; Non-salient Condition). Importantly, pitch contour is not used contrastively in English, and none of the infants had experience with a tone language. Only infants in the Non-salient Condition successfully mapped the novel labels to objects. Results suggest that the criteria for what makes a word a good object label involves a confluence of factors, including, but not limited to the acoustic salience of the contrast.

Table of Contents

List of Figures

Chapter 1

Introduction

The process of acquiring a native language begins very early in development and continues through early childhood and even into adolescent. For example, infant speech perception abilities change dramatically over the first year of life. From birth until around 6 months of age, young infants have ability to discriminate a large number of phonetic contrasts across the world's languages (Kuhl, 1987). At around 6 months of age infants begin to display language-specific perception of vowels (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). At around 9 months, young infants can already recognize the language-specific sound combinations (e.g., phonotactic regularities; Jusczyk, Friederici, Wessels, Jeanine, Svenkerud, & Jusczyk, 1993). Around 10-11 months infants show an increased ability to discriminate some native language consonant contrasts and show a declining ability to discriminate some non-native consonants (Werker & Tees, 2002). Although infants have gleaned a considerable amount of information about the categories of consonants and vowels used in their native language over the course of the first year of life, they still remain relatively flexible in in what they are willing to treat as an acceptable word form.

One of the foremost challenges infants face during early language acquisition is learning which sounds in the auditory environment are relevant for language. Our auditory environment is vastly complex, and includes speech sounds, natural animate sounds (e.g., a dog barking), natural inanimate sounds (e.g., trees rustling in the wind), and man-made sounds (e.g., the vroom of a car or the beep of an alarm clock). Never the less, infants eventually successfully isolate a small group of sounds as potentially lexically relevant. That is, infants eventually learn to reject different environmental sounds such as non-speech sounds (e.g., beep; Namy, 2001), communicative vocal sounds that are not word-like (e.g., *mmm*, *ahh*; MacKenzie, Graham, & Curtin, 2011), and speech sounds which violate native language phonotactic rules (MacKenzie, Curtin, & Graham, 2012; Graf Estes, Edwards, & Saffran, 2011) as labels for objects. How do infants determine which sounds are lexically relevant? The goal of the current research is to investigate the role of acoustic salience on infants' flexibility during early word learning.

Flexibility and Constraint in Early Word Learning

Young word learners demonstrate considerable flexibility in what they accept as a candidate object label (Acredolo & Goodwyn, 1988; Woodward & Hoyne, 1999; Hollich, Hirsh-Pasek, & Golinkoff, 2000; Namy, 2001; Namy & Waxman, 2000; Campbell & Namy, 2003). For example, numerous studies have demonstrated that early in the second year of life infants are willing to accept a wide variety of sounds as labels for novel objects, including symbolic gestures (Acredolo & Goodwyn, 1988; Namy & Waxman, 2000), pictograms (e.g., ; Namy, 2001), mouth noises (e.g., *psst*; Hollich, 2000) and non-speech sounds (e.g., squeak; Woodward & Hoyne, 1999). Given social-referential support (i.e., embedding target sounds in a labeling phrase) older infants (18-month-olds and older) will continue map both verbal- (e.g., *foppick*) and non-verbal sounds (e.g., two-tone beep) with to novel objects (Campbell & Namy, 2003).

As infants gain experience with the acceptable word forms in their native language, they appear to become less flexible in what they accept as an object label. For example, after about 18-month, infants become less willing to treat gestures as good object labels if they do not have prior experience with the gestures (Iverson, Capirci, & Capelli, 1994; Namy & Waxman, 1998), and 20-month-olds no longer accept novel non-word sounds (produced by a small noise maker) as object labels (Woodward & Hoyne, 1999).

In addition to becoming sensitive to acceptable basic word forms, infants become sensitive to the phonotactic patterns (the probability of occurrence of phonemes and phoneme combinations) of their native-language even before they begin speaking, and this early knowledge feeds into subsequent word learning (e.g., Graf Estes et al., 2011; MacKenzie et al., 2012). For example, at 12-months of age infants reject, as object labels, sound sequences that do not conform to native-language phonotactic patterns (e.g., the Czech word *ptak*). However, they continue to map non-native words (e.g., the Japanese words *sika* & *hashi*) that do not violate native-language phonotactics (MacKenzie et al., 2012). Similar effects of phonotactic probability on word learning can be seen at 17 months (Graf Estes et al., 2011) and beyond (Storkel, 2001). Twelve-month-old infants also reject communicative vocal sounds (e.g. *ooh*, *ssh*) as labels for novel objects, even though they can readily map phonotactially legal novel words onto novel objects (MacKenzie et al., 2011).

Minimal Pair Mapping

Interestingly, young word learners often have difficulty mapping words that differ minimally by a single phoneme (minimal pairs; e.g., *bin* and *din*) to novel objects, even if the words are phonotactically legal and the difference between the words is lexically contrastive in their native language. For example, Stager and Werker (1997) demonstrated that at 14 months of age, infants do not appear to attend to phonetic details in minimal pair words (e.g., *bih* and *dih*) when they are paired with objects even though they readily discriminate minimal pair words in an object-free task (i.e., when the words are presented with a checkerboard instead of an object). Follow-up studies have replicated this pattern of failure by 14-month-olds with numerous minimal pairs contrasts including *bin* and *din*, *bin* and *pin*, and *pin* and *din* (Pater, Stager, & Werker, 2004). But by 17 to 20 months of age, infants succeed in this minimal pair label-object association task (Werker, Fennell, Corcoran, & Stager, 2002). Werker and Curtin (2005) suggested that 14-month-olds might not have strong, rapid access to the phonetic details in the minimal pair words during cognitively taxing word-learning tasks (for related evidence see Fennell, 2012; Fennell & Waxman, 2010; Yoshida, Fennell, Swingley, & Werker, 2009).

The Role of Acoustic Salience in Word Mapping

Acoustic salience may play an important role in early label-object mapping. First, although infants are notoriously bad at mapping minimal pair words to novel objects, they readily map words with non-overlapping phonological features, such as *lif* and *neem* (Stager & Werker, 1997). Further, recent work by Curtin, Fennell and Escudero (2009) has demonstrated that infants rely more on acoustic cues for successful label-object mapping than phonemic differences between minimal pair words. Fifteen-month-old infants were presented two objects paired with minimal pair words that differed in a single vowel (i.e., *deet* and *dit*, *deet* and *doot*, or *dit* and *doot*). They found that infants succeeded in learning *deet* and *dit*, but failed on the other two contrasts. They suggest that at 15 months of age, infants rely on the acoustic salience associated with differences in vowel height in the /i/ - /I/ contrast to learn the object-label mapping. Finally, infants map words produced in infant-directed speech (IDS), which is an acoustically salient, to novel objects more readily than those produced in adult-directed speech (ADS) (Graf Estes & Hurley, 2013).

Other evidence that hints at the importance of acoustic salience in early word learning comes from recent work on English-learning infants' ability to interpret pitch contour as lexically contrastive in a label-object association task (Hay, Graf Estes, Wang, & Saffran, 2014). In English, pitch is an acoustically salient cue that, along with other cues, marks intonational/emotional meaning (Fernald, 1989; Moore, Spence, & Katz, 1997), phrase boundaries (Gussenhoven, 2004), and lexical stress (Ladd, 2008), and differentiates infant- from adult-directed speech (Fernald & Kuhl, 1987; Fernald, 1992), and female from male voices (Gunzburger, Bresser, & ter Keurs, 1987). However, in English, variations in pitch are not lexically contrastive, meaning that variations in pitch do not change the meaning of words in English. Thus, although pitch variation is communicatively relevant in English, it would be adaptive for English-learning infants to learn to ignore variations in pitch contour during word learning. In contrast, Tone languages use pitch contour to differentiate words. For example, in Mandarin Chinese there are four different tones used, which are Tone 1 (high-level), Tone 2 (high-rising), Tone 3 (low-dipping), and Tone 4 (high-falling) (Chao, 1948; Howie, 1976). For example, Mā (Tone 1, level) means mother, and Mǎ (Tone 3, dipping) means horse.

Pitch Contours and Word Mapping

Early in development, infants from non-tonal language backgrounds are able to perceive pitch changes. For example, Mattock and Burnham (2006) found that 6 month-old English-learning infants could discriminate Thai pitch changes. Further, in the second year of life, infants are able to attend to pitch contour information in newly learned words. Using a word-learning task, Singh, Hui, Chan, and Golinkoff (2014) found that 18-month-old English-learning infants were able to detect both pitch contour and vowel mispronunciations in newly learned words. However, by 24 months of age, English-learning infants failed to notice a mispronunciation in the label's pitch contour. Mandarin-learning infants at both ages noticed both the vowel and pitch contour mispronunciations. This result is consistent with recent findings by Quam and Swingley (2010) that demonstrated 30-month-old infants and adults do not associate pitch changes with changes in a word's meaning even though they remain sensitive to vowel changes.

Hay and colleagues (Hay et al., 2014) recently examined the role of language experience on infants' ability to map pitch contours to novel objects. Specifically, using a modified version of the Switch paradigm (Werker, Cohen, Lloyd, Casasola, & Stager, 1998) English-learning infants were habituated to two novel label-object pairings, where the labels were the syllable /kʊ/ produce with a rising pitch contour and /kʊ/ produced with a falling pitch contour. Following habitation infants were presented with Same trials, in which the label-object pairing from habituation were maintained, and Switch trials, in which the label-object pairing from habituation were switched (e.g., label A was paired with object B and vice versa). If infants notice the violation in the pairing they should look longer on Switch as compared to Same trails. In their first experiment Hay and colleagues found that 14-month-olds readily noticed the label-object violations, suggesting that they had learned the label-object mappings. In their second experiment, 17- and 19-month-old infants failed to map words differing in pitch contour to novel objects, even though they readily discriminated the pitch contours in an object-free task (Experiment 3). Hay and colleagues suggest English-learning infants undergo an interpretive narrowing by which they are initially willing to treat pitch contour contrastively, and then become more constrained in their interpretation as they gain experience that pitch contour is not relevant for word meaning.

One paradox in their results is why their 14-month-olds succeed when many studies have suggested that 14-month-old infants have a difficult time mapping consonant-based minimal pair words (e.g., *bih* and *dih*) to novel objects. One of their hypotheses is that pitch contours, which unfold over hundreds of milliseconds, represent a much more acoustically salient contrast than do voicing or place of articulation differences, which unfold over tens of milliseconds. However, their work never explicitly tests the role of acoustic salience in early label-object mapping.

Acoustic Salience of Pitch Contours

In the present study I further probed the role of acoustic salience by directly manipulating it in an associative word-learning study. I compared infants' perception of salient Mandarin tone pairs with non-salient ones. Fundamental frequency (F0) is the main acoustic marker that differentiates different tone categories in tone languages. Perfect Mandarin tone recognition (90% correct and above), by Mandarin listeners, can be achieved by keeping only the F0 of words with out the first, second or third harmonic (F1, F2, F3) (Liang, 1963). There are two main parameters that vary across different lexical tones, one is F0 contour/shape (e.g., rising or falling) and the other is F0 height/level (e.g., high, middle or low) (Howie, 1976; Tseng, 1990; Wu, 1986). Different tones also share some similar characteristics. As shown in Figure 1, Tone 2 (rising) and Tone 3 (dipping) have overlapping F0 at the onset and also display a similar trajectory. The majority of the differences come from the turning point of the contour (the time point at which the pitch contour changes from falling to rising, or vice versa). Figure 1 also shows that Tone 1 (level) and Tone 3 (dipping) have very dissimilar F0 contours. F0 height also appears to be important factor driving discriminating of different tones. The order of average F0 (from high to low) for Mandarin tones is as follows: Tone1 (level) ≈ Tone 4 (falling) > Tone 2 (rising) > Tone 3 (dipping) (Liu, Tsao, & Kuhl, 2007). According to Lu and colleagues, Tone 1 (level) and Tone 3 (dipping) also have the largest F0 height difference. The difference between Tone 2 (rising) and Tone 3 (dipping) was relatively small. Consistent with the idea that similarity may drive perceptual salience, Tsao (2008) showed that 10- to 12-month-old Mandarin-learning infants were better at discriminating Mandarin tone contrasts, Tone 1 (level) vs. Tone 3 (dipping), but were less accurate discriminating acoustically similar tone contrasts Tone 2 (rising) vs. Tone 3 (dipping) and Tone 2 (rising) vs. Tone 4 (falling). This result was partly confirmed by So and Best (2010). They found that Tone 1 (level) vs. Tone 3 (dipping), Tone 2 (rising) vs. Tone 4 (falling), and Tone 3 (dipping) vs. Tone 4 (falling) are all more easily discriminated than Tone 1 (level) vs. Tone 2 (rising), Tone1 (level) vs. Tone 4 (falling) and Tone 2 (rising) vs. Tone 3 (dipping). Acoustic salience also plays an important role for Mandarin-learning for infants; Tone pairs that are less acoustically distinctive are more often mispronounced than more salient ones (Li & Thompson, 1977). Thus, in the current experiment, I selected Tone 1 (level) vs. Tone 3 (dipping) as the salient pair and Tone 2 (rising) vs. Tone 3 (dipping) as the not salient pair.

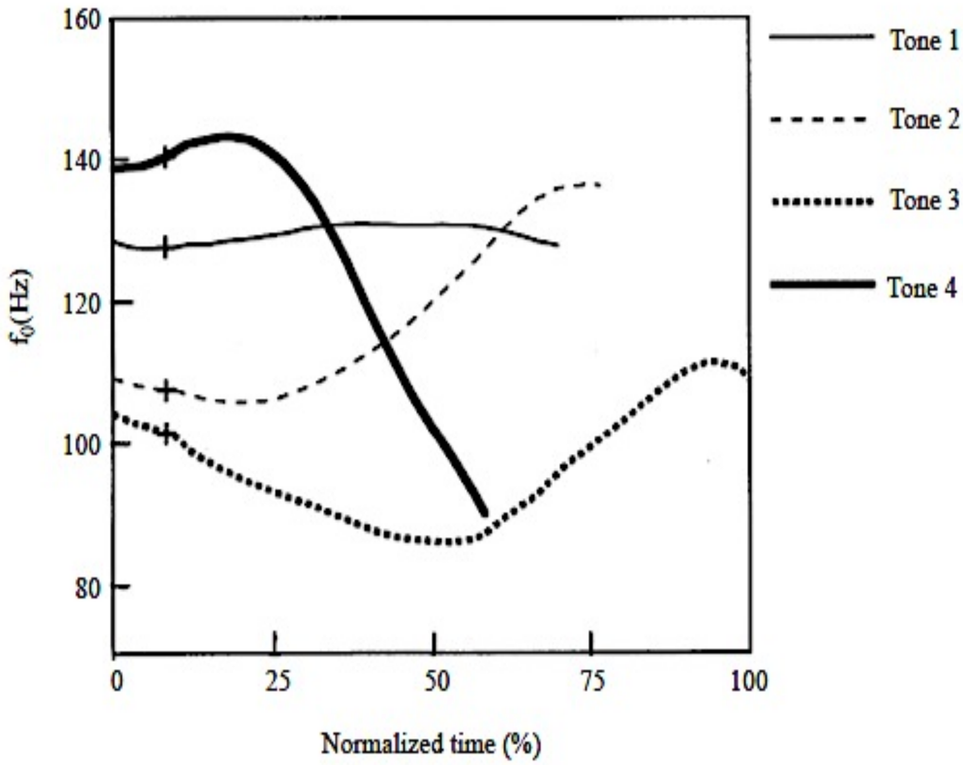


Figure 1. Pitch contour for tones in Mandarin Chinese (Xu, 1997, p. 67).

Current Study

In the current experiment, I used the Switch Paradigm to presented 13.5-month-14.5-month infants with two novel objects paired with two single syllable Mandarin nonsense words that differed only in pitch contour. In order to ensure that results are not dependent on the particular underlying consonant-vowel (CV) sequences, pitch contour were instantiated over two different CV), /kʊ/ and /di/. However each participant was only presented with a single CV sequence (i.e., /kʊ/ OR /di/). There were two conditions: Salient and Non-salient. In the Salient condition, the target CV labels were produce with salient Mandarin pitch pairs; /kʊ/ (Tone 1, level) and /kʊ/ (Tone 3, dipping) or /di/ (Tone 1, level) and /di/ (Tone 3, dipping). In the Non-salient condition the target CV labels were produced with less salient tone pairs: /kʊ/ (Tone 2, rising) and /kʊ/ (Tone 3, dipping) or /di/ (Tone 2, rising) and /di/ (Tone 3, dipping). I expected that English-learning infants should show better label-object mapping, as indicated by an increase in looking on Switch relative to Same trials, when the labels were more salient.

Chapter 2

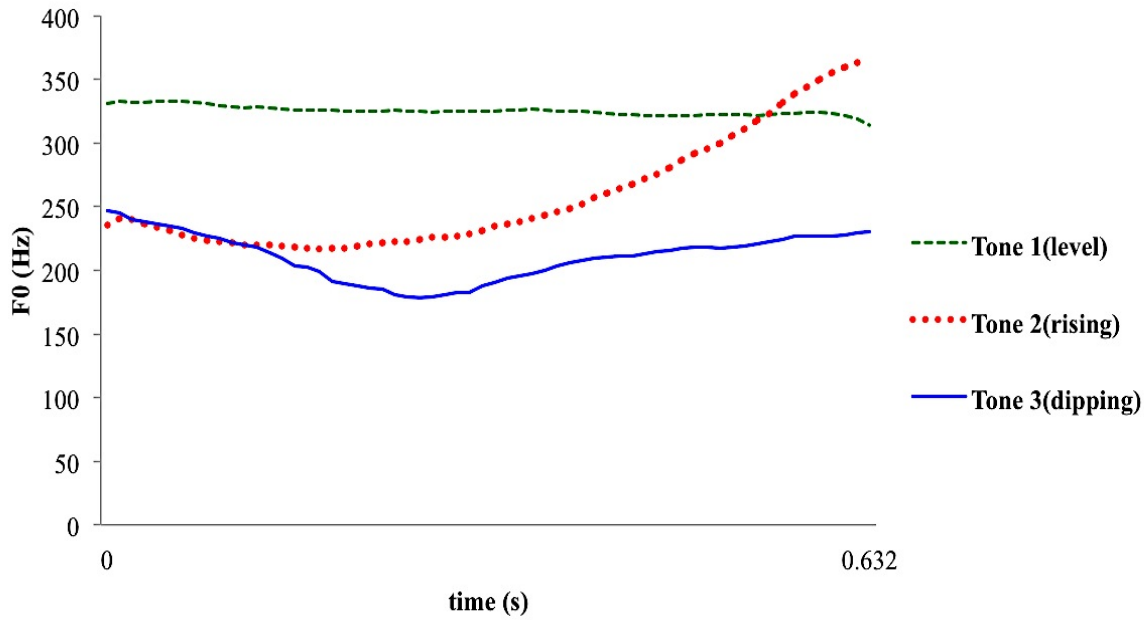
Method

Participants

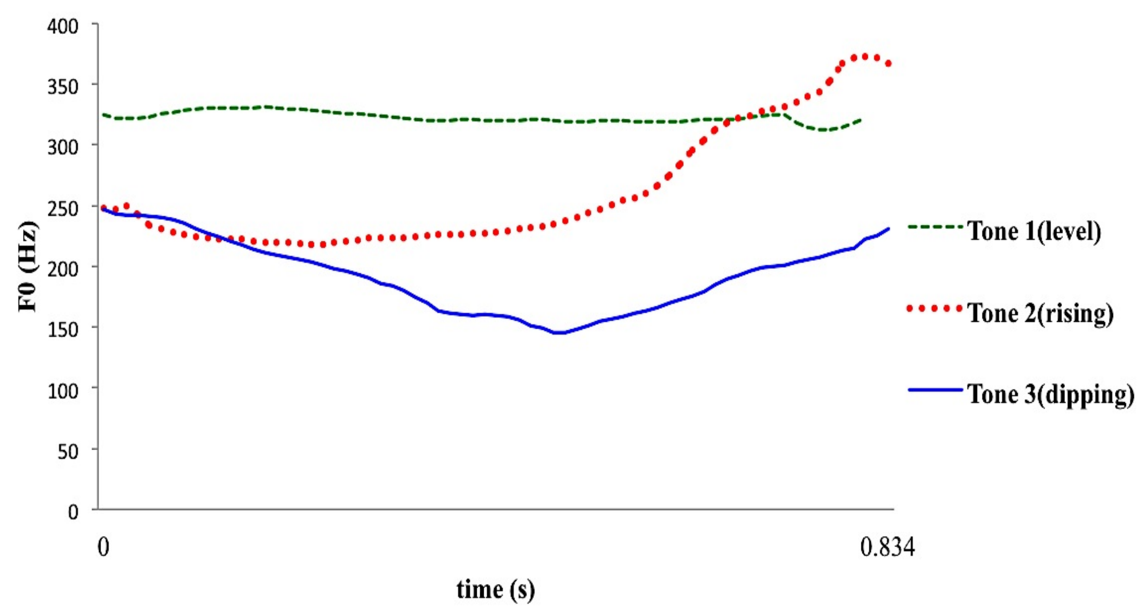
Twenty-nine 14-month-old (mean = 13.93 months, range = 13.47-14.59 months) monolingual English-learning infants participated in this study. Twelve infants were female. Participants were born full-term, had minimal if any second language exposure (i.e., foreign language expose was limited to less than one hour per week), had normal hearing and vision, and were free from any developmental delays, according to parental report. Five infants were run but were excluded from the analysis due to fussiness (n = 4) and parental interference (n = 1).

Materials

Two kinds of stimuli were included in this experiment: sound stimuli and picture stimuli. A female Mandarin native speaker pronounced the same single syllable CV nonsense word (i.e., /kʊ/ or /di/) that was phonotactically legal in both English and Mandarin, with three different pitch contours: Tone 1 (level), Tone 2 (rising), and Tone 3 (dipping)(see Figure 2). Another nonsense word, /mi/, was produced in a neutral tone and was used as pre-habituation and post-test stimulus. The sound was recorded in a soundproof booth. Sound stimuli were modified to have similar overall durations using Praat (Boersma, 2001) and were RMS matched for equal loudness in Adobe Audition 3.0 ™.



/kʊ/



/di/

Figure 2. Pitch contours used in current experiment.

Computer-generated 2D picture stimuli were used in this experiment (see Figure 3). In order to ensure that results are not dependent on the particular objects used, I chose two different object pairs for the current study. Objects 1 and 2 were always paired with each other. Similarly, objects 3 and 4 were always paired. A fifth object was used for the pre-habituation and post-test trials. On each trial, objects were presented against a grey background and bounced continuously across the screen to help maintain infants' attention.

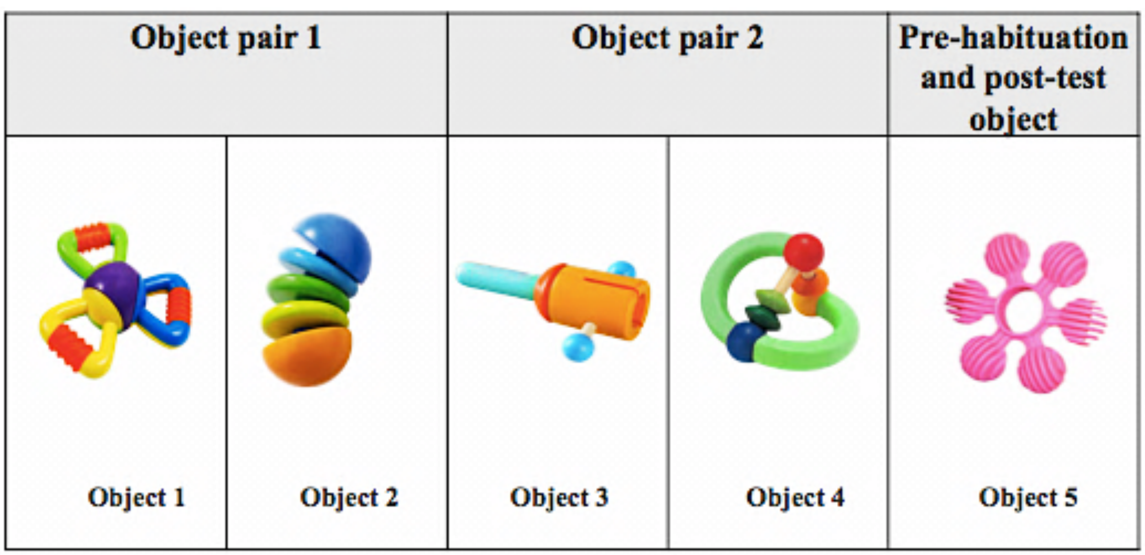


Figure 3. The objects used in the current experiment.

Apparatus

The study was conducted in a 2.3m×2.3m soundproof booth. The program Habit X 1.0 (Cohen, Atkinson, & Chaput, 2004), installed on a Mac, was used to control the experimental procedure and record infant looking time. Objects were presented on a centrally positioned 42-inch flat screen television. Two speakers hidden behind the screen were used to broadcast the object labels. A digital video camera was used to relay the visual image of infants' looking behavior to the control room. At the same time, a MacMini using iMovie was used to monitor the infants' behavior inside the booth and record the entire experiment.

Procedure

Infants were seated on a parent's lap in a sound attenuated booth, approximately 1 meter from a flat screen TV. An observer viewed infants' responses on a monitor and indicated looking times by pressing a button on the computer running Habit. To avoid potential bias, the observer was blind to the identity of the materials being presented. Parents listened to music through soundproof headphones in order to prevent them from inadvertently biases their infants' looking behavior.

I assessed object-label association using a modified version of the Switch paradigm (Werker et al., 1998). First, infants were habituated to two novel label-objects pairs. In the Salient Condition the labels were maximally acoustically distinctive: /di/ (Tone1, level) vs. /di/ (Tone3, dipping) or /kʊ/ (Tone1, level) vs. /kʊ/ (Tone3, dipping). In the Non-salient Condition the labels were minimally acoustically distinctive /di/ (Tone2, rising) vs. /di/ (Tone3, dipping) or /kʊ/ (Tone2, rising) vs. /kʊ/ (Tone3, dipping). For each label-object pair, the infants saw a 2D image bouncing on the screen, and at the same time they heard the sound used to label this object coming from the speakers.

Stimulus presentation continued until the infant looked away from the screen for more than one second or after 20 seconds had elapsed. Presentation of habituation trials was randomized by block. The whole phase ended after reaching habituation criterion (i.e., a 65% decrease in looking from the first to the last three habituation trials) or after 25 trails. Following habituation infants proceeded to the Testing phase. During the Testing phase infants were presented with two types of test trials: Same trials and Switch trials (see Figure 4). In the Same trails, the label-object pairings from the Habituation phase were maintained. In the Switch trails, the labels of the objects were switched, such that label A was presented with object B, and vice versa. If infants have learned the label-object pairing they should look long on Switch than on Same trials. Infants' attention was directed back to the screen between trials with an attention getter. The observer began each trail only after the infants looked at the attention getter. There were 8 test trials; four Switch trials and four Same trials counterbalanced in 8 different testing orders. In order to orient infants to the task infants were also presented with a pre-habituation trial where the label /mi/ was presented with an unrelated object. To ensure that infants maintained attention through the experimental procedure, the same label-object pairing was used for a final post-test trial.

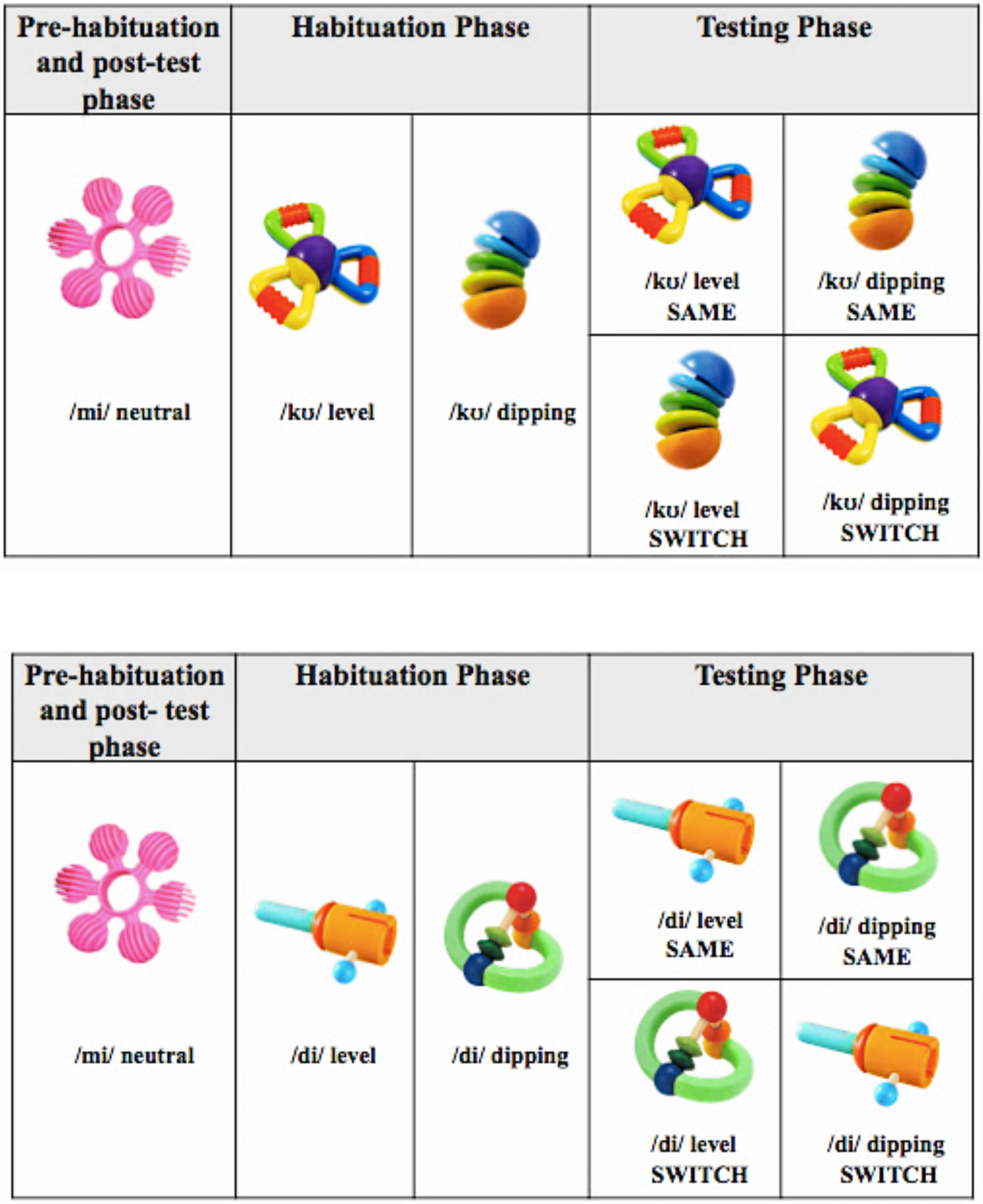


Figure 4. Example of experimental design with different labels and object pairs.

Chapter 3

Results

All the infants successfully habituated in my task. There were no differences between conditions in the number of trials to habituation (Salient = 10.6, *SD* = 5.56; Non-salient = 10.6, *SD* = 4.07), $F(1,27) = .000$, $p = .987$, $\eta_p^2 = .000$, or in the total time to habituate (Salient = 119.04 sec, *SD* = 519.49; Non-salient = 121.4, *SD* = 50.17), $F(1,27) = .007$, $p = .932$, $\eta_p^2 = .000$. This suggests that infants in both conditions showed similar levels of interest in the task. I compared the looking time of the pre-habituation and post-test trails to make sure the experiments were not influenced by the fatigue of the infants. Table 1 shows the mean looking times to pre-habituation and post-test trials and to the last block of habituation trials (i.e., the mean of last 2 habituation trails). Paired *t* test (all *t* tests reported are 2-tailed; effect sizes reported for *t* tests are Cohen's *d*) revealed no significant differences between pre-habituation and post-test looking times, $t(28) = 1.499$, $p = .145$, $d = .279$, suggesting that infants did not become fatigued by the task. A paired *t* test also was conducted to compare post-test looking times and the mean looking time of the last habituation block (average of the last two habituation trails). As expected, infants looked significantly longer to the post-test trail ($M = 15.33$, $SD = 4.85$) than to the last block of habituation trials ($M = 6.27$, $SD = 2.93$), $t(28) = -9.52$, $p = .000$, $d = 1.848$, indicating that the infants recovered from habituation during the Testing phase.

Table 1. Mean looking times (and standard deviation) by Condition for the pre-habituation and post-test trials, and to the last block of habituation trials.

Condition	Pre-habituation	Post-test	Last habituation block
Salient	16.69(4.34)	15.39(4.80)	6.55(2.64)
Non-salient	16.55(4.91)	15.64(5.15)	7.13(3.98)

Preliminary analyses revealed no significant main effects or interactions involving Gender, Label (/kʊ/ vs. /di/), or Object pair, with all *p* values larger than 0.2. Thus, all subsequent analyses were conducted collapsed over these variables.

In order to examine label-object mapping I ran a between Condition (Salient vs. Non-salient) × within Trial Type (Switch vs. Same) repeated measures ANOVA. There were no significant main effects of Condition, $F(1,27) = .022$, $p = .884$, $\eta_p^2 = .001$, or Trial Type, $F(1,27) = .215$, $p = .647$, $\eta_p^2 = .008$, however, there was a significant Condition X Trial Type interaction, $F(1,27) = 9.462$, $p = .005$, $\eta_p^2 = .26$. Follow-up paired sample *t* tests revealed that infants in the Non-salient condition looked significantly longer to Switch ($M = 8.95$, $SD = .93$) than Same trials ($M = 6.99$, $SD = .54$), $t(13) = 3.123$, $p = .008$, $d = 1.037$. There were no significant differences in looking between Switch ($M = 7.42$, $SD = .83$) and Same ($M = 8.86$, $SD = 1.10$) trials for infants in the Salient condition, $t(14) = -1.617$, $p = .128$, $d = .433$.

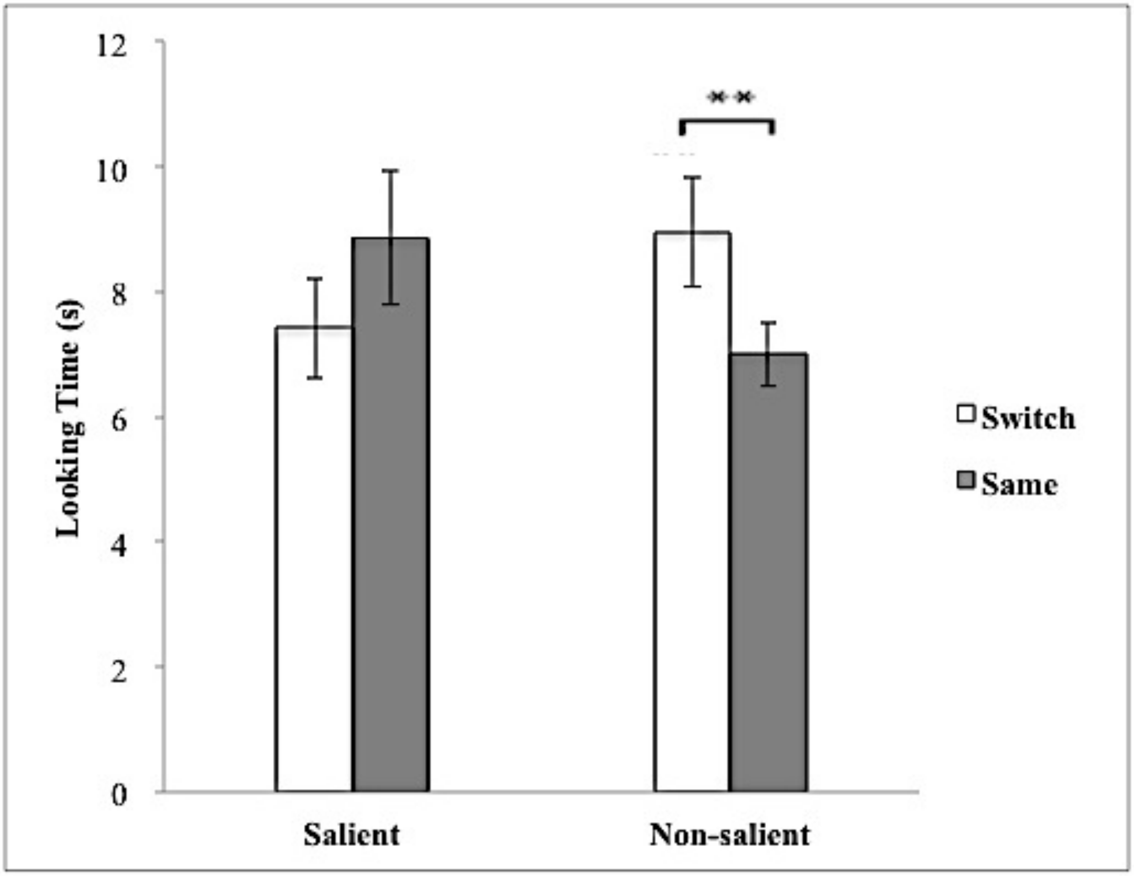


Figure 5. Mean looking times to Same and Switch trials for infants in the Salient and Non-salient conditions. Error bars represent standard error of the mean.

Chapter 4

Discussion

This experiment was designed to explore the role of acoustic salience in word mapping. Although 14-month-old infants have a notoriously difficult time mapping minimal pair words to novel objects (Stager & Werker, 1997), these effects may be partially mitigated if the contrast is sufficiently acoustically salient. Specifically, I predicted that 14-month-old infants would more easily map labels that were more acoustically distinctive onto novel objects than those that were less acoustically distinctive. Infants were taught two object label pairs where the objects labels differed only in pitch contour. Contrary to my predictions, infants demonstrated more successful learning of the 'Non-salient' (rising vs. dipping) label pairs than the 'Salient' (level vs. dipping) label pairs.

At first glance, the results do not appear to support my hypothesis. Thus, one conclusion that must be considered is that acoustic salient might not be a relevant or significant factor in early word learning. There are a number of previous studies that have failed to demonstrate word learning even when the contrasts are acoustically salient. First, work by Pater and colleagues (Pater et al., 2004) found that the number of features that vary in minimal pair words has no effect on minimal pair word learning at 14 months. Specifically, infants have as much difficulty mapping words that minimal contrast in both voicing and place of articulation (e.g., *pin* vs. *din*) to novel objects as words that vary in a single feature (i.e., voicing; *pin* vs. *bin*). Further, 12-month-old infants fail to map acoustically distinctive communicative non-speech sound (e.g. *ooh, ssh*) (MacKenzie et al., 2011) to novel objects and non-native the words that are distinctive but violate native-language phonotactic patterns (e.g. Czech word *ptak*; MacKenzie et al., 2012), despite the acoustic distinctiveness of both types on contrasts. Together, there studies suggest that acoustic salience may not be the sole criteria for determining whether infants are able to map novel labels to objects.

Although other studies have also failed to see word mapping of acoustically salient contrasts, my results do suggest that the acoustic characteristics of sounds affect interpretive-narrowing during early object label mapping. Presumably, infants learning tone language should maintain the ability to recognize and map different tones to different objects throughout development. However, for infants learning non-tone languages, it is adaptive to begin to ignore pitch contour as a cue to word meaning. Studies on perceptual narrowing provide support for the effects of language experience on the perception of pitch contour, a process that may begin as early as 4 months of age (Francis, Ciocca, Ma, & Fenn, 2008; Lee, Vakoch, & Wurm, 1996; So & Best, 2010; Yeung, Chen, & Werker, 2013). For word learning, there are conflicting results regarding when non-tone language infant begin to ignore pitch contour. Hay and colleagues (2014) suggest that this process begins between 14 and 17-19 months of age, whereas Singh and colleagues (2014) demonstrate attention to pitch contour in novel words in infants as old as 18 months of age, with declines seen by 24 months. Differences in experimental design may account for differences between these two studies; whereas Hay and colleagues (2014) used a task very similar to the current task, Singh and colleagues used a task that provided considerably more referential support which may have supported learning in older infants. Never the less, the current study suggests that declines in attention to pitch contour during word learning do not occur at the same time for every pitch contour contrast. Instead, some contrasts appear to be more resilient during early label-object association.

What makes one contrast more resilient to interpretive narrowing is still an open question. It is possible that acoustic salience is an important or at least an influential factor during word learning, but that my definition of acoustic salience may not have been appropriate to the task at hand. Contrary to previous discrimination research suggesting the Tone 1 (level) and Tone 3 (dipping) distinction is more acoustically salient than the Tone 2 (rising) and Tone 3 (dipping) distinction, my participants appear to map the non-salient pitch pairs to novel objects more readily than the salient pitch pairs. Thus, what is considered salient in a discrimination task, may differ from what is treated as salient in a word-learning task. In the current experiment, I defined salience according to previous research on acoustic measurements of F0 contour/shape and F0 height/level (Howie, 1976; Tseng, 1990; Wu, 1986) and on the discriminability of the contrasts (Liu et al, 2007; Tsao, 2008; So & Best, 2010). However, the salience of a given pitch contour contrast during word learning might not only be based on the pitch contour itself, but may also include other factors, such as the native language experience with pitch at the prosodic level. For example, rising F0 is more likely to be used to attract infants' attention by the mothers or caregivers (Ferrier, 1985; Fernald, 1992) and variation in pitch contour is more common in infant-directed-speech (IDS) than adult-direct-speech (ADS) (Cooper & Aslin, 1990; Fernald, 1985; Pegg, Werker, & McLeod, 1992; Schachner & Hannon, 2011). Further, questions tend to be marked by a rise in pitch, which may make rising pitch contours particularly salient (van Heuven & Haan, 2002). As early as 6 months of age, infants are very sensitive to distinctions between prosodic categories across multiple languages (Best, Levitt, & McRoberts, 1991) and can separate questions from statements not only in their native language, but also in another non-native language. Thus, pitch contours that contain variability in F0 may be more likely to draw infants' attention than level pitch contours. This variability in F0 might help the infants encode more details about the novel words, and thus make the word mapping easier to accomplish.

Variability in F0 might also add to the richness of information contained in each of the object labels and may thus have affected how easily the labels were mapped to novel objects. Shannon (1948) developed Information Theory to measure the transmission of information through noise. According to this theory, transmission of information between transmitters and receivers is based on the amount of uncertainty, or entropy, in the signal; information in the signal relies on the transmitters and receivers agreeing on the intended message, which is most likely to occur under conditions of minimal uncertainty. According to Shannon (1948), the larger the entropy the less the uncertainty and thus the more information is available in the signal. Thus, it is possible that the pitch contours used in the non-salient condition contained more ‘information’ than those used in the salient condition,

Additionally, the structure of labels themselves may have affected how easily the labels were mapped onto novel objects. Mandarin words are notably simpler than the types of words commonly faced by the English-learning infants in their daily lives. The words used in the current experiment were similar to English function words (e.g., the) in that they were monosyllabic and had some vowel reduction. Compared to content words, function words are less salient, acoustically and phonologically. Acoustically, function words are shorter, more likely to be reduced and normally are not stressed in sentences. Phonologically, they normally do not have complex syllable structures. Shi and colleagues (Shi et al., 1998; Shi & Werker, 2003) have demonstrated that function words do not attract infants’ attention to the same extent as content words. Further, at 12 months of age infants associate novel objects with content-like words (e.g., *fep*) but not with function-like words (e.g., *iv*) (MacKenzie et al. 2012). So at the beginning of language learning, the infants may pay more attention to content words, and may be more willing to treat them as object labels as compared to function words. It is possible that infants in the Salient condition may have been more successful in my task if I had used more complex word forms, more similar to English content words.

Instead of relying on unclear definitions of acoustic salience to drive predictions about flexibility in early word learning, perhaps my results may be better understood in terms of the role that stimulus complexity plays during early word learning. The more complex the word is, the more information may be available to the infants during word label-object mapping. This complexity might include several factors including the complexity of the F0 contour, the acoustic characteristics of the sounds, native language experience, and the amount of uncertainty/information in the signal (entropy).

Chapter 5

Conclusions

In the current experiment I demonstrated that 14-month-old infants could map two labels that are minimally different in their pitch contours onto two different objects. However, my prediction that they would more easily learn to associate level and dipping pitch contours with novel objects than rising and dipping contours was not confirmed. Instead, infants appear to treat the rising versus dipping pitch contour distinction as more lexically contrastive than level and dipping contrast. My findings suggest that the concept of acoustic salience may depend, at least partially, on the task demands (i.e., discrimination vs. word learning). At 14 months, young language learners have already gained some knowledge regarding what makes a good word in their language. However, they are still in the process of interpretive narrowing. Infants are faced with an enormous amount of information; to be efficient word learners they need to analyze and attend to the most relevant information and ignore irrelevant variation. The results of current research support the idea that word learning is a complex process whereby infants must consider numerous factors, including acoustic characteristics, information richness, and native language experience in determining what makes a good word. Further research needs to be done to determine the relative importance of various factors in early word learning.

References

Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development, 59,* 450–466.

Best, C.T., Levitt, A., & McRoberts, G.W. (1991). Examination of language-specific influences in infants' discrimination of prosodic categories. *Proceedings of the XIIth lnternational Congress of Phonetic Sciences, 4*, 162-165.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

Campbell, A. L., & Namy, L. L. (2003). The role of social referential cues in verbal and non-verbal symbol learning. *Child Development, 74*, 549–563.

Chao, Y. R. (1948). *Mandarin Primer.* Cambridge, UK: Harvard University Press.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0) [Computer software]. Austin: University of Texas.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*(5), 1584–1595.

Curtin, S., Fennell, C. T., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word-object associative learning. *Developmental Science, 12,* 725–731.

Fennell, C. T., & Waxman, S. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81*, 1376–1383.

Fennell, C. T. (2012). Object familiarity enhances infants' use of phonetic details in novel words. *Infancy, 17*(3), 339–353.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*(2), 181–195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10,* 279–293.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Development, 60*(6), 1497-1510.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, H. Jurgens, U., & Papousek, M. (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches*. Cambridge, UK: Cambridge University Press.

Ferrier, L. J. (1985). Intonation in discourse: talk between 12-month-olds and their mothers. In K. Nelson (Ed.). *Children's language. 5,* 35-60.

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

Graf Estes, K. M., Edward, J., & Saffran, J. R. (2011). Phonotactic constraints on infant word learning. *Infancy, 16*(2), 180–197.

Graf Estes, K. M., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy, 18*(5), 797-824.

Gunzburger, D., Bresser, A., ter Keurs, M. (1987). Voice identification of prepubertal boys and girls by normally sighted and visually handicapped subjects. *Language and Speech, 30*, 47-58.

Gussenhoven, C. (2004). *Phonology of Tone and Intonation.* Cambridge, UK: CUP.

Hay, J. F., Graf Estes, K., Wang, T. and Saffran, J. R. (2014), From Flexibility to Constraint: The Contrastive Use of Lexical Tone in Early Word Learning. *Child Development*. doi: 10.1111/cdev.12269

Heuven, V. J., & Haan, J. (2002). Temporal development of interrogativity cues in Dutch. In C. Gussenhoven & N. Warner (Eds.), *Papers in laboratory phonology VII,* 61– 86. Berlin: Mouton de Gruyter.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Re-search in Child Development, 65*(3, Serial No. 262).

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones.* Cambridge, UK: Cambridge University Press.

Iverson, J. M., Capirci, O. & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development 9,* 23-43.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Jeanine, M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*, 402-420

Kuhl, P. K. (1987). *Handbook of Infant Perception.* New York: Academic Press.

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

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*(5), 527–542.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin – speaking children. *Journal of Child Language*, *4*, 185-199.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant- directed speech. *Developmental Psychology*, *43*, 912-917.

MacKenzie, H., Graham, S. A., & Curtin, S. (2011). Twelve- month-olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science*, *14*, 249–255.

MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-month-olds' phonotactic knowledge gruides their word-object mappings. *Ghild Development*, *83*, 1129-1136.

Moore, D. S., Spence, M. J., & Katz, G. S. (1997). Six-month-olds' categorization of natural infant-directed utterances, *Developmental Psychology, 33,* 980-989.

Namy, L. L., & Waxman, S. R. (1998). Words and symbolic gestures: Infants' interpretation of different forms of symbolic reference. *Child Development, 69*, 295-308.

Namy, L. L. & Waxman, S. R. (2000). Naming and Exclaiming: Infants' Sensitivity to Naming Contexts. *Journal of Cognition and Development, 1*(4), 405–428.

Namy, L. L. (2001) What's in a name when it isn't a word ? 17-month-olds' mapping of non- verbal symbols to object categories. *Infancy 2*, 73-86.

Pater, J., Stager, C. L., & Werker, J. F. (1998). Additive effects of phonetic distinctions in word learning. *Proceedings of the 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America,* 2049-2050. Sewickley, PA: Acoustical Society of America Publications.

Pater, J., Stager, C. L., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*, 384–402.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: evidence from 7-week-old infants. *Infant Behavior and Development, 15*(3), 325–345.

Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Develop mental Psychology, 47*(1), 19–25.

Shi, R., Morgan, J., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language, 25,* 169 –201.

Shi, R., & Werker, J. F. (2003). The basis of preference for lexical words in 6-month-old infants. *Developmental Science, 6*(5), 484-488.

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science, 17*(1), 94-109.

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech, 53*(2)*,* 273-293.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321–1337.

Tsao, F. M. (2008). The effect of acoustically similarity of lexical-tone perception of one-year-old Mandarin-learning infats. *Chinese Journal of Psychology, 50*(2)*,* 111-124.

Tseng, C. Y. (1990). *An acoustic phonetic study on tones in Mandarin Chinese*. Taipei, Taiwan: Institute of History & Philology Academia Science.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associa- tions by 14-month old infants. *Developmental Psychology, 34,* 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. *Infancy*, *3*, 1–30.

Werker, J. F., & Tees, R. C. (2002). Cross- language speech perception: Evideence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 25,* 121-133.

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

Woodward, A. L., & Hoyne, K. L. (1999). Babies' learning about words and sounds in relation to objects. *Child Development, 70,* 65-77.

Wu, Z. J. (1986). *The Spectrographic Album of Mono-syllables of Standard Chinese*. Beijing, China: Social Science Press.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25,* 61–83.

Yeung, H. H., Chen, K. H., & Werker, J. F.(2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and language, 68*, 123-139.

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

VITA

Qian Zhao was born in Dalian China. She obtained Bachelor of Science degree in Psychology back in summer 2009 from Sun Yat-sen University, Guangzhou China. She then continued education at University of Tennessee Knoxville working on a Master of Science Degree in Experimental Psychology. She has been conducting research in Prof. Hay's lab. She wrote a thesis based on the research she has done in the field of infant word mapping. She is planning to graduate summer 2014 with a Master of Science Degree in Experimental Psychology from University of Tennessee Knoxville.