The Perceptual Draw of Prosody: Infant-Directed Speech within the Context of Declining Nonnative Phoneme Perception

Wendy L. Ostroff

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Psychology

Robin Panneton Cooper, Chair
David W. Harrison
Martha Ann Bell

October 7, 1998
Blacksburg, Virginia

Keywords: Infant-Directed Speech, Speech Perception, Infant Attention, Phonemic Perception
The Perceptual Draw of Prosody: Infant-Directed Speech within the Context of Declining Nonnative Phoneme Perception  
Wendy L. Ostroff

(ABSTRACT)

Infant speech perception develops within the context of specific language experience. While there is a corpus of empirical evidence concerning infants’ perception of linguistic and prosodic information in speech, few studies have explored the interaction of the two. The present investigation was designed to combine what is known about infants’ perception of nonnative phonemes (linguistic information) with what is known about infant preferences for ID speech (prosodic information). In particular, the purpose of this series of studies was to examine infant preferences for ID speech within the timeline of the phonemic perceptual reorganization that occurs at the end of the first postnatal year. In Experiment 1, 20 Native-English 10- to 11-month-old infants were tested in an infant-controlled preference procedure for attention to ID speech in their native language versus ID speech in a foreign language. The results showed that infants significantly preferred the ID-native speech. In Experiment 2, the preferred prosodic information (ID speech) was separated from the preferred linguistic information (native speech), as a means of discerning the relative perceptual draw of these types of speech characteristics. Specifically, a second group of 20 10- to 11-month-old infants was tested for a preference between ID speech in a foreign language and AD speech in their native language. In this case the infants exhibited a significant preference for ID-foreign speech, suggesting that prosodic information in speech has more perceptual weight than linguistic information. This pattern of results suggests that infants attend to linguistic-level information by 10- to 11-months of age, and that ID speech may play a role in the native-language tuning process by directing infants’ attention to linguistic specifics in speech.
Dedication

This project is dedicated to Rob Genova—
for many months of unconditional support
and an incessant capacity to listen
Author’s Acknowledgments

Many thanks Jamie Cooper, Carrie Green and everyone in the lab for help with data collection. Very special thanks to Dr. Robin Cooper, whose endless enthusiasm, dedication and energy are highly contagious.
Table of Contents

Chapter 1-Introduction and Literature Review ____________________________________________1

Infant Attention to the Linguistic Aspects of Speech ________________________________1
Infant Attention to Prosodic Aspects of Speech ________________________________3
The Relationship between Phoneme Perception and ID Speech Preference__________7

Chapter 2-Experiment 1: The Relationship between Phoneme Perception and ID Speech Preference__8

Method _______________________________________________________________9
Results and Discussion __________________________________________________15

Chapter 3-Experiment 2: Preference for ID-Foreign Speech over AD-Native Speech ____________19

Method _______________________________________________________________20
Results and Discussion __________________________________________________22

Chapter 4-General Discussion _________________________________________________________25

References ________________________________________________________________________31
Appendices ________________________________________________________________________37
Vita ______________________________________________________________________________43
Multimedia Objects

Table 2.1
Matched Utterances for Experiment 1 Stimulus Tapes: ID-English with ID-Foreign______11

Figure 2.1
Schematic of the Auditory Preference Protocol______________________________14

Figure 2.2
Study 1: Mean Infant Looking Times: ID-English Versus ID-Foreign _________________16

Table 3.1
Matched Utterances for Experiment 2 Stimulus Tapes: AD-English with ID-Foreign______21

Figure 3.1
Study 2: Mean Looking Times: AD-English Versus ID-Foreign_______________________23
Chapter 1
The Perceptual Draw of Prosody: Infant-Directed Speech within the Context of Declining Nonnative Phoneme Perception

The language input that infants encounter has important consequences for the development of speech perception. Because infants experience speech in the form of prosodic (rhythm, intonation) and linguistic (words, syntax) information, defining the relative prominence of these features is a key step toward discerning how the speech environment can afford the perception of language. Generally speaking, there is a considerable amount of evidence concerning infants’ sensitivity to both the linguistic and the prosodic elements of speech. However, these bodies of research have not been melded together to form a cohesive account of speech perception. The present investigation was designed to take a step in that direction, by combining what is known about infants’ perception of nonnative phonemes with what is known about their preference for exaggerated prosody. In particular, this pair of studies examined the interaction between infants’ perceptual tuning to native phoneme contrasts, and their preference for infant-directed (ID) speech.

Infant Attention to the Linguistic Aspects of Speech

The linguistic component of speech consists of phonemes (sounds), and words which phonemes combine to form. Segments of sound must become linguistically discrete if they are to be systematically used to convey differences in word meanings (Best, 1994). Since understanding language involves the eventual matching of words with the objects and events that they represent, a system of rules for sound combinations must be understood for meaning to be attached (Aslin, Jusczyk & Pisoni, 1996). This is quite a feat, considering that the acoustical form of every utterance varies tremendously across and within individual speakers, rates of speech, and contexts (Werker, 1994).

Adult language-users overcome such variability by perceiving speech categorically
(Aslin et al., 1996). That is, they hear slightly differing sounds as equivalent (part of a single ‘category’), and can distinguish those sounds from the sounds of another category. It has been demonstrated in the laboratory that adults fail to discriminate different native speech sounds that are adjacent along an acoustic continuum, if those sounds come from one perceptual category. However, adults are quite proficient at discriminating adjacent sounds that straddle two different perceptual categories (Liberman, Harris, Hoffman & Griffith, 1957).

Interestingly, infants seem to perceive speech categorically well before they have begun to produce speech. A classic study by Eimas, Siqueland, Jusczyk and Vigorito (1971) found that 1- to 4-month-olds showed superior discrimination for phonetic contrasts straddling two different perceptual categories (as opposed to contrasts from within the same category). This finding has been replicated several times using various sound distinctions and in various contexts (for reviews, see Aslin, 1987; Jusczyk, 1985).

To look for the developmental origin of infant categorical phoneme sensitivity, several researchers began examining cross-language speech perception (Werker & Lalonde, 1988). Naturally, if a human infant is to learn a native language (and its set of sound categories), his/her perceptual system must initially be equipped to acquire the universal range of phonetic combinations used in any of the world’s languages. A seminal study by Streeter (1976) demonstrated that young Kikuyu-learning infants could discriminate an English voiced/voiceless distinction even though it does not phonemically exist in their language. Since then, a corpus of research has shown that infants under 6-months of age do not need specific language experience to discriminate phonetic comparisons—whether or not they are contrasted phonologically in their native language (Aslin, Pisoni, Hennessy & Perey, 1981; Best, McRoberts & Sithole, 1988; Eimas, 1975; Lasky, Syrdal-Lasky & Klein, 1975; Trehub, 1976; Werker, Gilbert, Humphrey & Tees, 1981; Werker & Lalonde, 1988, Werker & Tees, 1984). However, if infants are to separate out the sound categories and sound patterns appropriate to their native language, this universal capacity must eventually be influenced by language experience (Aslin et al., 1996).

Unlike infants, adults do have considerable difficulty categorizing and discriminating contrasts that do not exist in their native language (Pisoni, Lively & Logan, 1994; Tees & Werker, 1984;
Trehub, 1976; Werker, 1994; Werker & Tees, 1984). For example, monolingual Japanese and Korean speakers have problems distinguishing the English /r/ and /l/ sound contrast (Sheldon & Strange, 1982). Although the decline in ease of discrimination seems to be more reorganization than loss (Diamond, Werker & Lalonde, 1994), specific linguistic experience clearly constrains the perception of nonnative contrasts in some way.

In the early 1980’s Werker and her colleagues discovered that young children also have trouble discriminating nonnative phonemes, and began examining infancy for changes in this capacity. Through a series of studies Werker and Tees (1984) revealed that the decline in sensitivity to nonnative speech patterns occurs by the latter half of the first postnatal year. Infants were tested for their ability to distinguish between two phonemes sounding almost alike, but from different ‘perceptual categories’. The stimuli were from three languages: the English contrast /ba/ vs. /da/, the Hindi contrast /ta/ vs. /tɑ/, and the Nthlakapmx contrast /k’i/ vs. /q’i/. At the age of 6-8 months, English-reared infants were able to discriminate all three contrasts. At 8-10 months, only a small percentage of English-reared infants were able to discriminate the nonnative contrasts, and at 10-12 months, English-reared infants were performing as poorly on the nonnative contrasts as native-English adults (Werker & Tees, 1984).

This decline in nonnative sensitivity has been replicated using a different procedure (Best, 1993), as well as a longitudinal sample (Werker & Tees, 1984). More recently this effect was also replicated using synthetically produced stimuli (Werker & Lalonde, 1988). Taken together these results point to a robust developmental phenomenon: listening experience, within the first year of life (and perhaps prenatal life as well) facilitates a reorganization in perceptual capabilities which is evident by ten months of age. This reorganization manifests itself in infants’ declining abilities to discriminate sounds not represented in their native language.

Infant Attention to Prosodic Aspects of Speech

Prosody is the melody of the speech signal. The rhythmic quality of a voice holds “emotion-identifying” information that infants need and appear to enjoy (Locke, 1993, p. 142). Thus, it seems logical that adults exaggerate the prosody in the speech they direct toward infants.
Prosody is important to the study of speech perception because infants appear to understand the message in the tone of an utterance before they can understand the specific meaning of the words. More directly, it has recently been argued that infant-directed (ID) speech enhances infants’ perception of language-relevant information (Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg & Lacerda, 1997).

Adults from all parts of the world modify their speech in distinctive ways when they address preverbal infants. Specifically, they elevate the pitch, elongate the vowels, expand the pitch contours and increase the rhythmicity of their speech (Cooper, Abraham, Berman & Staska, 1997; Fernald, 1985; Fernald & Kuhl, 1987; Papousek, Papousek & Bornstein, 1985; Werker & McLeod, 1989). The use of ID speech has been documented in adults of both sexes (Jacobson, Boersma, Fields & Olson, 1983), parents as well as nonparents (Gleason, 1975), and children (Sachs, 1977). Likewise, its practice has been recognized within a variety of cultures, and in several types of languages including German (Fernald & Simon, 1984), Italian, British-English, American-English, French, Japanese (Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies & Fukui, 1989), and Mandarin (Grieser & Kuhl, 1988).

Moreover, behavioral preferences for ID speech over adult directed (AD) speech have been empirically exhibited throughout the first year of life: in newborns and 1-month-olds (Cooper & Aslin, 1990; Cooper et al., 1997), 7-week-olds (Pegg, Werker & McLeod, 1992), 4-month-olds (Cooper et al., 1997; Fernald, 1985; Werker & McLeod, 1989), and 7-9-month olds (Glenn & Cunningham, 1983; Werker & McLeod, 1989). Because ID speech is highly accessible to the infant listener, it has become important to study whether experience with such speech enhances and/or induces some aspect of the infant’s development (Cooper & Aslin, 1990). Three categories of functional value have been proposed: (1) the engagement and maintenance of arousal and attention, (2) the regulation of the infant’s emotion and/or the communication of the speaker’s emotion, and (3) the facilitation of language perception (see Fernald, 1992 for discussion).
Several researchers have also noted that the way adults address infants changes as a function of age and/or developmental stage (e.g., Ratner, 1996; Snow, 1977; Stern, Speiker, Barnett, & MacKain, 1983). Consequently, claims have been made that the role of ID speech in the infant’s development changes with time. Fernald (1992) proposed that newborns are biologically predisposed to respond to ID speech. As their perceptual capabilities improve over the first few months of life, ID speech becomes effective in eliciting and maintaining arousal and attention. Later, when multimodal displays (e.g., face and voice) become increasingly integrated and understood by the infant, the role of ID speech becomes one of socio-emotional understanding. At this point the infant can use pitch contours to discern the intended meaning of the speech (e.g., Fernald, 1993; Ferrier, 1985; Papousek, Papousek & Symmes, 1991; Stern, Speiker & McKain, 1982).

Ultimately, the prosody of ID speech comes to impart specific linguistic information to the infant, which may facilitate the acquisition of language. In a series of experiments by Karzon (1985), 1- and 4-month-old infants were able to discriminate between the phonetic contrasts embedded in the three syllable words \textit{malana} and \textit{marana} only when the middle syllable was accentuated by ID speech. The infants failed to discriminate the two words when they were spoken in adult-directed (AD) speech, when they were spoken flatly (i.e., with three equally stressed syllables), and when the first syllable was highlighted with ID speech (i.e., higher fundamental frequency, longer duration and greater amplitude). In other words, in order for the infants to make the phonemic distinction, the ID speech modifications had to coincide with the target syllable. Karzon concluded that prosodic characteristics act as important cues in the perception of speech by drawing attention to sound differences that need to be discriminated.

Other studies have also shown that prosody in speech may cue infants to the organization of linguistic information. In fact, data suggests that infants are sensitive to acoustic correlates of syntactic units within sentences. For instance, Hirsh-Pasek and her colleagues noted that the boundaries between clauses in sentences are often marked by prosodic change. They examined whether infants can use those acoustical markers to identify units of speech following the grammatical rules of clauses (Hirsh-Pasek, Kelmer Nelson, Jusczyk, Wright Cassidy, Druss, &
Kennedy, 1987). In a head-turn preference procedure the researchers presented 7- to 10-month-old infants with a choice between ID speech paragraphs with pauses inserted at the clause boundaries (as they are normally), and ID speech paragraphs with pauses inserted in spots not coincidental with the clauses. The infants displayed a preference for the speech with pauses in the grammatically correct places. This preference was observed even when all other naturally existing pauses in both speech streams were removed—indicating that the infants did not simply prefer the speech with a greater amount of pauses. In a follow-up study, Kelmer Nelson, Hirsh-Pasek, Jusczyk and Wright Cassidy (1989) discovered that 7- to 10-month-olds did not show the same preference for speech with coincident pauses if the paragraphs were spoken in AD speech—evidence that ID speech functions to ease the task of language learning. In this instance, ID speech brackets the speech input into the meaningful units needed for the acquisition of a grammar.

In another investigation by Jusczyk, Hirsh-Pasek, Kelmer Nelson, Kennedy, Woodward and Piwoz (1992) speech segmentation was studied by systematically examining infants’ sensitivity to prosodic markers of *phrases* in sentences. Again, using a head-turn preference procedure, Jusczyk and his colleagues conducted a series of experiments. The major finding that emerged was that 9-month-old infants listened significantly longer to sentences interrupted at the appropriate phrase boundary, than they did to sentences interrupted within a phrase. This pattern of results was observed in five groups of infants and with two sets of stimuli. The pattern was also observed when the speech samples were low-pass filtered\(^1\) to eliminate linguistic content (Jusczyk et al., 1992).

Apparently exaggerated prosody also helps infants recognize the sound patterns of familiar words. In a series of studies using a head turn preference procedure Jusczyk and Aslin (1995) showed that 7 ½-month-old infants could recognize familiar words when they were embedded within the context of an ID speech stream. Moreover, Fernald, McRoberts and
Herrera (cited in Werker et al., 1996) discovered that 15-month-old infants displayed better comprehension for videotaped objects accompanied by words when the words were spoken in ID speech. By 19-months of age the infants comprehended equally well with ID and AD words, indicating that ID speech might assist in word perception at a time when the infant is ‘tuning in’ to the specific linguistic characteristics of their native language.

Furthermore, Fernald and Mazzie (1991) discovered that mothers of 14-month-olds utilize specific patterns of prosody to highlight featured words for their infants. In particular, when asked to emphasize target words in a story to their infants, the women tended to place those words on the peak pitch location within the sentence, as well as to put them at the end of the utterance. When telling the same story to adults, the women treated target words much less predictably. These authors speculated that caretakers capitalize on infants’ perceptual biases

\[\text{When speech is low-pass filtered, all frequencies below the set point of the filter are 'passed', whereas frequencies above this set point are eliminated. Generally, low-passed speech contains pitch information (i.e., F}_0), but few higher harmonics, making the words more difficult to understand.}\]
The Relationship between Phoneme Perception and ID Speech Preference

Because ID speech functions to highlight relevant linguistic aspects of the speech signal, it would seem likely that the preference for ID speech in a foreign language would decline over the first year of life, as does the ability to discriminate nonnative phonemes. Interestingly, this does not appear to be the case. Recently, Werker, Pegg and McLeod (1994) observed both attentional and affective preferences in 4.5- and 9-month-olds for Cantonese ID speech over Cantonese AD speech even though the infants were from native English backgrounds. Likewise, Best, Calderon, Jones, Avery, Correa and McRoberts (1997) found attentional preferences by native-English infants for both Japanese ID speech over AD speech, and Xhosa ID speech over AD speech. Again, the preference was exhibited at 4-6- and 8-10-months of age. This persistent preference for ID-speech appears to reflect a “socially relevant, language-general bias” which stays intact, despite any language-specific tuning (Werker, et al., 1994, p. 330).

Still, as the words contained in a native language become important, infants’ universal sensitivity to nonnative speech sounds should become increasingly unnecessary. Thus, attention should become selectively focused on native language characteristics, causing a reorganization in the capability to distinguish nonnative categories. According to this logic, it is also expected that the perceptual draw of ID prosody should diminish enough to allow infants to attend more to linguistic information in caretakers’ speech. In other words, although exaggerated tone, pitch and rhythm may initially serve to draw infant attention toward specific sounds and words of the native language (thereby aiding in the reorganization process)—at some point in language development the words likely take precedence. However, we know relatively little about changes in the perceptual draw of prosodic and linguistic information in speech directed toward infants.

Chapter 2

Experiment 1: Preference for ID-Native Speech

Over ID-Foreign Speech
This investigation was designed to view the function of ID speech within the timeline of perceptual reorganization of native and nonnative phonetic sensitivity. Infants were tested at 10- to 11-months of age, based on seemingly disparate findings that infants this age can no longer discriminate nonnative phonemes (Werker & Tees, 1984), and yet they still prefer ID speech in nonnative languages (Best et al., 1997; Werker, Pegg, & McLeod, 1994). These previous studies have all examined infants’ preferences for nonnative ID speech when compared to nonnative AD speech. The finding that infants preferred nonnative ID speech under these conditions speaks to the perceptual salience of ID prosody, but does not address the issue of whether infants were also attending to information at the linguistic level. This study was designed to present infants with the prosody of ID speech within the context of both a familiar and an unfamiliar language, but with their native linguistic structure only available in one case. Specifically, in the first experiment (Experiment 1) 10- to 11-month-olds were tested for a preference between ID speech in their native language (English) and ID speech in a foreign language. Given that infants of this age are becoming perceptually attuned to their native language linguistic information, it was hypothesized that infants would prefer ID speech in their native language (i.e., English).
Method

Participants

A database of local parents and infants (compiled from hospital and newspaper birth announcements in the New River Valley community) was used to identify and recruit infants of the appropriate age. Parents and infants were invited to participate first by letter (see Appendix A), then by phone. Infants from monolinguistic homes were scheduled to come into the lab for testing at a time of convenience. On the day of testing, the parents were contacted by phone to confirm the appointment. Twenty healthy, full-term 10- to 11-month-old infants (11 males and 9 females) from English, monolingual homes made up the final sample of this experiment (M age = 44.75 weeks, SD = 1.16 weeks). Prenatal and postnatal health and monolinguicity were confirmed by parental report at the time of testing. The demographics of the final sample were as follows: 90% percent were white/Caucasian, while 10% were of mixed racial background; 100% were from married homes; 90% were delivered vaginally; 90% were breast-fed; 45% were first born. The average age of the infants’ mothers was M = 30.5 years, SD = 3.99 years; the average combined formal education level of the infants’ parents was M = 5.06 years post high school, SD = 2.10 years; the average combined annual household income of infants’ parents was M = 46,000 dollars, SD = 15,000 dollars.

Speech Recordings

The ID speech samples for this experiment were obtained by recording 8 mothers (4 English, 4 non-English) talking in their native languages to their infants. Seven out of the 8 recordings used were from mothers talking to infants in the latter half of the first year of life, based on evidence that adults change the way they address infants throughout development (e.g.,
Ratner, 1996; Snow, 1977; Stern, Speiker, Barnett, & MacKain, 1983). American English served as the ID-native recording. The ID-foreign recordings were from tonal (Cantonese and Mandarin Chinese) and non-tonal (French and Swedish) languages. This was an important detail for several reasons. First, in tonal languages, varying the pitch pattern actually changes the meaning of words. Thus, out of necessity, women whose native language is tonal may exaggerate their speech differently when speaking to their infants. Also, in tonal languages prosodic stress is neither used for emphasis nor for the interrogative form of a sentence, as it is in non-tonal languages (e.g., English). Furthermore, parents from non-Western cultures (who often speak tonal languages) typically strive to keep their infants in a calm state, while parents in Western cultures tend to encourage higher states of arousal (Fernald et al., 1989). Not surprisingly, an exploratory study by Fernald and her colleagues (1989) reported that women whose native languages were tonal used less modulated, lower pitched speech when addressing their infants. Both tonal and non-tonal languages were used in the present study to rule out the possibility that infant attention to speech in one language over another was a function of these types of prosodic nuances.

Endless cassette tapes (20 sec each) were created for infant testing. Utterances from the four English-speaking women were acoustically analyzed using Micro Speech Lab II software.

---

2 The Mandarin Chinese speech samples were recorded from a woman speaking to her 2-month-old infant, because another Asian woman with an older infant could not be located. Nevertheless, this speech sample was included because it was determined that the utterances used were no different from those spoken to older infants. The mean pitch for Mandarin utterances was $M = 234$ Hz, $SD = 21.15$ Hz; the combined mean pitch of utterances in the other four languages was $M = 219.71$ Hz, $SD = 31.1$ Hz. An unpaired t-test (2-tailed) revealed that these differences were not statistically significant, $t(30) = .88, p > .05$, n.s. Likewise, an unpaired t-test revealed that the pitch variability of the Mandarin utterances ($M = 75.5$ Hz, $SD = 18.6$ Hz) was not significantly different from the pitch variability of the utterances in the other four languages ($M = 66$ Hz, $SD = 17.9$ Hz), $t(30) = .958, p > .05$, n.s. The average duration of Mandarin utterances was $M = 1.44$ sec, $SD = .37$ sec, and the combined average duration of the four other languages was $M = 1.52$ sec, $SD = .40$ sec. These too were not significantly different from each other $t(30) = .41, p > .05$, n.s.
and hardware for the IBM computer, and matched as closely as possible with utterances of a non-
English speaker based on mean pitch and duration (See Tables Below).

Table 1.1. Matched Utterances for Experiment 1 Stimulus Tapes: ID-English with ID-Foreign

<table>
<thead>
<tr>
<th>Tape 1</th>
<th>Utterance #</th>
<th>Language</th>
<th>Mean Pitch (Fo)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.</td>
<td>English</td>
<td>220 Hz</td>
<td>1.50 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>French</td>
<td>238 Hz</td>
<td>1.50 sec</td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td>English</td>
<td>259 Hz</td>
<td>.96 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>French</td>
<td>260 Hz</td>
<td>.92 sec</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>English</td>
<td>282 Hz</td>
<td>.90 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>French</td>
<td>286 Hz</td>
<td>1.32 sec</td>
</tr>
<tr>
<td></td>
<td>4.</td>
<td>English</td>
<td>223 Hz</td>
<td>2.02 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>French</td>
<td>247 Hz</td>
<td>2.06 sec</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tape 2</th>
<th>Utterance #</th>
<th>Language</th>
<th>Mean Pitch (Fo)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.</td>
<td>English</td>
<td>160 Hz</td>
<td>1.69 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swedish</td>
<td>182 Hz</td>
<td>1.76 sec</td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td>English</td>
<td>249 Hz</td>
<td>1.66 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swedish</td>
<td>226 Hz</td>
<td>1.64 sec</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>English</td>
<td>218 Hz</td>
<td>1.18 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swedish</td>
<td>217 Hz</td>
<td>1.84 sec</td>
</tr>
</tbody>
</table>
### Tape 3

<table>
<thead>
<tr>
<th>Utterance #</th>
<th>Language</th>
<th>Mean Pitch (Fo)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>239 Hz</td>
<td>1.06 sec</td>
</tr>
<tr>
<td></td>
<td>Mandarin</td>
<td>244 Hz</td>
<td>1.10 sec</td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>210 Hz</td>
<td>1.78 sec</td>
</tr>
<tr>
<td></td>
<td>Mandarin</td>
<td>215 Hz</td>
<td>1.90 sec</td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>217 Hz</td>
<td>1.34 sec</td>
</tr>
<tr>
<td></td>
<td>Mandarin</td>
<td>218 Hz</td>
<td>1.20 sec</td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>250 Hz</td>
<td>1.76 sec</td>
</tr>
<tr>
<td></td>
<td>Mandarin</td>
<td>259 Hz</td>
<td>1.57 sec</td>
</tr>
</tbody>
</table>

### Tape 4

<table>
<thead>
<tr>
<th>Utterance #</th>
<th>Language</th>
<th>Mean Pitch (Fo)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>186 Hz</td>
<td>2.46 sec</td>
</tr>
<tr>
<td></td>
<td>Cantonese</td>
<td>182 Hz</td>
<td>1.96 sec</td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>182 Hz</td>
<td>1.32 sec</td>
</tr>
<tr>
<td></td>
<td>Cantonese</td>
<td>199 Hz</td>
<td>1.34 sec</td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>203 Hz</td>
<td>1.52 sec</td>
</tr>
<tr>
<td></td>
<td>Cantonese</td>
<td>190 Hz</td>
<td>1.48 sec</td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>207 Hz</td>
<td>1.22 sec</td>
</tr>
<tr>
<td></td>
<td>Cantonese</td>
<td>210 Hz</td>
<td>1.28 sec</td>
</tr>
<tr>
<td>5.</td>
<td>English</td>
<td>208 Hz</td>
<td>1.84 sec</td>
</tr>
<tr>
<td></td>
<td>Cantonese</td>
<td>202 Hz</td>
<td>2.02 sec</td>
</tr>
</tbody>
</table>

Sets of approximately four utterances in each language were recorded onto separate channels of a single cassette tape. This resulted in 4 different test tapes with one channel of English ID speech and a second channel of non-English ID speech (2 tonal, 2 non-tonal). Each infant was randomly assigned to one of the four test tapes, with no replacement. Thus, each of the four tapes was used with five infants (for a total of 20 subjects). During testing the visual stimulus was a videotaped, computer generated display of colored concentric circles which progressively increased in contrast toward the center, to direct infant attention toward the middle of the television screen.
Apparatus

Each infant was placed on the lap of the parent, facing the front panel of a three-sided (80 cm x 80 cm x 60 cm) black plywood enclosure. A 12.7 cm by 8 cm television screen (Magnavox model RX 4030 WA 02) was positioned within a window cut into the front panel, offset 7.6 cm to the right, about 35 cm away from the seated infant. One small speaker (Jamo compact 60) was located directly below the television screen in a second window. The lens of a video camcorder (Panasonic model AG-180) was positioned within a circular cut-out in the front wooden panel, slightly to the infant’s left, to provide a full-face view of the infant for the observer on an 11.5 cm by 9 cm television monitor (Sony Solid State model 21 CFR) (See Figure 1).

The observer controlled a hand-held microswitch connected to a Power Macintosh computer and custom-built relay board controlling independent access to two channels of a cassette recorder (Tascam Porta model 05) as well as the output of a VCR (Sanyo VHS model VHR-5214) (see Figure 1). The audio output from the interface was amplified (Harman/Kardon model PM 635) and presented via the speaker at approximately 70 dB (A-scale), as determined by a sound level meter positioned at the head of the infant (Radio Shack catalog number 33-2050). The A-scale was chosen because it assigns less weight to those frequencies to which the human ear is less sensitive (i.e., very high and very low frequencies). Giving less weight to the energy at inaudible frequencies resulted in a more realistic estimate of perceived sound intensity (Werner & Marean, 1996). The ambient noise level at this same position was approximately 60-62 dB (A-scale).
Figure 1.1. Schematic of the Auditory Preference Protocol

Procedure

The sequential auditory preference procedure used by Cooper and Aslin (1994) was employed in the present investigation, with visual fixation as the primary index of infant attention to the auditory recordings. Infants were tested at the Infant Speech Perception Lab of Virginia Polytechnic Institute and State University. Upon arrival, parents were asked to sign a consent form, and to fill out a family demographic sheet, an infant temperament inventory, and a developmental milestone questionnaire (see Appendices B-E). The latter two inventories were added post hoc, for the purpose of examining the data for individual differences in infants’ temperament, level of mobility, and achievements in language perception and production.

Awake and alert infants were placed onto the lap of their parent, and the overhead lights were turned off. Both the parent and the observer were “deaf” to the sounds being presented to the infant on a given trial by wearing headphones over which continuous vocal music was played. The observer started the session by pushing the microswitch button to activate the visual display. Infant visual attention to the concentric circle display, as determined by observed corneal reflection off the infant’s eyes, resulted in the observer pressing the microswitch a second
time. This activated one of the two channels of the tape recorder (randomly determined before testing began). The speech recording continued to play until the observer judged that the infant looked away from the visual display. For Trial 2, the observer again turned on the visual display. When the infant was judged to be looking at the concentric circle display, the speech recording was activated, playing the second channel of the cassette recorder. In this way, the two speech recordings were alternated across trials, with half of the infants receiving one of the two orders of speech presentation. The computer recorded the length of time the infants’ looked to the visual display during each trial. Infant looking times were recorded for a total of 12 alternating speech trials (6 trials of each speech type). The average inter-trial interval was 2.56 seconds.

Results and Discussion

To determine whether the infants preferred to listen to one type of speech over the other, their individual looking times were collapsed across the trials of each speech type, and then divided by the number of their respective trials. A Mixed Analysis of Variance (ANOVA) was chosen for statistical analysis of the data, for the purpose of examining systematic differences between the conditions. Specifically, mean looking times were analyzed with order (ID-native first vs. ID-foreign first) as the between-subjects factor and speech type (ID-native vs. ID-foreign) as the within subjects factor. The results of the ANOVA indicated a significant main effect for speech type, with the infants looking significantly longer to the visual display when it was paired with ID-native speech (M = 8.54 sec, SD = 2.86 sec) than when it was paired with ID-foreign speech (M = 6.82 sec, SD = 1.90 sec); F (1, 18) = 8.46, p < .01 (see Figure Below). No other significant effects were found. The Effect Size (i.e., estimate of the degree to which the null hypothesis was false) of this statistical test was medium to large according to Cohen’s (1992) standards (ES = .47). Moreover, out of the 20 infants in this study, 14 exhibited preferences for ID-native speech over ID-foreign speech (p < .04; binomial 1-tailed test).
In the sequential auditory preference procedure half of the infants are randomly assigned to hear one type of speech on the first trial and the other half are assigned to hear the other type of speech on the first trial. The infant looking times on the first trial were analyzed individually for a few reasons. First, research has demonstrated that novel objects and events tend to engage infant attention (e.g., Columbo & Bundy, 1983). In addition, previous studies in our lab have found that infants have a tendency to look longer at the visual display on the first trial, especially as a function of what they are hearing (e.g., Cooper et al., 1997; Cooper & Aslin, 1994). As a means of finding potential differences between the first looks of infants who heard native ID speech first and the first looks of infants who heard nonnative ID speech first, an unpaired t-test (2-tailed) was carried out. However, no significant difference was found between the looking times on the first trials associated with native speech ($M = 12.79$ sec, $SD = 9.62$ sec) versus those associated with foreign speech ($M = 8.77$ sec, $SD = 6.12$ sec), $t(18) = -1.11$, $p > .05$, n.s.

Even though the analysis on the first trials was not statistically significant, it was the case that the first trials associated with ID-English speech were longer (on average) than those associated with ID-foreign speech. Thus, to best gauge whether the novelty of the first trial was influencing infant preferences, the previous ANOVA was carried out again, this time with the first trials left out of mean calculations. Again, a mixed ANOVA was conducted with order (ID-native first vs. ID-foreign first) as the between-subjects factor, and speech type (ID-native minus
the first trial vs. ID-foreign minus the first trial) as the within subjects factor. The effect remained, with looking times significantly longer when the visual stimulus was paired with ID-native speech minus the first trial (M = 8.12 sec, SD = 2.75 sec) than when it was paired with ID-foreign speech minus the first trial (M = 6.67 sec, SD = 1.60 sec), F (1, 18) = 5.64, p < .05.

Because the infants preferred the native ID speech to nonnative ID speech, it was important to determine whether acoustic differences between the tapes existed, potentially biasing infants’ attention. An unpaired t-test (2-tailed) revealed no difference between the mean pitch of English (M = 219.56 Hz, SD = 30.97 Hz) and non-English utterances (M = 223.44 Hz, SD = 30.16 Hz), t (30) = -.36, p > .05, n.s. Likewise, in an unpaired t-test (2-tailed) I found no difference between the mean pitch of tonal (M = 213.22 Hz, SD = 24.80 Hz) versus non-tonal languages (M = 224.74 Hz, SD = 31.89 Hz), t (30) = -.97, p > .05, n.s. There was no significant difference between the pitch variability of English (M = 74.06 Hz, SD = 18.53 Hz) and non-English utterances (M = 74.56 Hz, SD = 19.10 Hz), t (30) = -.075, p > .05, n.s.; and an unpaired t-test (2-tailed) revealed no significant difference between the pitch variability of tonal-language utterances (M = 62.33 Hz, SD = 13.02 Hz) versus non-tonal languages (M = 79 Hz, SD = 18.44 Hz), t (30) = -2.470, p > .05, n.s. Furthermore, the average duration of each type of utterance was compared. There was no difference between the average duration of English speech (M = 1.51 sec, SD = .42 sec) and the average duration of non-English speech (M = 1.56 sec, SD = .35 sec), t (30) = -.31, p > .05, n.s. Likewise there was no difference between the average duration of tonal (M = 1.54 sec, SD = .35 sec) versus non-tonal languages (M = 1.53 sec, SD = .40 sec), t (30) = .04, p > .05, n.s.

A mixed ANOVA was carried out with tone type (tonal language vs. non-tonal language) as the between-subjects factor and speech type (ID-native vs. ID-foreign) as the within subjects factor, as a means of determining whether infants looked significantly longer to hear one type of language over another. There was no significant difference between infant attention to tonal languages (M = 7.02 sec, SD = 2.32 sec) and that to non-tonal languages (M = 8.35 sec, SD = 2.53 sec), F (1,18) = 2.23, p > .05, n.s., indicating that infant preference was not a function of tone type. In short, infants’ preference for ID speech in their native language was not a function
of one particular speaker’s voice, one particular type of language (tonal vs. nontonal), or the acoustic properties (pitch, pitch variability, duration) of the specific speech utterances.

In this experiment, infant age was used to estimate the perceptual reorganization phase of language development. However, it became clear throughout the course of the study that many of the 10- to 11-month-old infants were not at the same point developmentally. In fact, several of the babies may not have begun to perceptually attune to their native language. For that reason, I examined other variables that could potentially be related to perceptual reorganization (namely, speech perception and production), as indices of performance. In order to compare infants on these dimensions at the individual subject level, their preference scores had to be converted to a single metric. In particular, a ratio of preference magnitude was computed for each infant by dividing their mean looking time to ID-native speech by the sum of their mean looking times to ID-native speech and ID-foreign speech. Thus, infants with a ratio greater than .5 preferred ID-native speech and infants with a ratio of less than .5 preferred ID-foreign speech.

Infants’ preference ratios were not significantly correlated with their speech perception scores (i.e., the number of words that they were able to understand, as determined by maternal report; see Appendix E), \( r = .54 \), \( p > .05 \), n.s., or with their speech production scores (i.e., the number of words that they were able to speak, as determined by maternal report; see Appendix E), \( r = .27 \), \( p > .05 \), n.s. It should be noted that because this was a post hoc consideration, speech perception and production (via the Developmental Milestone Questionnaire; see Appendix E) were only assessed in 40% of the subjects in Experiment 1. Furthermore, the Developmental Milestone Questionnaire may not have been a sensitive measure of infant speech perception and production at this age.

The main goal of this study was to examine the relationship between the linguistic and prosodic influences on infants’ developing speech perception. More specifically, I sought to discern whether 10- to 11-month-old infants would attend enough to the specific linguistic structure of their native language to exhibit a preference for it, over ID speech in a foreign language. The primary hypothesis was that 10- to 11-month-old infants would prefer ID speech
in their native language over ID speech in a foreign language, based on the ‘perceptual tuning’ to native phonetics which occurs late in the first year of life. This hypothesis was supported, in that the 10- to 11-month-old infants preferred ID speech in English to ID speech in a foreign language.

Given that pitch contours have been shown to convey meaning in speech (Papousek, Papousek & Symmes, 1991), and that 8- to 10-month-olds have been shown to exhibit preferences for ID speech in foreign languages (Best et al., 1997; Werker, Pegg & McLeod, 1994), one might predict that infants would find any exaggerated speech salient. However, the present findings do not support this expectation. One strong interpretation of why the infants preferred ID native speech to ID nonnative speech, is that by 10- to 11-months of age, infants have become perceptually attuned to the linguistic information in caretakers’ speech. A more moderate interpretation of the current results however, is that infants at this age can attend to such linguistic information, but only in the context of ID speech. In order to discern whether the context of ID prosody was necessary for eliciting and maintaining infant attention to word-level (linguistic) information, Experiment 2 was designed. In this experiment, the linguistic and prosodic characteristics ID native speech were separated, and the relative ‘perceptual weight’ of linguistic and prosodic information was examined (see Gottlieb, 1985).

Chapter 3
Experiment 2: Preference for ID-Foreign Speech Over AD-Native Speech

In Experiment 2 the preferred prosodic information of the infants (ID speech) was detached from the preferred linguistic information (native speech). In particular, a second group of 10-to 11-month-old infants was presented with adult-directed (AD) speech in English pitted against ID speech in a foreign language (French, Swedish, Mandarin Chinese or Cantonese).
Method

Participants

The same database of local parents and infants was used to identify and recruit infants of the appropriate age. Again, parents and infants were invited to participate first by letter (see Appendix A), then by phone. Infants from monolinguistic homes were scheduled to come into the lab for testing at a time of convenience. On the day of testing, the parents were contacted by phone to confirm the appointment. Twenty healthy, full-term 10- to 11-month-old infants (12 males and 8 females) from English, monolingual homes made up the final sample of this experiment (M age = 45.26 weeks, SD = 1.14 weeks). Prenatal and postnatal health and monolinguicity were confirmed by parental report at the time of testing. The demographics of the final sample were as follows: 90% percent were white/Caucasian, whereas 5% were African American and 5% of mixed racial background; 95% were from married homes; 80% were delivered vaginally; 90% were breast-fed; and 40% were first born. The average age of the infants’ mothers was M = 31 years, SD = 4.66 years; the average combined formal education level of the infants’ parents was M = 4.18 years post high school, SD = 2.66 years; the average combined annual household income of infants’ parents was M = 48,000 dollars, SD = 15,000 dollars.

Speech Recordings

The nonnative ID speech samples used in Experiment 1 were used again in Experiment 2. Native adult-directed (AD) speech samples were obtained by recording women speaking English to another adult. Again, native and nonnative utterances from the four English-speaking women were acoustically analyzed using Micro Speech Lab II software and hardware for the IBM computer. For this experiment, utterances were matched as closely as possible in duration only, as the mean pitch of ID speech and AD speech are different by definition (see Table Below).
### Table 3.1. Matched Utterances for Experiment 2 Stimulus Tapes: AD-English with ID-Foreign

<table>
<thead>
<tr>
<th>Tape 1</th>
<th>Utterance #</th>
<th>Language 1</th>
<th>Duration 1</th>
<th>Language 2</th>
<th>Duration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>4.60 sec</td>
<td>Cantonese</td>
<td>4.60 sec</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>1.36 sec</td>
<td>Cantonese</td>
<td>1.25 sec</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>2.88 sec</td>
<td>Cantonese</td>
<td>2.90 sec</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>3.44 sec</td>
<td>Cantonese</td>
<td>3.61 sec</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>English</td>
<td>2.64 sec</td>
<td>Cantonese</td>
<td>2.90 sec</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>English</td>
<td>2.58 sec</td>
<td>Cantonese</td>
<td>2.32 sec</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tape 2</th>
<th>Utterance #</th>
<th>Language 1</th>
<th>Duration 1</th>
<th>Language 2</th>
<th>Duration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>5.02 sec</td>
<td>Mandarin</td>
<td>6.11 sec</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>3.68 sec</td>
<td>Mandarin</td>
<td>3.42 sec</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>2.92 sec</td>
<td>Mandarin</td>
<td>3.00 sec</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>3.70 sec</td>
<td>Mandarin</td>
<td>3.42 sec</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tape 3</th>
<th>Utterance #</th>
<th>Language 1</th>
<th>Duration 1</th>
<th>Language 2</th>
<th>Duration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>5.23 sec</td>
<td>French</td>
<td>5.36 sec</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>1.88 sec</td>
<td>French</td>
<td>1.56 sec</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>2.02 sec</td>
<td>French</td>
<td>2.00 sec</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>1.66 sec</td>
<td>French</td>
<td>1.56 sec</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>English</td>
<td>2.68 sec</td>
<td>French</td>
<td>2.88 sec</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>English</td>
<td>2.24 sec</td>
<td>French</td>
<td>2.16 sec</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tape 4</th>
<th>Utterance #</th>
<th>Language 1</th>
<th>Duration 1</th>
<th>Language 2</th>
<th>Duration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>English</td>
<td>.88 sec</td>
<td>Swedish</td>
<td>1.00 sec</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>English</td>
<td>2.32 sec</td>
<td>Swedish</td>
<td>1.98 sec</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>English</td>
<td>3.80 sec</td>
<td>Mandarin</td>
<td>3.68 sec</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>English</td>
<td>2.23 sec</td>
<td>Mandarin</td>
<td>1.98 sec</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>English</td>
<td>3.38 sec</td>
<td>Mandarin</td>
<td>3.40 sec</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>English</td>
<td>2.94 sec</td>
<td>Mandarin</td>
<td>1.98 sec</td>
<td></td>
</tr>
</tbody>
</table>
A second set of endless cassette tapes (20 sec each) was created for infant testing. Again, approximately five utterances in each language were recorded onto separate channels of a single cassette tape. This resulted in 4 different test tapes with one channel of English AD speech and a second channel of non-English ID speech (2 tonal, 2 non-tonal). Each infant was randomly assigned to one of the four test tapes, with no replacement. Thus, each of the four tapes was used with five infants (for a total of 20 subjects). The average inter-trial interval was 2.66 seconds. The visual stimulus was identical to that of Experiment 1.

**Apparatus and Procedure**

The apparatus and procedure were identical to that of Experiment 1.

**Results and Discussion**

Looking times that corresponded to each type of speech were collapsed across trials and averaged across the 20 subjects. A mixed analysis of variance (ANOVA) was conducted as a means of determining whether there were systematic differences in infants’ attention to one type of speech over the other. Specifically, mean looking times were analyzed with order (ID-foreign first vs. AD-native first) as the between-subjects factor and speech type (ID-foreign vs. AD-native) as the within subjects factor. The results of the ANOVA indicated a significant main effect for speech type, with the infants looking significantly longer to the visual display when it was paired with ID-foreign speech ($M = 7.95$ sec, $SD = 3.15$ sec) than when it was paired with AD-native speech ($M = 6.47$ sec, $SD = 2.18$ sec): $F(1, 18) = 7.6, p < .01$ (see Figure 3). No other significant effects were found. The Effect Size (i.e., estimate of the degree to which the null hypothesis was false) of this statistical test was medium to large according to Cohen’s (1992) standards ($ES = .42$). Furthermore, out of the 20 infants in this study, 16 exhibited preferences for ID-foreign speech over AD-native speech ($p < .01$; binomial 2-tailed test).
Just as in Experiment 1, additional analyses were conducted on the first trials to gauge whether there were differences between the first looks of infants who heard nonnative ID speech first and those of infants who heard native AD speech first. In fact, there was no statistically significant difference between the first looks to nonnative speech ($M = 10.41$ sec, $SD = 6.6$ sec) and the first looks to native speech ($M = 10.35$ sec, $SD = 7.29$ sec), $t(18) = -.02$, $p > .05$, n.s.

Furthermore, a second mixed analysis of variance (ANOVA) was performed, with the first trials left out of the mean calculations, to determine whether responding on the first trial was influencing infant preferences. Order (ID-foreign first vs. AD-native first) was the between-subjects factor and speech type (ID-foreign minus the first trial vs. AD-native minus the first trial) was the within subjects factor. Even without the first trials in mean calculations, the infant looking times were significantly greater to the visual stimulus when it was matched with ID-foreign speech ($M = 7.78$ sec, $SD = 3.24$ sec) than when it was matched with AD-native speech ($M = 6.05$ sec, $SD = 1.99$ sec), $F(1, 18) = 8.68$, $p < .01$.

Because the properties of the speech used in Experiment 2 were different by definition (i.e., AD speech and ID speech), acoustic differences in the tapes were expected. Still, it was
important to determine statistically that the speech presented was in fact characteristic of infant-directed and adult-directed speech. For that reason, an unpaired t-test (2-tailed) was carried out between the mean pitch of English AD speech (M = 174.83 Hz, SD = 18.27 Hz) and the mean pitch of foreign ID speech (M = 226.03 Hz, SD = 31.36 Hz). The pitch of the ID speech was significantly greater than that of the AD speech, t(51) = -6.97, p < .0001. Likewise, an unpaired t-test (2-tailed) between the pitch variability of English AD speech (M = 45.52 Hz, SD = 11.55 Hz) and the pitch variability of foreign ID speech (M = 83.30 Hz, SD = 17.16 Hz) was conducted. The pitch variability of ID speech was significantly greater than that of AD speech, t(51) = -9.09, p < .0001. Thus, the characteristics of the ID and AD speech samples used in Experiment 2 were in keeping with defined differences between ID and AD speech in the literature (e.g., Cooper, Abraham, Berman & Staska, 1997; Fernald, 1985; Fernald & Kuhl, 1987; Papousek, Papousek & Bornstein, 1985; Werker & McLeod, 1989).

The average duration of the utterances used in Experiment 2 were also compared. An unpaired t-test (2-tailed) revealed that the mean duration of the English tapes (M = 2.91 sec, SD = 1.10 sec) was not significantly different from the mean duration of the non-English tapes (M = 2.87 sec, SD = 1.30 sec), t(51) = .112, p > .05, n.s. Therefore, the possibility that preferential attention to ID nonnative speech was a function of the length of the utterances was ruled out.

As in Experiment 1, ratios of preference magnitude were calculated for each infant, as a way of comparing their performance with developmental determinants other than age. In particular, each infant’s mean looking time for ID-foreign speech was divided by the sum of their mean looking times for ID-foreign speech and AD-native speech. Infants with a ratio above .5 preferred ID-foreign speech, and those with a ratio below .5 preferred AD-native speech. Post hoc tests were then carried out, to compare these preference ratios with other potential measures of perceptual reorganization (i.e., speech perception and production). Unlike Experiment 1, speech perception and production (via the Developmental Milestone Questionnaire; see Appendix E) were assessed in 100% of the infants in this experiment. Parentally reported speech production (i.e., the number of words that they were able to speak, as determined by maternal report; see Appendix E) was negatively correlated with infants’ preference ratios, (r = -.53), p <
.05, meaning that the more words an infant was producing, the more likely they were to preferentially attend to AD-native speech over ID-foreign speech. On the other hand, speech perception (i.e., the number of words that they were able to understand, as determined by maternal report; see Appendix E) was not significantly correlated with preference ratios, ($\tau = -.34$), $p > .05$. However, recall that the Developmental Milestone Questionnaire may not have been a sensitive enough measure of infant speech perception and production at this age.

To summarize, in Experiment 2 the prosodic information that infants typically encounter (ID speech) was separated from the linguistic information that they encounter (native speech), in an effort to understand the relative prominence of these types of information. The 10- to 11-month-old infants exhibited a preference for ID speech in a foreign language over AD speech in English, indicating that the prosodic information had more perceptual draw than the linguistic information at this age.

Chapter 4
General Discussion

This pair of experiments was primarily designed to address the function of ID speech within the context of perceptual tuning to native linguistic structure. In Experiment 1, I found that 10- to 11-month-old infants preferred ID speech presented in their native language (English) to ID speech presented in a foreign language. There are three reasons why this may have been so. First, perhaps by 10- to 11-months of age infants have narrowed their focus to only those sound categories that exist in their native language, and can no longer discriminate sounds that exist outside of that language (i.e., have fully perceptually attuned to it). If this were the case, the infants’ would be attending to speech at the level of linguistics, and would prefer speech in their native language due to their recognition and/or understanding of it. A more moderate explanation is that 10- to 11-month-olds can attend to linguistic-level information in their native language, but only when it is couched in the exaggerated prosody of ID speech. Finally, it is possible that the 10- to 11-month olds preferred ID native speech simply because it contained the characteristic prosodic melody of the language to which they have been exposed throughout the first year of their life. In sum, the critical issue to address in explaining Experiment 1 was
whether the 10- to 11-month-olds attended preferentially to ID native speech based on some level of linguistic processing, or merely based on their sensitivity to the prosodic patterning typical of their native language.

The possibility that the preference for ID native speech in Experiment 1 was due to prosody alone gains some support from evidence that infants perceive and respond to the tone and rhythmicity of their native language from the first days of life. Moon, Cooper and Fifer (1993) demonstrated this with 2-day-olds. Infants of English- and Spanish-speaking parents were presented with both languages in an infant-controlled sucking procedure. The infants sucked longer during a native language recording than during a foreign language recording, suggesting that they preferred to listen to familiar (native) speech. In a similar investigation by Bahrick and Pickens (1988), 5-month-old infants of English-speaking parents that were habituated to an English sentence dishabituated more often to a Spanish sentence, than to a novel English sentence. In another study, Mehler, Dupoux, Nazzi and Dehaene-Lambertz (1996) discovered that although French 4-day olds could distinguish changes among syllables (the elementary unit of French), they could not discriminate between morae (the elemental unit of Japanese). Finally, a study by Mehler and his colleagues reported that French 4-day-olds could discriminate French from Russian speech, even when the language targets were low-pass filtered, indicating that the infants could recognize their native language based on prosody alone (Mehler, Jusczyk, Lambertz, Halstead, Bertoncini & Amiel-Tison, 1988). Clearly infants this young were not basing their preferences on word-level comprehension.

On the other hand, support for a word-level explanation of the observed preference in Experiment 1 comes from empirical evidence that infants can process linguistic-level information by the end of the first postnatal year. In fact, it seems that infants become attuned to various aspects of the phonetic organization of their native language at about the same time that they lose their ability to discriminate nonnative phonemes (Jusczyk, Frederici, Wessels, Svenkerud & Jusczyk, 1993). For instance, data suggests that by 9-months of age infants can identify both the sounds that exist in their native language (phonetic information) and the ordering of those sounds that is permitted in their native language (phonotactic information). In
particular, Jusczyk and his colleagues (1993) conducted a set of experiments to test whether infants treat words following native phonetic and phonotactic structure differently from nonnative words. To minimize the possibility that infants could respond to prosodic cues, the researchers chose languages with very similar prosodic organization (i.e., English and Dutch), and used words rather than sentences as stimuli. In addition, the specific English and Dutch words presented contained phonetic items that were unavailable in the respective languages. For example, in English, the consonant \(d\) can occur at the end of a syllable, whereas in Dutch it cannot; in Dutch, consonant sequences such as \(zw\) and \(kn\) can begin a syllable, whereas in English they cannot (Jusczyk et al., 1993, p. 405). Using a head-turn preference procedure, the authors discovered that both American and Dutch 9-month-olds listened longer to words in accordance with the phonetic and phonotactic standards of their respective native languages. When the word lists were low-pass filtered (removing all linguistic information), neither American nor Dutch 9-month-olds exhibited a preference for the words in their native language. Thus, the original preference for native words was not based on the characteristic prosody of that language.

Furthermore, both English and Dutch 6-month-olds failed to exhibit a preference for native over nonnative words (either normal or filtered), implying that the sensitivity to native linguistics emerges at the same time as the decline in sensitivity to nonnative phonemes (Jusczyk et al., 1993). Interestingly, when American infants were presented with a contrast language having more pronounced differences in stress patterning (i.e., Norwegian), both 6-month-olds and 9-month-olds preferred English. Apparently the distinct prosody highlighted linguistic information, making the native language structure apparent to even 6-month-olds.

In another study, Jusczyk, Luce and Charles-Luce (1994) examined whether 9-month-old infants attend to the typical phonetic sequence distribution in speech. The researchers reasoned that the development of word recognition might initially rely on recognizing phonetic sequences that are most probable. Specifically, using a head-turn preference procedure they presented infants with a choice between nonsense words following a sound pattern typical in English versus nonsense words with nontypical sound patterns. The 9-month-olds listened longer to the
To gain a more clear understanding of the role that linguistic information played in the infants’ preference for ID native speech in Experiment 1, the next empirical step I took was to separate the preferred linguistic and prosodic features from one another, and discern whether one had more ‘perceptual draw’. Specifically, in Experiment 2 the linguistic and prosodic features of ID-native speech were pitted against one another. When offered the choice between listening to speech that contained native language words versus speech that was melodic and exaggerated, the 10- to 11-month-olds preferred the prosodically exaggerated speech. In other words, although the infants had supposedly already begun to ‘tune in’ to the linguistic specifics of their native language, the prosodic information in speech held more perceptual weight than the linguistic information.

It makes sense that infants would be sensitive to the global prosodic characteristics of their native language prior to showing sensitivity to the finer details, as categorizing a limited number of recurring intonational contours is easier than matching a multitude of words with the situations in which they arise (Werker, Lloyd, Pegg & Polka, 1996). Furthermore, perceptual development is characterized by increasingly finer-grained analyses of the available perceptual array (Gibson, 1969). Since prosody is perceived from earlier on, it is likely more pronounced than linguistic information, and could remain a reliable, powerful signal throughout language development. The exaggerated prosody in ID speech may even direct infant attention to relevant sounds that exist in the native language, and by doing so, it may play a role in the perceptual reorganization process. Karzon’s (1985) study suggests this is so. In another set of experiments, Mandel, Jusczyk and Kelmer Nelson (1994) discovered that prosodic structure in speech actually increased infants’ memory for specific phonetic information. In particular, using a high amplitude sucking habituation procedure, these researchers found that 2-month-old infants better remembered phonetic properties of words when they were presented in a prosodic unit (i.e., a
clause) than when they were presented in a fragmented list. Furthermore, their second study revealed that 2-month-olds better remembered contiguous words that occurred in the same prosodic unit (i.e., clause) than contiguous words that straddled prosodic units. These findings were the first to show that infants are not only sensitive to prosodic cues in speech, they actually exploit them in the processing of speech (Mandel et al., 1994).

Future research efforts should try to more effectively get at the question of whether the infants in Experiment 1 based their preference for ID native speech on information at the level of linguistics. One way to ascertain whether linguistic information was involved in drawing infants’ attention to ID native speech over ID foreign speech, would be to present 10- to 11-month-olds with the same contrast (ID native vs. ID foreign), but in low-pass filtered speech. If infants still prefer ID native speech to ID foreign speech, reliance on the characteristic prosody of their native language cannot be ruled out as a source of influence.

To get at this same question from another angle, the contrast from Experiment 1 (ID native vs. ID foreign) could be presented, but with ID speech in the various languages matched for tone structure. That is, the women recorded speaking to their infants would be instructed to read a script using a specific inflection pattern. In this case a preference for ID speech in English over ID speech in a foreign language would indicate that infants attend to linguistic information to make their choice. Perhaps in this latter instance the preference for ID speech in English would not emerge until later, when such fine distinctions between words could be made.

As another follow up study, it would be informative to determine if and when infants’ attention is focused on linguistic information at the expense of prosody (i.e., the age at which infants begin to prefer AD speech in their native language, even when it is contrasted with ID speech). The perceptual draw of ID prosody most likely has to diminish enough at some point to allow infants to attend more to linguistic information in the speech around them. Maybe once infants have begun to produce their native language, they would begin to prefer it to speech in another language. In fact, the negative correlation that was found in Experiment 2 between the number of words infants produce and their tendency to prefer ID speech in a foreign language,
supports this contention. In a future study, the contrast from Experiment 2 could be presented to 18-month-old infants who are just beginning the stage of language production known as the ‘vocabulary spurt’ (Bloom, 1989).

From a theoretical perspective, if ID speech supports the reorganization of phonemic perception, then a lack of exaggerated prosodic input may have pronounced consequences. For example, depressed mothers have an increased tendency to use flat, dull (non-ID) speech when talking to their infants. One study showed that depressed mothers were six times more likely than non-depressed mothers to speak to their infants in an unexaggerated, prosodically constrained manner (Bettes, 1988). Perhaps the infants of depressed mothers are unable to utilize prosodic cues as a highlighter of linguistic details, and as a result, are delayed in their perceptual reorganization. It follows that these infants would also be delayed in the perception and production of their native language. In order to gain a more clear understanding of the potential role of prosody in fostering perceptual reorganization, one could study the infants of this population at various ages, and then compare their preferences for native versus nonnative speech with those of infants from non-depressed caretakers.

Overall, this pair of experiments shows that ID speech continues to be a powerful event for infants, even at the end of the first postnatal year. Since the language input that infants encounter remains specialized both linguistically and prosodically, it likely impacts the development of speech perception. After all, the perceptual reorganization (tuning to native language sounds) that occurs at approximately 10- to 12-months of age (Werker & Tees, 1984) occurs within the context of this specialized speech register. Perhaps infants are only able to ‘tune in’ to the specific linguistic nuances of their native language because that information is couched in the context of ID speech.
References


Appendix A

Infant Speech Study Program
Department of Psychology
Virginia Tech

Dear Parent(s):

Soon after infants are born, they can recognize many different sounds and voices. For instance, we now know that babies only a few days old would rather listen to their own mother’s voice than to the voice of another woman. All during the first year of life, infants hear lots of different sounds in their environment. In the Department of Psychology at Virginia Tech, we are interested in the amazing feat of language learning. In particular, we would like to know how babies use the cues in their environment to learn to talk.

You and your baby are invited to participate in our latest project! Currently, we are investigating what particular sounds infants find most interesting. To do this, we play recorded speech sounds for the infants in various languages, and observe what they like to listen to most. Your participation would involve just one visit to the Infant Speech Study Program (located next to Bogen’s restaurant in Blacksburg; a map is enclosed for your convenience) when your baby is between 10 and 11 months old, so that we can observe your infant’s responsiveness to different voices. The test only lasts about 15 minutes, but we like to schedule a full hour appointment with you to give you and your baby time to get settled without feeling rushed. We schedule this appointment at a time that is most convenient for your (and your baby’s) schedule. If you have older children, you are welcome to bring them along. We have a waiting room right next to the observation room with lots of toys, and we offer free babysitting for your convenience.

If you would like to schedule an appointment for your infant or find out more about our program, please feel free to call us at either 231-3972 or 231-5938. We hope to see you and your baby soon!

Sincerely,

Wendy L. Ostroff
Graduate Researcher

Robin Panneton Cooper, Ph.D.
Associate Professor
Appendix B

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
INFORMED CONSENT FORM

TITLE: Infant-Directed Speech within the Context of Declining Nonnative Phoneme Perception

PRINCIPAL INVESTIGATORS: Robin Panneton Cooper, Ph.D. & Wendy L. Ostroff

I. Purpose of Research Project

The purpose of this project is to investigate 10- to 11-month-old infants’ perception of speech sounds in various languages.

II. Procedure

Your infant will be tested for approximately 15 minutes, provided that he/she is awake, alert and quiet. You will hold your baby in your lap, facing a video screen onto which a display of colored concentric circles will appear. When your infant looks at this screen, a speaker located below the screen will present a recording of a woman speaking. Your infant can control how long he/she gets to hear this voice, by how long they look at the circle display. On some trials the voice will be in English, and on other trials the voice will be speaking another language. If your infant prefers one type of speech over another, we expect that he/she will pay more attention to the visual display when that voice is playing.

The sound levels of the voices played to your infant are no louder than the sounds present in the typical home environment (i.e., 70 dB). If your infant cries or falls asleep, testing will stop. Also, each infant will be videotaped during the session (these tapes will be stored in our lab, and erased after 5 years). There are no apparent risks to you or your baby resulting form participating in this study.

III. Benefits of this Project

Your baby’s participation in this project benefits the field of infant speech perception. Specifically, this study will further our knowledge of the cues in the environment which help infants learn language.

IV. Extent of Anonymity and Confidentiality

All of the information gathered in this study will be kept confidential and the results will not be released without your consent. The information your baby provides will be identified by a number only (no names). Your informed consent will be kept separate from your infant’s information. The results of this study may be presented at scientific meetings, and/or published in a scientific journal. If you would like, you will be sent a summary of this work when the project is completed.
V. Freedom to Withdraw

You have the right to terminate your involvement in this project at any time and for any reason, if you so choose.

VI. Approval of Research

This project has been approved by the Human Subjects Committee of the Department of Psychology and the Institutional Review Board of Virginia Tech.

VII. Subject’s Permission

I have read and understand the informed consent and conditions of this project. I have been given an opportunity to ask further questions about the procedure, and I understand that I have the right to end this session for any reason if I so choose. If I have any questions regarding this research and its conduct, I should contact one of the persons named below. Given these procedures and conditions, I give my permission to Dr. Cooper and her graduate students to test my infant.

________________________________________                                   ________________
Signature of Parent         Date

Dr. Robin Panneton Cooper, Principal Investigator  231-5938
Wendy L. Ostroff, Graduate Investigator         961-1457
Dr. R.J. Harvey, Chair, Human Subjects Committee  231-7030
Dr. Thomas Hurd, Chair, Institutional Review Board  231-9359
Appendix C

Family Information Sheet

(All information is strictly confidential)

Mother’s Age: ________

Mother’s Occupation: ____________________________

Father’s Occupation: ____________________________

Mother’s Education (in years): ____________________

Father’s Education (in years): ____________________

Estimated Family Income: _________________________

Race:  White/Caucasian  African American  Hispanic  Asian  Native American  Other

Marital Status:  Married  Separated  Divorced  Single

For your most recent pregnancy, please note the following:

Method of Delivery:  Vaginal  Caesarean

Method of Feeding:  Breast  Bottle

Estimated Gestational Age at Birth (in weeks): __________

Please list the gender and age of your older children (if any):

1. ________________

2. ________________

3. ________________

4. ________________
Appendix D

**Infant Temperament Questionnaire**

*For the following questions, please circle the number that is most typical of your baby.*

1. **How easy is it for you to calm or soothe your baby when he/she is upset?**
   
   1. very easy
   2. about average
   3. difficult

2. **How easy is it for you to predict when your baby will go to sleep?**
   
   1. very easy
   2. about average
   3. difficult

3. **How easy is it for you to know what’s bothering your baby when he/she cries or fusses?**
   
   1. very easy
   2. about average
   3. difficult

4. **How many times per day, on the average, does your baby get fussy and irritable—for either short or long periods of time?**
   
   1. never
   2. 1-2
   3. 3-4
   4. 5-6
   5. 7-8
   6. 10-14
   7. over 15

5. **How much does your baby cry and fuss in general?**
   
   1. very little
   2. about average
   3. quite often

6. **How easily does your infant get upset?**
   
   1. not easily
   2. about average
   3. very easily

7. **When your baby gets upset, how vigorously or loudly does he/she cry?**
   
   1. very mildly
   2. about average
   3. very loudly

8. **How much does your baby want to be held?**
   
   1. very often
   2. sometimes
   3. not very often
Appendix E

Developmental Milestone Questionnaire

Has your baby begun walking?

Yes  No

If so, how long ago (in weeks) did he or she take the first steps?

______________

Has your baby begun speaking any words?

Yes  No

If so, approximately how many?

______________

Approximately how many words does your baby seem to understand?

______________

Is your infant exposed to any language other than English on a regular basis (i.e., daily)?

Yes  No
VITA
Wendy L. Ostroff

Personal Information
Birth date: August 9, 1973

Current Address: 721 Hutcheson Drive
Blacksburg, VA 24060

E-mail: wostroff@vt.edu
Phone: (540) 961-1457

Educational History
Stoughton High School, Stoughton, MA
National Honor Society Highest Academic Honors, June 1991

University of Connecticut, Storrs, CT
Bachelors degree in Psychology, Magna cum Laude, May 1996

Virginia Polytechnic Institute and State University, Blacksburg, VA
APA Accredited Doctoral program in Psychological Science,
Developmental Concentration (currently enrolled)
**Recent Research Activity**

- **Undergraduate Research - University of Connecticut Infancy Lab**
  
  Supervisor: Gwen Gustafson, Ph.D.
  
  **Senior Thesis:**
  
  *Adult Perception of Infant Communication*
  
  - Developed and conducted an original, independent research project.
  
  - Presented as a poster at 10th Biennial International Conference on Infant Studies, Providence, RI: April 1996.

- **Graduate Research - Infant Speech Perception Program of Virginia Tech**
  
  Graduate Mentor: Robin Panneton Cooper, Ph.D.
  
  *The Importance of Rate for Infant-Directed Communication*
  
  - Conducted a replication study, with one-month-old infants.
  

**Thesis: Master of Science Degree**

*The Perceptual Draw of Prosody: Infant-Directed Speech within the Context of Declining Nonnative Phoneme Perception*

- Defended October 7, 1998


**Recent Conference Activity**

- **10th Biennial International Conference of Infant Studies** - Providence, RI: April, 1996.
  
  - Attended
  
  - Presented poster:

  The role of the face in perception of infant sounds: A first look.
  - Attended

  - Attended
  - Presented poster:
    

  - Attended
  - Presented poster:
    
  - Received third place in the division of social science research

  - Presented poster:
    

  - Attended

• **Biennial Meeting: Society For Research in Child Development** – Albuquerque, NM: April, 1999.
  - Planning to attend
  - Submitted poster:
    
Awards

Graduate Research Development Project Grant

-Awarded $300.00 for master’s thesis research by the Graduate Student Assembly of Virginia Polytechnic Institute and State University - January, 1998.

Travel Fund Program Award

-Awarded $275.00 for conference travel expenses by the Travel Fund Committee of the Graduate Student Assembly of Virginia Polytechnic Institute and State University - January, 1998.

14th Annual Virginia Tech Research Symposium Award

-Awarded $75.00 for placing third in the division of social science research - April, 1998.

29th Carnegie Symposium on Cognition Travel Fellowship


Teaching Experience

Introductory Psychology Teaching Assistantship

-Taught five recitation classes of approximately 30 undergraduate students each

-Fifteen hour per week graduate assistantship with tuition waiver

Academic Advisory Experience

Department of Psychology Undergraduate Information Office

-Currently serving as an academic advisor to approximately 800 undergraduate psychology majors, organizing pre-registration, spring commencement ceremony, Freshman orientation and Transfer Student registration

-Twenty hour per week graduate assistantship with full tuition waiver
**Academic Activities**

Developmental Science Society

- Graduate student organization and discussion forum
- Treasurer and cofounder

Psychological Sciences Area - Graduate Student Liaison

- Liaison between graduate students in the Psychological Sciences area and the Area Director, Dr. Helen J. Crawford.

**Professional Organization Membership**

American Psychological Association: Graduate Student Affiliate

International Society for Infant Studies

American Psychological Society

**Graduate Level Course Work**

<table>
<thead>
<tr>
<th>Biological Bases of Behavior</th>
<th>Principles of Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Development</td>
<td>Psychophysiology</td>
</tr>
<tr>
<td>Cognitive Psychology</td>
<td>Research Methods</td>
</tr>
<tr>
<td>Developmental Psychobiology</td>
<td>Social Development</td>
</tr>
<tr>
<td>Developmental Psychology</td>
<td>Statistics for Social Science Research I &amp; II</td>
</tr>
<tr>
<td>Personality Processes</td>
<td></td>
</tr>
</tbody>
</table>
**Other Experience**

- **Volunteer**

  *Norwich Psychiatric Hospital* - fall 1994/Spring 1995: Norwich, CT
  
  - Facilitated and supervised weekly patient activities

- **Study Abroad**

  *Charles University* - fall 1995: Prague, Czech Republic

  - Studied Czech language, music, film, theater and literature

  - Experienced Eastern European lifestyle and culture