PLASTICITY IN INFANTS' SPEECH PERCEPTION: A ROLE FOR ATTENTION?

by

Katherine Aya Yoshida

B. Sc., McGill University, 2002

M.A., the University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2008

ABSTRACT

Phonetic perception becomes native-like by 10 months of age. A potential mechanism of change, distributional learning, affects the perception of 6-8-month-old infants (Maye et al., 2002). However, it was anticipated that perception may be more difficult to change by 10 months of age, after native categories have developed. In fact, some evidence suggests that by this age, the presence of social interaction may be an important element in infants' phonetic change (Kuhl et al., 2003). The current work advances the hypothesis that infants' level of attention, which tends to be higher with social interaction, may be a salient factor facilitating phonetic change. Three experiments were designed to test infants' phonetic plasticity at 10 months, after phonetic categories have formed. A non-social distributional learning paradigm was chosen, and infants' attention was monitored to probe whether a facilitating role would be revealed.

In Experiment 1, 10-month-old English-learning infants heard tokens from along a continuum that is no longer discriminated at this age that formed a distribution suggestive of a category boundary (useful distinction). The results failed to reveal evidence of discrimination, suggesting that the distributional information did not have any effect. A second experiment used slightly different sound tokens, ones that are farther from the typical English pronunciation and are heard less frequently in the language environment. Infants still failed to discriminate the sounds following the learning period. However, a median split revealed that the high attending infants evinced learning. Experiment 3 increased the length of the learning phase to allow all infants to become sufficiently high attending, and revealed phonetic change. Thus, after phonetic categories have formed, attention appears to be important in learning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

My greatest gratitude is owed to my advisor, Janet Werker. She has provided immeasurable support and guidance over the past years, and I would have never completed the programme without her encouragement and above-and-beyond levels of understanding, academic and otherwise.

I am also grateful for the opportunity to interact and collaborate with other faculty, primarily Geoff Hall, Steven Heine, Ara Norenzayan, Alan Kingstone, Todd Handy, and Ralph Hakstian. I thank Geoff, Ara, Todd and Anita DeLongis for their input on this thesis.

I am indebted to a lab full of wonderful individuals, several of whom deserve particular mention. I was lucky to work closely with Chris Fennell on my first project, and with Ferran Pons on the first experiments of what would grow into this dissertation. And nothing would ever be accomplished without the efforts of Laurel Fais and Ramesh Thiruvengadaswamy. Laurie, Henny Yeung, Judit Gervain, and Athena Vouloumanos have provided helpful comments on previous versions of this work. Tania Zamuner contributed greatly to the infant participant pool.

I thank a series of amazing lab coordinators: Erin Moon, Marisa Cruickshank, Marie Jette, Vashti Garcia, Jasmine Cady, Emily Chevrier, and especially Vivian Pan; and the equally amazing Early Development Research Group recruitment coordinators, Jess Deglau and Sarah Heller.

There were a number of research assistants who contributed directly to this work. Thanks to Stephanie Helm, Jasmine Cady, and Nazanin Akmal, who tested infant participants, and to Clarisa Markel, Moko Chen, Sarah Jo, and Kat Gunion, who coded videos and kept tabs on CDIs.

Two of my favourite people in the world: Eli Puterman and Mijke Rhemtulla.

I thank Athena, who gave me a second lab-home during the writing of this dissertation, and Antonia Joannides and T.J. Siegal, for keeping me happily caffeinated throughout.

Finally, I thank my family, for everything that they have given me (which is everything): Vince Yoshida, Wendy Yoshida, Greg Yoshida, Tina Tran, my aunties Sherry, Kaz, Christine, and Margaret, my late Obachan, and Harry and Sue Miyamoto.

INTRODUCTION

One of our defining human attributes is the ability to communicate. Some species also have specialized communicative signals such as monkey calls that carry specific meaning (*Predator above!*) or the dance of the honeybee that conveys precise spatial direction. However, no other animal code comes even close to the sophisticated ability of human language to communicate abstract thoughts and ideas. This productivity is accomplished through infinite combinations of a finite number of units (Hockett, 1960). Though obviously cognitive in process, language is social in context, always learned through communication and interaction with other individuals. The presence of social interaction implies the presence of many different factors, including attention, which is drawn by social interaction. A facilitative role for social interaction has been previously suggested in language learning. This thesis will argue that attention is, independent of social interaction, an important factor in learning language.

The structure of language is complex, composed of many different levels of information including semantic (meaning), syntactic (grammar), morphological (word), and phonetic (sound). This thesis focuses on infants' learning of the most basic of these levels, examining the development of phonetic perception (speech perception). In this introduction, I first discuss phonetic perception and its development in infancy. I next review some of the statistical language learning literature, including distributional learning, a proposed mechanism of phonetic development. The social context within which language is learned is considered, highlighting a study in which infants in a social condition show phonetic learning and infants in non-social conditions do not. Attention is suggested as a potential alternative explanation for results of this social manipulation, especially given the fact that infants in that study had developed native phonetic categories, possibly making phonetic learning more difficult. Three experiments address these issues and ultimately reveal that infants' perception is indeed less plastic after phonetic category formation, but that perceptual change is effected with sufficient attention.

Terminology

Speech sounds are known as phones, which cluster into language-specific categories called phonemes. Phonemes contrast with one another to distinguish meaning: if a phoneme is added, deleted, or changed in a word, the meaning of the word will be altered (e.g., *in*, *bin*, *pin*, *pinch*). The existence of minimally-contrasting word pairs reveals the phonemic inventory of a language. For example, the word pair *bin-pin* reveals that the *b* and *p* sounds (transcribed [b] and [p^h][1]) belong to separate phonemes in English.

Adult perception can be discussed in terms of phonemes; however, young infants, the focus of this thesis, do not yet know a criterial number of minimal pair words to be considered to have awareness of the lexically contrastive phonemic aspect. Therefore, even insofar as infants' perception may be adult-like, it will be discussed in terms of phonetic categories (Werker & Curtin, 2005). Phonetic categories are distinguished on the basis of the physical (sound) properties of phonemes, without implying lexical knowledge. These phonetic categories change and develop in accordance with the native

---

[1] Square brackets will be used to denote phones and slashes will be used to denote phonemes.

language phonemic inventory (discussed in detail below), to later coalesce into phonemes.

Phonetic Perception

Experience changes the salience of perceptual dimensions (Nosofsky, 1986). Language experience pulls non-contrasting phones closer together, resulting in more difficult discrimination, and pushes contrastive phones farther apart, resulting in easier discrimination (Jusczyk, 1993; Kuhl, 1993). In this way, phonemes ease phonetic perception by enhancing distinctions that are meaningful and diminishing those that are not meaningful in language. As a result, adults perceive and identify phones in terms of their phonemic category ("this is a *b*, not a *p*"), rather than along a gradient ("somewhere between a *b* and a *p*"). In labeling sounds along a continuum, identification switches from one category to the next in a sudden rather than a gradual manner, revealing sharp boundaries between categories (e.g., Liberman, Harris, Hoffman, & Griffith, 1957).

Phonemic categories are language-specific, differing across different language environments. For example, in Japanese, [r] and [l] do not belong to separate phonemes, and both sounds are perceived by Japanese speakers as instances of the single Japanese category (Strange & Jenkins, 1978). Similarly, Hindi has two separate *d* categories, with one articulated with the tongue tip contacting the top row of teeth (the dental /d/) and the other articulated with the tongue curling up and around towards the back of the mouth, hitting the soft palate (the retroflex /D/). These Hindi phonemes are collapsed into a single *d* in English (an alveolar /d/, which is articulated with the tongue hitting the front of the palate, medial to the Hindi articulations). When exposed to a [ba] – [da] sound continuum, adult speakers of both Hindi and English clearly identify separate /ba/ and

/da/ categories, but Hindi speakers additionally separate the latter sound space into separate /da/ and /Da/ categories (Werker & Lalonde, 1988). How do individuals learn which phones cluster together into a single phonetic category and which phones contrast with each other, denoting differences of meaning in their native language?

Phonetic Development

Very young infants already have some phonetic organization. Infants of 1 and 4 months of age perceive sounds categorically, discriminating sound changes that cross a phonetic boundary, but not within-category changes of equal magnitude (Eimas, Siqueland, Jusczyk, &Vigorito, 1971; see also Dehaene-Lambertz & Baillet, 1998; Dehaene-Lambertz & Peña, 2001). This early categorical behaviour is language-general as infants also discriminate non-native even non-speech categorical changes (Aslin, Pisoni, Hennessy, & Perey, 1981; Lasky, Syrdal-Lasky, & Klein, 1975; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Streeter, 1976; Werker, Gilbert, Humphrey, & Tees, 1981).

It was generally expected that the transition from language-general (early infancy) to language-specific perception (adulthood) would occur sometime during childhood or puberty (Werker & Tees, 1983). An examination of English-learning infants confirmed that infants of 6-8 and 8-10 months were still able to discriminate two non-native contrasts. However, infants of 10-12 months did not discriminate the sounds, revealing that infants' perceptual abilities develop in the first year of life (Werker & Tees, 1984a). These effects have been replicated and extended using artificial speech stimuli (Werker & Lalonde, 1988) and a number of other consonantal contrasts, including an affricate-fricative Mandarin Chinese contrast (Kuhl, Tsao, & Liu, 2003), a tone contrast (Mattock & Burnham, 2006), and /ra/ - /la/ in a Japanese population (Kuhl et al., 2006; Tsushima et al., 1996), all implicating the second half of the first year of life as the time of perceptual change.

Perceptual reorganization has been documented at slightly earlier ages with vocalic data. Infants of 6 months learning either English or Swedish were tested with vowels that were prototypical or not prototypical of each language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). The perceptual space of the English-learning infants appeared warped so that the English vowel tokens were perceived as more alike, just as the perceptual space of the Swedish-learning infants appeared warped in accordance with the Swedish vowels. Neither group showed any such reorganization to the vowels of the other language (see also Polka & Werker, 1994 for evidence of development by 6-8 months and Cheour et al., 1998 for event-related potential evidence of slightly later development, between 6 months and 1 year). Together, these studies make clear that the first-year reorganization is specific to the native language that the infant is learning.

Perceptual reorganization is not only brought about by decline of non-native discrimination but also by enhanced discrimination of native contrasts. The first evidence of this was facilitation in discrimination of an English contrast at some point between 10-12 months of age and adulthood (Polka, Colantonio, & Sundara, 2001). Further cases of facilitation have been documented earlier, in the same time frame as decline of non-native contrasts. English-learning infants evince enhanced discrimination of the native /r/ - /l/ contrast from 6-8 to 10-12 months of age (Kuhl et al., 2006), and though infants learning a Filipino language do not discriminate the native nasal /na/ - /nga/ contrast at 6-8-months, its discrimination is clear by 10-12 months of age, in a case of either facilitation or perhaps even induction (Narayan, 2006; Narayan, Werker, & Beddor, under review).

Another kind of phonetic refinement is realignment, where an initial phonetic boundary does not line up exactly with native boundaries and later shifts to align. For example, the voice onset time (VOT)[2] boundary separating the /b/ and /p/ categories is placed at slightly different places across languages. Spanish-learning infants of 4.5-6 months discriminate the English, but not the Spanish VOT contrast (Lasky et al., 1975), suggesting a later shift in discrimination. A realignment was documented between 6-8 and 10-12 months of age in English-learning infants, who initially discriminated the French VOT boundary rather than the English one, but by 10-12 months infants had realigned perception to discriminate the same VOT boundary as English adults (Burns, Yoshida, Hill, & Werker, 2007).

Thus, although the initial sensitivities of infants allow language-general phonetic discrimination, subsequent experience refines these sensitivities in a number of ways, enabling facilitated perception of their native language. How does experience modify phonetic perception? One powerful mechanism that is available to infants in language learning is statistical tracking. Infants are able to perform sophisticated statistical analyses on the language they encounter, helping them to set the parameters of various aspects of their native language.

---

[2] VOT is the length of time from the "release" of the speech sound until the vocal chords begin to vibrate.

Statistical Language Learning

Word segmentation.

Statistical language learning mechanisms were first described in the realm of word segmentation. It is difficult to determine where one word ends and another begins because boundaries are not perceptually well marked; pauses are not systematically found between adjacent words. For adults, segmenting new words from the stream of speech is a relatively easy task because they are familiar with most of the surrounding words and syntactic frames. Infants can also take advantage of familiar words to help segment unfamiliar words from a stream of sound (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). However, infants know a relatively small proportion of the words that they hear in their language environment. Another source of information that could be leveraged in segmentation may come from tracking the frequency with which syllables co-occur (Hayes & Clark, 1970). For example, the syllable *pre* is often followed by *ty*, and *bay* is often followed by *by*, whereas *ty* is not often followed by *bay*. Thus, it is statistically likely that *pretty* and *baby* are cohesive words, whereas *ty bay* likely traverses a word boundary. Adults (Saffran, Newport, & Aslin, 1996) and infants (Saffran, Aslin, & Newport, 1996) are sensitive to these statistics, and after familiarization with a stream of syllables, treated those that often co-occurred as more likely to be cohesive words than syllables that did not often follow each other (even if the words and part-words occurred with the same absolute frequency; Aslin, Saffran, & Newport, 1998).

The ability to track these probabilities is not specific to language. If syllables are replaced with musical tones in a stream of sound, adults and 8-month-old infants learn the most probable sequential groupings of the tones (Saffran, Johnson, Aslin, & Newport, 1999). Rats are able to use the statistical speech segmentation information, suggesting that this sensitivity is not specific to language learning (Toro & Trobalon, 2005). Indeed, in the visual domain, 2-, 5-, and 8-month-old infants learn the statistically likely groupings of sequentially-presented objects (Kirkham, Slemmer, & Johnson, 2002). Relatedly, 9-month-old infants are able to learn the underlying combinations of objects that comprise multi-element scenes (Fiser & Aslin, 2002).

To what extent do infants actually rely on statistical information in word segmentation? Another, potentially competing cue is the prosodic characteristics of the stream. In English the first syllable of a word tends to be stressed, and syllables within the same word tend to be coarticulated. Infants develop sensitivity to the native prosodic pattern between 6-9 months of age (e.g., Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Houston, & Newsome, 1999). Often, transitional probabilities and prosodic cues will converge on the same word boundary, but if the two are pitted against one another, 6-month-old infants segment in accordance with the statistical over prosodic cues (Thiessen & Saffran, 2003), but 8-month-olds segment in accordance with the prosodic cues (Johnson & Jusczyk, 2001). This suggests that the statistics initially determine the word boundaries, which then help reveal the prosodic structure of the language.

Grammar learning.

Rule learning requires that infants analyze the structure of the input at a higher level. The different sequences *ga ti ti* and *li na na* both follow the format ABB, and the novel sequence *wo fe fe* would be consistent whereas *wo fe wo* would violate the structure. Neonates (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008) and 7-month-old infants (Marcus, Vijayan, Bandi Rao, & Vishton, 1999) are able to make these higher-order calculations, generalizing patterns to novel sequences.

As in statistical learning, rule-like learning has been demonstrated outside of language, with 7-month-old infants generalizing patterns of visually-presented dogs and cats (Saffran, Pollak, Seibel, & Shkolnik, 2007). However, it appears that infants are better able to generalize non-speech sound patterns if they are first instantiated in speech; the patterns in non-speech sequences are recognized if they are first learned in speech but not vice versa (Marcus, Fernandes, & Johnson, 2007).

If given a stream that is impregnated with both statistical segmentation probabilities and rule structure, the stream must first be segmented to uncover the grammatical regularities. Adults are able to segment statistically, but the grammar is only learned if pauses, even so brief as to not be detected, are inserted between the groupings consistent with the statistics (Peña, Bonatti, Nespor, & Mehler, 2002). It appears that adults most easily calculate rules across vowels and statistics most easily across consonants (Toro, Mehler, Nespor, & Bonatti, 2008; but see Berent, Marcus, Shimrom, & Gafos, 2002 and Newport & Aslin, 2004 for counter-evidence). This is consistent with the proposal that vowels and consonants play different roles in language, with vowels denoting grammatical structure (syntax) and consonants denoting word meaning (Nespor, Peña, & Mehler, 2003).

Infants are also sensitive to more complicated grammars. Eleven-to-12-month-old infants can learn a finite-state grammar, even generalizing to a new vocabulary (Gomez & Gerken, 1999). Learning of non-adjacent grammatical dependencies (e.g., aXb where *a* predicts *b*) is promoted by higher variability of the intervening element (larger X set size). Fifteen and 18-month-old infants are only able to learn the dependencies when the set size of X was 24 elements, but not 2 or 12 elements (Gomez, 2002; Gomez & Maye, 2005).

Statistics in word learning.

Words are classified as either content (nouns, verbs, adjectives, etc.) or function words (prepositions, articles, etc.), and this broad classification can be determined perceptually on the basis of a number of acoustic and phonetic cues (Shi, Morgan, & Allopenna, 1998), as well as by relative frequency (Gervain et al., 2008). From birth, neonates can summarize statistics across a number of perceptual dimensions to sort words into broad grammatical classes (Shi, Werker, & Morgan, 1999), and by 6-months of age this ability develops into a preference for content over function words (Shi & Werker, 2001), useful in acquiring a meaningful vocabulary.

Statistics are also useful in attaching words to their proper meaning. It is unlikely that every time a word is heard its proper referent is in view (or that it is where attention is directed), and not every time a referent is being attended to is its label heard. Nevertheless, it is possible that a word label is heard most frequently in the presence of its referent, and the word most frequently heard in the presence of a referent is its label. Adults (Vouloumanos, 2008) and 18-20-month-old infants (Vouloumanos & Werker, under review) presented with word-object pairings that co-occur with varying consistency are sensitive to these probabilities, using the most frequent pairings to learn the most likely word labels.

Distributional learning.

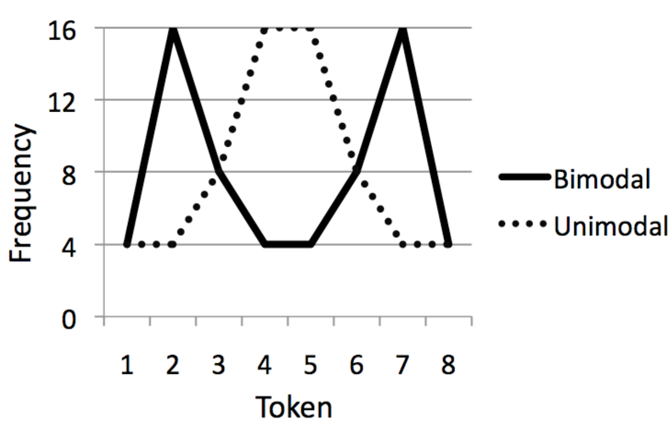
Given the broad range of sophisticated statistical mechanisms at infants' disposal to analyze and acquire the surrounding language, it stands to reason that one potential mechanism of phonetic development may arise through a statistical analysis of the sounds that are heard in the surrounding language environment. One way in which this could occur is through a relative frequency-based analysis.

The language environment surrounding the infant is hypothesized to contain certain distributional characteristics that arise from the phonetic structure of the language. Infants hear adults produce phones that are realized in a number of different ways. The precise articulation of each sound varies depending on a number of factors such as the other sounds being produced (coarticulation), the individual speaker producing the sounds (sex, high or low voice), emotional context, speaking rate, and so forth. As a result, each speech token heard sounds slightly different, but the range of sounds is encompassed by one phonemic category. The sounds that are heard the most frequently are those at the centre of the category, whereas sounds that are heard least frequently would indicate a category boundary. Thus, where two speech sound categories are distinguished, their frequency distribution should be bimodal: a mode marking each category centroid and a drop in frequency at the boundary between the two categories. On the other hand, if the sounds belong to the same category, there should only be one mode at the category centroid, with the sounds decreasing in frequency the farther away they are from the mode, forming a unimodal distribution.

The hypothesis that the frequency distribution of the sounds can reveal native phonetic category structure was tested in a cross-cultural analysis of the vowel distributions in English and Japanese. English vowels differ primarily in colour (formant frequency) and Japanese vowels differ in length (short or long). Therefore if distributional frequency information is available for infants to learn about the categories in their native language, the Japanese vowel length difference should form a bimodal distribution in the production of Japanese mothers, and the vowel colour difference should form a bimodal distribution in the production of English mothers. Measurements of mothers' productions revealed that the distributions were more varied in length from the Japanese mothers and more varied in colour from the English mothers. Further, the distributions of length were more predictive of two categories in Japanese whereas the distributions of colour were more predictive of two categories in English (Werker et al., 2007). Computational modeling of these data revealed that the distributions were sufficient for the sounds to be sorted into the language-specific categories, even when the model was not restricted to classification into one or two categories only (Vallabha, McClelland, Pons, Werker, & Amano, 2007). This work reveals that distributional information is present in the language environment, consistent with the idea that infants could use this information to learn about the speech sound category inventory in their language.

Infants' sensitivity to distributional frequency information as a possible statistical learning mechanism for changing phonetic categories was tested by Maye, Werker, and Gerken (2002). Six to-8-month-old infants were exposed to a distribution of speech sound tokens that are discriminable by English adults and young infants (Pegg & Werker, 1997). The distributional frequencies of the tokens presented differed between two groups of infants, such that those given to one group formed a unimodal distribution (tokens in the middle of the continuum were presented most frequently) and those presented to the other group formed a bimodal distribution (tokens closer to the endpoints were the most frequent) (see Figure 1). All infants were tested on their ability to discriminate the endpoint tokens, which had been presented with equal frequency in both conditions. If distributional learning is a mechanism of perceptual change, then familiarization with the unimodal distribution should mimic exposure to a single category and lead to decreased discrimination of its endpoints. The bimodal distribution should mimic a useful distinction between two categories (in this case, consistent with a priori discrimination), and maintain discrimination of the continuum endpoints. Indeed, only those infants familiarized with the bimodal distribution discriminated the two categories, indicating that the statistical information successfully collapsed discrimination of the contrast for infants in the unimodal condition.

FIGURE 1. BIMODAL AND UNIMODAL FREQUENCY DISTRIBUTIONS.
Redrawn from Maye, Werker, and Gerken (2002).



Another study examined infants' ability to enhance discrimination of a contrast that was not previously discriminated (Maye, Weiss, & Aslin, 2008). The stimuli used were either a /ga/ - /ka/ or a /da/ - /ta/ distribution. Infants in the bimodal but not in the unimodal condition later discriminated the contrast, indicating that statistical frequency information is used by infants to enhance, as well as collapse discrimination. This work further tested infants' ability to generalize the phonetic learning to other contrasts. Infants were familiarized to one continuum (e.g., /ga/ - /ka/) and then habituated and tested on the other (/da/ - /ta/). Infants were able to apply the distributional frequency information from one continuum to the different stimuli. This suggests that infants may be learning at the level of individual feature (i.e., voicing, place-of-articulation, manner) as opposed to the level of a phone (which synthesizes the features), and demonstrates the easy transfer of learning, suggestive of plasticity in the perceptual system. Together, these results strongly suggest that infants are able to keep track of and use the statistical distribution of phones in the language exposure.

Adults have also been tested on their ability to use distributional frequency information to change speech perception (Maye & Gerken, 2000). The contrast tested was /da/ - /ta/, the same as in the first infant demonstration of distributional learning of speech sound categories (Maye et al., 2002). English-speaking adults were exposed to the sounds in either a bimodal or unimodal distribution for 9 m, with participants asked to check a box every time they heard a sound stimulus presented. Following this learning phase, participants who heard the bimodal distribution were more likely to discriminate the endpoint sounds than those who heard the unimodal distribution. This learning did not appear to be generalizable like the infant learning (Maye & Gerken, 2001), suggesting that it is more difficult to affect adults' than infants' speech perception.

Distributional learning is not specific to the auditory domain. Language in the signed modality develops similarly to the auditory modality. Hearing infants reveal a developmental trajectory where 4-month-olds but not 14-month-olds, perceive hand shapes categorically (Baker, Golinkoff, & Petitto, 2006). Distributional learning appears to be an effective mechanism for learning handshape categories: Six-to-8-month-old infants who view a bimodal frequency distribution of handshapes, but not those who view a unimodal frequency distribution, subsequently discriminate two handshape categories (Werker, Pan, & Baker, in prep).

Support for distributional language learning.

Supportive evidence for a statistical frequency mechanism of perceptual reorganization is found in the earlier discriminatory decline of nonnative contrasts comprised of phones that appear more frequently in the language exposure (Anderson, Morgan, & White, 2003). The developmental trajectory of English-learning infants' discrimination was compared for two nonnative contrasts that appear in English with different relative frequencies: the more commonly appearing Hindi /t/ - /T/, and the rarer Thompson /k'/ - /q'/. Neither are discriminated by English-speaking adults. If statistics-tracking does in fact influence the development of the phonetic system, infants should stop discriminating earlier in life the contrasts that are comprised of phones which appear (noncontrastively) more frequently in the English language (stronger statistical evidence) than those that are less frequently found. In fact, although younger infants of 6.5-months were found to discriminate both the less and the more frequently appearing contrasts, 8.5-month-olds had more difficulty with the more frequent Hindi contrast than they did with the less frequent Thompson contrast. This suggests that the frequency of appearance does affect the trajectory of perceptual development, and is consistent with its development via a statistical learning mechanism.

The quality of language exposure may also affect infants' discrimination (Liu, Kuhl, & Tsao, 2003). Mandarin-learning infants of 6-8 and 10-12 months were tested on discrimination of a native phonetic contrast, and this performance was compared with the vowel space of their mothers when producing infant-directed speech (IDS). The area formed by the mothers' tokens significantly correlated with their infant's discrimination (greater acoustic distance between two vowels in maternal speech was predictive of better infant discrimination). Expanded vowel space is suggestive of exaggerated bimodal distributions, which would ease the statistical load, making the meaningful categories easier to distinguish (see Bosch & Sebastián-Gallés, 2003). Therefore, the correlational evidence of Liu et al. (2003) is consistent with a statistically-driven hypothesis.

Sensitivity to distributional information of speech sounds is not restricted to human beings. Rats are also able to use distributional speech sound category information to guide category discrimination (Pons, 2006), and similar distributional learning effects have been shown on European starlings' vowel perception (Kluender, Lotto, Holt, & Bloedel, 1998). Thus this ability extends beyond humans and is not specific to language learning. Indeed, evidence of distributional learning can be found in other domains.

Visual distributional learning.

The use of distributional information is also apparent in visual perception. In one study, adults were sequentially exposed to vertical lines of varying thickness (Rosenthal, Fusi, & Hochstein, 2001). The frequency distribution with which each grade of thickness appeared affected participants' category judgments, with the most frequently presented stimuli generally judged as category centroids and the least frequently presented stimuli judged as boundaries. Five- and 7-year-old children's memory is also affected by the frequency distribution of object sizes: Memory of object size is biased more towards the category centroid when viewing a unimodal distribution than a flat distribution (Duffy, Huttenlocher, & Crawford, 2006).

Limitations of distributional learning.

In contrast to the bottom-up statistical hypothesis, a top-down mechanism of phonetic development stems from lexical knowledge (Hayes-Harb, 2007). Adults were trained on distributional statistics, and revealed evidence of phonetic learning, distinguishing a contrast that was presented in a bimodal distribution but not a unimodal distribution (Hayes-Harb, 2007; see also Maye, 2000). However learning was even stronger when a single sound from each category was consistently associated with a unique visual item (e.g., sound-A with a pot and sound-B with a mouse), and learning was impaired when the sounds were associated with the same item (sound-A and sound-B both with a pot). This implies that assigning distinct meaning to different sounds may be a powerful cue to their individuation, whereas shared meaning leads to collapse of the distinction.

Lexical associations have been shown to affect phonetic learning at 9 months of age, when phonetic reorganization is underway. When different sounds are consistently attached to the same objects (i.e., sound-A is always paired with object-A; sound-B with object-B), infants discriminate the sounds. But when the pairing is inconsistent (sound-A is sometimes with object-A and sometimes with object-B), the infants fail to distinguish the sounds even though the same distributional information is present (Yeung & Werker, 2006; under review).

Similarly, visual (facial) information contributes to infants' phonetic identification (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999; Rosenblum, Schmuckler, & Johnson, 1997). If 6-month-old infants hear a unimodal distribution of speech sounds such that each of the sounds is paired with a single facial configuration the infants will, unsurprisingly, not discriminate the sounds. However, if sounds from one half of the unimodal distribution are associated with one facial configuration and sounds from the other half are associated with a different facial configuration, infants will discriminate the contrast (Teinonen, Aslin, Asku, & Csibra, in press). Thus, distributional learning is likely not powerful enough on its own to overcome conflicting sources of visual information regarding phonetic categories.

Presumably, all of the above cues often converge in the natural language environment. Phonetic categories are marked by bimodal distributions, different facial configurations, and different meanings. The evidence that infants are sensitive to all of these cues during perceptual organization suggests that the sources of information may be synthesized to help infants learn speech sound categories.

Attention

Preferential attention to language.

Attention to language information is thought to aid in its acquisition. Infants are biased to attend to language sounds (Vouloumanos & Werker, 2004; 2007a; 2007b). Two-, 5- and 7-month-old infants look longer at a visual attention-getter when it is paired with speech sounds than non-speech sounds (Vouloumanos & Werker, 2004). The language bias is not specific to the auditory domain, but extends to the visual communicative domain. Six-month-old infants prefer looking at unfamiliar linguistic (American Sign Language) hand shapes over unfamiliar non-linguistic control stimuli (Krentz & Corina, 2008), paralleling auditory language preferences.

Selective attention to language sounds is not found in individuals with ASDs, who experience language difficulties (Baron-Cohen, 1995). Toddlers with ASDs fail to show a preference for speech sounds (Paul, Chawarska, Fowler, Cicchetti, & Volkmar, 2007) and children with ASDs may actually disprefer speech (Kuhl, Coffey-Corina, Padden, & Dawson, 2005). A preference for speech in individuals with ASDs is associated with enhanced language abilities, as measured by neural register of a phonetic change (Kuhl et al., 2005) and receptive vocabulary (Perl et al., 2007), highlighting the importance of a bias for speech to learn language.

In addition to a general preference to listen to language sounds, attention is further focused specifically on the native signal via infants' preference for the rhythm of the native language. Sound is processed by third-trimester fetuses (Birnholz & Benacerraf, 1983), leading to familiarity with the rhythm of their mother's speech (DeCasper & Spence, 1986; DeCasper, Lecanuet, Busne, Granier-Deferre & Maugeais, 1994). Infants are able to discriminate different languages based on their rhythm (Nazzi, Bertoncini, & Mehler, 1998; Nazzi, Jusczyk, & Johnson, 2000). This rhythmical preference forms the basis of a preference to attend to the native language (Dehaene-Lambertz & Houston, 1998; Mehler et al., 1998; Moon, Cooper, & Fifer, 1993).

Does attention aid in learning?

The processing of frequency information may require attention (Hasher & Zacks, 1984; Sanders, Gonzalez, Murphy, Liddle, & Vitina, 1987), with no age-related differences found between childhood and adulthood, implying that attention is similarly required at all ages (Hasher & Chromiak, 1977; Hasher & Zacks, 1979). The level of the stimuli at which attention is focused impacts frequency processing: Word frequency judgments suffer if participants are instructed to focus their attention on the number of consonants and syllables (Rowe, 1974) or graphemic features, but not acoustic or semantic features (Rose & Rowe, 1976).

Attention also appears to play a role in conceptual categorization processes. Comparison paradigms involve a series of forced choices of which item is, for example, larger (or smaller) than the other (Cech & Shoben, 2001). Categorization of the items influences the speed and accuracy of judgments. For example, judgments are faster (easier discrimination) for pairs that are more distant from each other, and judgments are slower (more difficult discrimination) for pairs that are closer to the centre of a category than those near the edges (Banks, 1977; Moyer & Dumais, 1978). Performance is enhanced if the items are recategorized; even if all of the items are fairly small, reconceptualizing the relatively larger ones as “large” will create categorical differences, facilitating performance on the task. Participants do not successfully regroup (as revealed by faster, more accurate performance) when filler questions are included (order, duration) in addition to size questions, directing attention to different aspects of stimuli, and hence disrupting sustained attention to size (Cech & Shoben, 2001).

In language learning, initial studies revealed no strong evidence that attention plays a role in statistical word segmentation. Adults and children who drew while the experimental sounds played in the background were found to have attenuated performance as compared to attentive familiarization, although performance was still above chance (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, when attention was targeted with a more demanding attention-distracting task, another pattern emerged. Adults in the high-load condition completed unrelated tasks while listening to the speech stream, pressing a button either every time they detected auditory or visual image repetitions in concurrently-presented information streams, or every time the pitch changed in the target speech stream (Toro, Sinnett, & Soto-Faraco, 2005). Performance in the high-load conditions was not significantly above chance, whereas statistical learning was still effective in incidental conditions (no concurrent task). Further, infants' performance is facilitated when attention-getting (infant-directed) speech is used (Thiessen, Hill, & Saffran, 2005). This pattern of results suggests that attention plays an important role in the statistical segmentation of speech sounds.

Attention in phonetic learning.

Evidence that attention might be important in phonetic learning is found in the adult second language learning literature. Learning appears to be facilitated with explicit feedback, where labeling the sounds helps to highlight specific acoustic cues, allowing them to be weighted more heavily (Francis, Baldwin, & Nusbaum, 2000; Francis & Nusbaum, 2002).

If attention is important, focus on the phonetic form should result in easier learning than when attention is directed towards other aspects of stimuli. This was experimentally assessed by directing adult participants' focus on either the phonetic or semantic (meaning) properties of nonnative word stimuli (Guion & Pederson, 2007). Three different nonnative contrasts were tested. For the easiest contrasts, discrimination was already at ceiling at pretest, or improved equally in both the phonetic and semantic conditions. Thus, for easier nonnative contrasts, phonetic learning occurs irrespective of attention. However attention did have an effect on the most difficult contrast, a Hindi voiceless aspirated dental – retroflex contrast (/ta/ - /Ta/), for which discrimination improved much more in the phonetic than semantic condition. This suggests that though attention may not be necessary in all phonetic learning, it is important in learning difficult nonnative contrasts.

Social Interaction Draws Attention in Phonetic Development

Work on social interaction in phonetic learning also supports a role for attention. Infants are born with social biases that direct attention to the sources of language, for example preferring to look at face-like stimuli (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991) and faces that have direct rather than averted eye gaze (Farroni, Csibra, Simion, & Johnson, 2002; Southgate, Csibra, Kaufman, & Johnson, 2008). Situations with social interaction thus likely also attract greater attention.

Birdsong.

Birdsong has given us some insight into the role of attention and social interaction in language learning. Both human and bird species' communicative signals rely on a combination of experience and guiding innate-like mechanisms. Where human infants have a preference for speech sounds (Vouloumanos & Werker, 2004; 2007), young birds discriminate their conspecific (species') song from non-conspecific song, showing differential heart rate in response to each (Dooling & Searcy, 1980). Young birds also respond with more numerous vocalizations to conspecific song than to non-conspecific song (Nelson & Marler, 1993), suggesting that young birds have a preference for the signal of their own species that predisposes them to its acquisition (Marler, 1990).

Birds' bias for the conspecific signal is supplemented by social factors in learning song. One way in which this is observed is in the effect of social interaction on the age at which song develops. There is a sensitive period for learning song, during which young birds must receive exposure for normal song to develop (Thorpe, 1958). However, the documented sensitive periods were found using tape-tutors, and live tutors were found to be much more effective than tape-tutors, with the live-tutored birds learning song even after the close of the tape-tutored sensitive period (Baptista & Petrinovich, 1984; Nordby, Campbell, & Beecher, 2001). Further, the live-tutored birds even learn non-conspecific song that is typically rejected when tape-tutored (Baptista & Petrinovich, 1984). The effect of social interaction on learning is so strong that zebra finches, a particularly social species, learn the non-conspecific song of a tutor who is feeding them, overriding the concurrently presented signal of the otherwise-preferred conspecific song (Immelmann, 1969). Zebra finches fail to learn even conspecific song from live tutors if they cannot see or interact with them (Eales, 1989), but they learn if the live tutors are in the same cage and perform grooming behaviors, even if they cannot see them (Adret, 1993). This work suggests an important influence of social interaction (and therefore greater attention) on song development: Song is more easily learned and can be learned later in life with social interaction and greater attention.

Social interaction and attention in phonetic development.

A possible facilitative effect of attention in human phonetic learning is also found in work contrasting infants' learning from live versus taped language exposure (Kuhl et al., 2003). English-learning infants were exposed to Mandarin Chinese from 9 months of age, at point when perceptual reorganization in response to native sounds is underway (Werker & Tees, 1984a). In the live condition, infants were exposed to Mandarin via a native speaker who read books and played with toys. There were two pretaped conditions, an audio-video (AV) condition where video of an interactive person was recorded from the viewpoint of the infant, and an audio (A) condition, which contained only the soundtrack. All three groups received 5 hours of Mandarin over 12 sessions in 4 weeks, after which infants were tested on their ability to discriminate a Mandarin-specific contrast. Infants in the live condition paid significantly more attention than infants in the pretaped conditions. Only those infants who had heard the Mandarin in the context of live interaction (and therefore paid greater attention) discriminated the Mandarin contrast better than infants in a control condition, who had the same types and amounts of language exposure with an English speaker.

Presumably, the Mandarin exposure (that was similar in all three experimental conditions) contained tokens of the nonnative Mandarin contrast that formed a bimodal distribution, in contrast to the tokens in the ambient English language environment. Some minimally-paired Mandarin words may have also been heard, although it is unclear whether they would have been associated with different visually-presented objects (Yeung & Werker, under review). Thus, the taped conditions of Kuhl et al. (2003) presented infants with information that is sufficient in changing perception at younger ages. However, in the absence of live social interaction and the correspondingly higher levels of attention, at this age infants did not change their perception.

A follow-up to the Kuhl et al. (2003) experiment provided further evidence that attention may be the more specific factor in phonetic change. Infants of 9.5-10.5 months were taught Spanish-contrast minimal-pair names for different toys through live interactions over a month. Infants' eye gaze was monitored, coding for whether the infants' gaze shifted from the toy to the experimenter's face when a new toy was presented. Presence of gaze shifting was significantly correlated with ERP measure of the nonnative phonetic change two weeks after the exposure and marginally correlated with behavioural discrimination of the nonnative contrast one week after the exposure (Conboy, Brooks, Taylor, Meltzoff, & Kuhl, 2008). Together, these results suggest that attention may be an important facilitating factor in distributional phonetic learning.

Phonetic Plasticity

Attention may not be important in phonetic learning at all ages. Younger infants' phonetic discrimination is affected in paradigms that do not explicitly draw attention. The infants in the Maye et al. (2002) distributional learning study were 6 – 8 months of age and those in the Yeung and Werker (under review) object-speech sound pairing study were 8.5 – 9.5 months. Both of those groups are slightly younger than the infants in Kuhl et al. (2003) who were reported to be, on average, 9.3 months of age when they began receiving the Mandarin exposure, which lasted 4 weeks. The phonetic discrimination test then took place between 2 – 12 days after the exposure was complete; thus, infants probably averaged about 10.5 months of age when their discrimination was tested. Similarly, the infants in Conboy et al. (2008) were reported to be 9.5-10.5 months. These age differences are significant because by 10 months of age, infants' perception of consonant contrasts has attuned to native phonetic categories. It is possible that attention may only be important after perceptual reorganization, and less important earlier in life.

Why might phonetic learning require attention only after reorganization? Although perceptual attunement is extremely valuable in first language (L1) perception (Werker & Tees, 1999), when the inventories of native and L2 phonetic distinctions conflict, the minimization of L1-irrelevant phonetic differences that are relevant in the L2 makes L2 perception more difficult. Plasticity in speech sound category discrimination before 10 months of age should be easier to demonstrate because native perceptual categories would not need to be overridden for any experimentally induced change. The infants in Kuhl et al. (2003) and Conboy et al. (2008) would have been exposed to the Mandarin stimuli at an age at which their perception was less plastic than those tested in Maye et al. (2002) and Yeung and Werker (under review). This raises the possibility that the difference in findings, where Kuhl et al. (2003) find that phonetic learning may be facilitated by attention (in contrast to the other two studies), could be due to a need for attentional or social factors only when changing already-established phonetic categories.

Once phonetic categories have formed it is more difficult even to perceive nonnative phones, let alone to develop discrimination of the contrasts. According to the Perceptual Assimilation Model (PAM; Best, 1995; Best, McRoberts & Sithole, 1988), nonnative speech sounds are perceived in terms of the L1 categories, and discrimination depends on the particular sounds in question. Unfamiliar phones are perceptually assimilated to existing categories in one of three ways, determined by their degree of difference from the closest existing L1 categories. The sounds could be mapped onto a native category, considered an uncategorizable speech sound, or not assimilated to speech at all. Excellent discrimination can result if each nonnative sound is assimilated to a different L1 category, or even if they are both mapped onto a single L1 category but one is a much better exemplar than the other ("category-goodness difference"). However, if the two sounds are mapped onto the same L1 category with a slight or no goodness difference between them, discrimination will likely be quite poor. This model does not incorporate any possibility of learning the nonnative speech sound categories, which are only perceived through the filter of the developed L1 perception.

In contrast, the Speech Learning Model (SLM; Flege, 1995) allows for the possibility of novel category formation, depending on the degree of dissimilarity between the L1 and L2 phonemes, and the age of learning the L2 (AOL). Greater dissimilarity between the categories and an earlier AOL are theorized to lead to a higher probability of new category formation. For example, the single liquid manifested in Japanese is perceptually more similar to the English /l/ than /r/ for native Japanese speakers (Sekiyama &Tohkura, 1993; Takagi, 1993). The SLM predicts that native Japanese speakers can more easily form a category for the more distant English /r/ as opposed to the closer, more similar /l/, and when exposed to English /r/-/l/ words (e.g., *rock - lock*), native speakers of Japanese were able to correctly identify more words containing /r/ than /l/ (Flege, Takagi, & Mann, 1995).

Adults are sensitive to distributional phonetic information, discriminating sounds better after exposure to a bimodal distribution (Hayes-Harb, 2007; Maye & Gerken, 2001), but in these experiments they were explicitly instructed to attend to the stimuli. Together with the attentional difference in the social manipulation of Kuhl et al. (2003), the existing literature suggests that phonetic change is possible after native category formation, but that it may require attention. However, this question has not yet been directly tested.

The primary questions addressed by the current research are: (1) how does the native language perceptual reorganization affect perceptual learning, and (2) is there a facilitative role for attention in phonetic learning after phonetic reorganization?

OVERVIEW OF EXPERIMENTS

Perceptual Plasticity

Plasticity is characteristic in phonetic learning in infants younger than 10 months, with perception affected by distributional frequency information at 6-8 months (Maye et al., 2002; 2008). At these ages, perceptual categories are changing in response to input from the ambient language. Thus, changing perception at 10 months of age is a fundamentally different question than that addressed by the previous research testing 6-8-month-olds, whereby testing infants after perceptual reorganization can reveal the effects of native category formation on phonetic plasticity. The distributional learning mechanism appears to be functional through childhood and adulthood in the visual domain (e.g., Duffy et al., 2005; Rosenthal et al., 2001), suggesting that if it is a mechanism of phonetic development, change maybe possible after perceptual reorganization. Distributional learning has been shown to affect discrimination in adults who are clearly post-perceptual reorganization (Hayes-Harb, 2007; Maye & Gerken, 2000, 2001), however, these data were obtained following a longer familiarization phase (9 mins versus 2 mins in the infant studies) in which participants were explicitly instructed to pay attention. This work leaves open the question of how easily phonetic change can be effected later in life.

Different perceptual contrasts are used across the experiments. All of the stimuli are drawn from a Hindi dental-retroflex place-of-articulation continuum, but Experiment 1 uses the sounds that occur with greater frequency in English, and are closer to the English category centroid. Infants ultimately fail to change perception in this initial experiment, therefore Experiments 2 and 3 use sounds from along the same continuum that are farther from the English category centroid and heard less frequently in English. The sounds in Experiment 1 are likely more challenging for infants to distinguish for two reasons. First, in categorical perception, two items near the centre of the category are more difficult to discriminate than two items nearer to the category boundary (Liberman et al., 1957). Second, a statistical explanation for perceptual development would suggest that more experience would result in greater difficulty of perceptual change. Nonnative contrasts that are heard more frequently are collapsed earlier than those that are heard less frequently (Anderson et al., 2003), and contrasts that are collapsed earlier would be likely require more information to overcome the learned distributions.

Attention

Given the role that attention has been found to play in statistical word segmentation (Thiessen et al., 2005; Toro et al., 2005), infants' level of attention was monitored during the learning phase to test whether attention plays a role in speech sound category change at this age. As noted, previous work has found that adults are able to use distributional speech sound category information. However, in those experiments, participants were explicitly instructed to attend to the stimuli, making a check mark every time they heard a token. This leaves open the possibility that adults' distributional learning may require attention, and it is unknown whether incidental exposure would also have led to a change in adult discrimination.

In these experiments, attention to the auditory stimuli was measured by looking time to a visual attention-getter. Auditory and visual attention correlate, with 6- and 9-month-old infants who pay more attention to a visual stimulus showing greater auditory discrimination, indicative of greater auditory attention (Panneton & Aslin, in preparation). Visual attention correlates with other measures of attending such as heart rate (Colombo, Richman, Shaddy, Follmer Greenhoot, & Maikranz, 2001; Richards, 1987) and event-related potentials (Reynolds & Richards, 2005).

EXPERIMENT 1

This first experiment was the initial query into whether infants' speech perception is changeable after native phonetic categories have developed. Ten-month-old infants were tested, an age at which native consonant categories are evident. Infants' levels of attention during the learning phase were also monitored to reveal whether attention plays a role, with the hypothesis that infants who pay more attention show stronger evidence of learning.

Method

Participants.

Twenty-four healthy, 10-month-old monolingual English-learning infants participated in this study (12 females; *M* age = 10 months 13 days, range = 10 months 0 days – 11 months 0 days). An additional 8 infants were tested and excluded for the following reasons: excessive fussiness (6) and parental interference (2).

Apparatus.

The experiment was controlled by a Macintosh G4 computer with the Habit 2000 program (Cohen, Atkinson, & Chaput, 2000) from a remote room. The study was conducted in a 280x226 cm sound-attenuated room dimly lit by two 60W shaded floor lamps, one to the left and one to the right. Images were viewed on a LCD television screen to the front, with black cloth covering the rest of the front wall. A digital video camera peeked out of a small hole in the cloth below the screen, relaying an image of the infant to the control room and recording looking behaviour. The video recordings were later digitized using Final Cut Pro for off-line coding of looking behaviour in a frame-by-frame analysis (30 frames per second). Audio speakers were hidden behind the cloth to the left and right of the television, playing sounds between 68-72 dB SPL.

Procedure.

Infants were tested individually on their parents' lap. Parents wore headphones playing female vocal music and, like the experimenter, were unable to hear the auditory stimuli being played.

Each trial was initiated only when the infant fixated on the flashing attention-getter stimulus. The first, pretest trial served to acquaint infants with the audio/visual set-up. This was followed by the 2 min familiarization phase, and finally the test phase.

English-learning infants of 10 months have collapsed discrimination of the contrast, therefore the manipulation of interest was a categorical induction, requiring a bimodal distribution. All infants in this experiment received familiarization to the tokens in a bimodal distribution (see Figure 2).

The test phase consisted of two types of test trials (Maye et al., 2002, adapted from Best & Jones, 1998). Alternating trials featured alternation of tokens from opposite ends of the continuum. Nonalternating trials featured only a single token repeated. If infants discriminate the contrast they should show greater looking to one of the two types of trials. Many methods used to test infants rely on their well-described preference first for familiarity and then for novelty (Fantz, 1964; Hunter & Ames, 1988). In the alternating/non-alternating paradigm, a preference for the alternating trials may have more face validity because they present a greater variation in stimuli and therefore may be more novel and interesting. Indeed, in the absence of a familiarization phase, preference for alternating trials is typically found (Best & Jones, 1998). However, in phonetic learning paradigms, infants' discrimination is revealed via preference for nonalternating trials (Maye et al., 2002; Yeung & Werker, 2006, under review). In these learning phases, infants hear tokens from two different categories, and the alternating trials would perpetuate this two-category auditory experience. Therefore the nonalternating trials are actually more novel because they present tokens from only one category, resulting in infants' novelty preference for the nonalternating trials (Maye et al., 2002). Therefore, a preference for the nonalternating trials is predicted.

Stimuli.

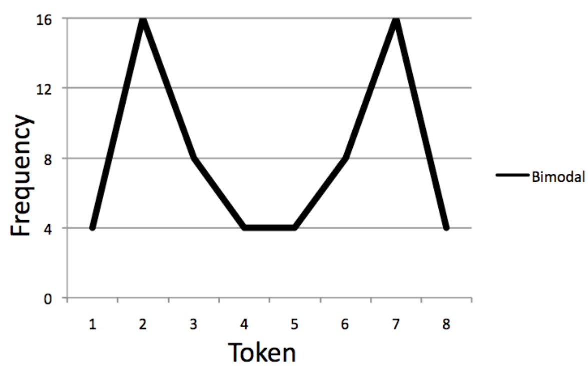
The contrast chosen was a Hindi voiced place-of-articulation dental-retroflex contrast. This is a difficult contrast which English-learning infants no longer discriminate by 10-12 months of age (Werker & Tees, 1984a), and on which they are already experiencing difficulty by 8.5 months (Anderson et al., 2003). The 8 target tokens were tokens 7 – 14 from a 16-step synthetic continuum where tokens 10 and 11 fell at the point judged by the Hindi speakers as the dental-retroflex category boundary (Werker & Lalonde, 1988).[3] The continuum was created by varying the second and third formants (F2 and F3)[4] where the starting frequency of F2 varied in equal 50 Hz steps from 900 to 1600 Hz, and the starting frequency of F3 varied in 48 Hz steps from 2240 to 2912 Hz. The vowel /a/ was selected to form consonant-vowel syllables (/da/ - /Da/), each 276 ms in length.

Filler tokens were included from along a synthetic /ra/ - /la/ continuum to help maintain infants' interest (Maye et al., 2002). These were also 276 ms long. The target

[3] From this point forward the tokens will be relabeled as 1-8 to accurately reflect their use in the current studies.
[4] Formants are individual frequency components of the speech signal.

and filler tokens were strung together into blocks of 24 tokens with an ISI of 500 ms (Maye et al., 2002; Werker & Tees, 1984b). Sixteen of the 24 tokens in each block were target tokens. The remaining 8 tokens in each block were fillers (one of each). Six blocks such were concatenated to create a familiarization sequence totaling 1 minute 52 s.

FIGURE 2. BIMODAL DISTRIBUTION OF TOKENS.



The test phase presented four trials of each of the two types (alternating and nonalternating) that were interleaved with one another. The alternating trials featured an alternation between tokens 1 and 8. The nonalternating trials featured a single token, either token 3 or token 6. Each trial was 10 s long with ISIs of 1 second.

A 12 s audio file containing 10 pronunciations of a female speaker uttering the nonsense word *pok* was used for the pretest.

Four images were used. A spinning toy waterwheel was the pretest image. During the familiarization phase infants viewed an image of a digital string of colourful dots forming and reforming a flower shape. In the test phase a static, unbounded black-and-white checkerboard was presented. A flashing, colour-changing ball was the attention-getter stimulus shown before every trial.

Results

Familiarization.

Infants' looking times averaged 63.1s (*SD*=15.9), with a median looking time of 61.9s. To investigate whether level of attention during the familiarization played a role in infants' learning, a median split was performed on the familiarization looking times, separating infants into high attending and low attending groups (*N*=12 each). The high attending infants attended for significantly longer than the low attending infants [$t(22)=-4.52$, $p<0.001$, Cohen's $d=-1.853$].

Test trials.

To assess discrimination of the sounds, looking times during the test trials were compared using a paired-samples *t*-test (see Table 1 for descriptives). Taken as one group, infants did not discriminate the two types of test trials [$t(23)=0.638$, *ns*; Cohen's $d=0.098$].

When split into high and low attenders, neither group discriminated the sounds [high-attenders: $t(11)=-0.039$, *ns*; Cohen's $d=-0.013$; low-attenders: $t(11)=-0.92$, *ns*; Cohen's $d=-0.19$]. To explore whether there was an underlying linear relationship between amount of attending and phonetic learning, familiarization looking time was correlated with a "nonalternating bias" score, created by subtracting looking time to alternating trials from looking time to nonalternating trials. The correlation was not significant ($r=.019$, *ns*).

Table 1. Experiment 1 Results: Looking Time (s)

All Infants

	Familiarization	Test Trials: Alternating Trials	Test Trials: Nonalternating Trials
	mean (SD)	mean (SD)	mean (SD)
	63.10 (15.86)	5.27 (1.63)	5.12 (1.43)
By Attention Level			
High Attending	73.87 (13.33)	5.22 (1.61)	5.24 (1.46)
Low Attending	52.32 (9.75)	5.31 (1.72)	5.01 (1.46)

Discussion

The goal of Experiment 1 was to probe whether 10-month-old infants could change their speech sound category discrimination in response to distributional frequency information. These infants have developed native speech sound categories, and thus change may be more difficult than at 6-8 months of age, the age at which an ability to use distributional information to change phonetic discrimination was revealed (Maye et al., 2002, 2008).

The results of this experiment suggest that after perceptual reorganization, brief familiarization to distributional speech sound category information is not sufficient to change discrimination. It had been hypothesized that infants who paid more attention may show more evidence of learning, but this was not borne out by the data. Infants who paid more attention did not show greater evidence of phonetic learning. Infants' failure to learn in this paradigm tentatively supports the notion that social interaction is necessary in phonetic change.

These results do not rule out the possibility that different phonetic distinctions may still be malleable at this age. For example, exposure to a less-frequently heard continuum could reveal plasticity. A statistical explanation of perceptual development predicts that it is easier to change discrimination of sounds that are less frequently heard, as exposure to a new distribution would more easily overwrite the initially-existing distribution for which there is less data, thus the new distribution has greater prominence. Further, models of second language speech perception such as PAM (Best, 1995) and SLM (Flege, 1995) state that the greater the perceptual distance between the familiar native categories and the new L2 categories, the more easily the L2 categories will be perceived. Experiment 2 tests the hypothesis that infants may more easily change perception of sounds that are less frequently heard in the native language.

EXPERIMENT 2

The infants in Experiment 1 failed to change their perception in response to distributional speech sound information. The sounds used in that experiment appeared frequently in the native language input, and were close to the centre of the native category. The present experiment attempted to find evidence of phonetic plasticity at 10 months of age by altering the sound contrast. This experiment used sound tokens that English-learning infants hear with less frequency in the native language environment, and that are farther from the English category centroid.

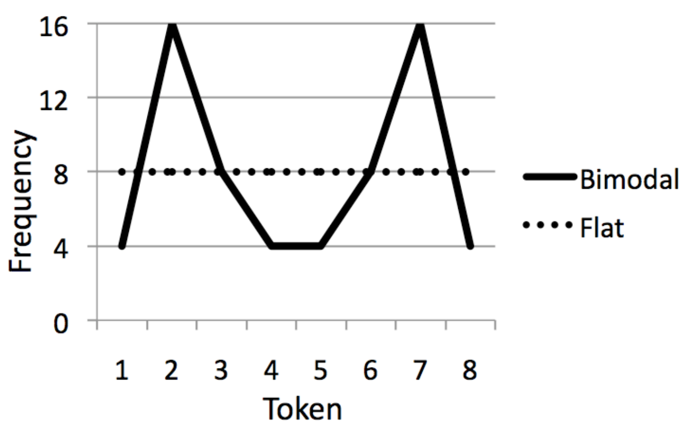
Method

Stimuli.

The stimuli used were identical to Experiment 1, with only a few exceptions. The 8 target tokens were chosen from the same larger continuum (Werker & Lalonde, 1988) also used in Experiment 1. However, the chosen tokens were tokens 9-16 of the original 16-step continuum, in contrast to Experiment 1, which used tokens 7-14. This resulted in a continuum of 8 tokens that was shifted slightly from those of the first study towards the retroflex end of the continuum. These are less typical instances of the English *d* phone, and they are stimuli with which English-learning infants would have less experience.

The manipulation of interest was the same as Experiment 1, a categorical induction, requiring a bimodal distribution. As a between-subjects control, a flat distribution was also tested, including an equal number of tokens from each point on the continuum. This distribution was expected to be treated as noise by infants in this condition, but at a minimum the flat distribution would bias infants less towards collapsing the distinction than would a unimodal distribution (Duffy et al., 2005), helping to confirm that infants' a priori perception did not discriminate between the nonnative sounds. Filler tokens, ISIs, and visual stimuli were the same as Experiment 1.

FIGURE 3. BIMODAL AND FLAT FREQUENCY DISTRIBUTIONS.



Apparatus and procedure.

These were identical to Experiment 1, with the addition of the MacArthur-Bates Communicative Development Inventory (MCDI; Fenson et al., 1994), which was mailed out when each infant reached 15 months of age. The MCDI consists of a vocabulary checklist where parents mark off whether their child understands and/or produces (says) each lexical item. This provided a follow-up assessment of infants' vocabulary development.

Participants.

Forty-eight healthy, 10-month-old monolingual English-learning infants participated in this study. Twenty-four were in the bimodal condition (12 females; *M* age=10 months 12 days, range=10 months 3 days – 10 months 30 days), and 24 were in the flat condition (12 females; *M* age=10 months 14 days, range=10 months 0 days – 10 months 28 days). An additional 14 infants were tested and excluded for the following reasons: excessive fussiness (10), equipment failure (2), ear infection (1), and parental interference (1).

Results

Familiarization.

In the familiarization phase, visual attention did not differ between the bimodal and flat conditions [$t(46)=0.28$, *ns*; Cohen's $d=0.080$). As in Experiment 1, infants in each condition were split into groups of high and low attenders based on the median looking times of their condition (see Table 2). The median looking times fell at 60.2 s (bimodal) and 61.8 s (flat). In each condition the high attending group looked for longer than the low attending group [bimodal condition: $t(22)=-6.66$, $p<0.001$, Cohen's $d=-2.72$; flat condition: $t(22)=-4.73$, $p<0.001$, Cohen's $d=-1.93$]

Test trials.

To assess discrimination of the sounds, looking times during the test trials were submitted to a 2 (condition: bimodal versus flat) x 2 (test trial type: alternating versus nonalternating) between-within ANOVA. There was a main effect of condition [$F(1, 46)=3.98$, $p=0.052$; $\eta_p^2=0.080$], with infants in the flat condition looking more overall in the test trials than infants in the bimodal condition. There was no main effect of test trial type [$F(1, 46)=1.49$, *ns*; $\eta_p^2=0.031$], and the interaction was not significant [$F(1, 46)=0.029$, *ns*; $\eta_p^2=0.001$]. Follow-up *t*-tests reveal that neither the bimodal group [*t*(23)=-1.18, *ns*; Cohen's *d*=-0.11] nor the flat group [*t*(23)=-.650, *ns*; Cohen's *d*=-0.090] distinguished between the two types of test trials.

Paired-samples *t*-tests were performed on the high- and low attending groups, and revealed that the infants in the high attending bimodal group discriminated the sounds [*t*(11)=-2.33, $p = 0.04$; Cohen's *d*=-0.23] whereas the other groups did not [low attending bimodal condition: *t*(11)=0.05, *ns*; Cohen's *d*=0.017; high attending flat condition: *t*(11)=-0.788, *ns*; Cohen's *d*=-0.16; low attending flat condition: *t*(11)=0.094, *ns*; Cohen's *d*=0.016]. A "nonalternating bias" score was created by subtracting looking time to alternating trials from looking time to nonalternating trials, and its correlation with familiarization looking time was not significant in either the bimodal [*r*=0.028, *ns*] or flat conditions [*r*=-0.067, *ns*].

Table 2. Experiment 2 Results: Looking Time (s)

All Infants	Familiarization	Test Trials	
		Alternating Trials	Nonalternating Trials
	mean (SD)	mean (SD)	mean (SD)
Bimodal Condition	65.10 (16.80)	3.99 (1.86)	4.19 (1.71)
Flat Condition	63.49 (23.08)	4.98 (1.76)	5.14 (1.76)
By Attention Level			
Bimodal (high)	78.55(12.32)	4.88 (1.91)	5.28 (1.63)
Bimodal (low)	51.65 (6.64)	3.09 (1.32)	3.11 (0.96)
Flat (high)	79.53 (14.66)	5.11 (1.66)	5.39 (1.77)
Flat (low)	47.44 (18.36)	4.85 (1.93)	4.88 (1.66)

MCDIs.

Univariate ANOVAs were performed on the MCDIs that were returned from parents of 34 of the 48 infants (*N*=15 in the bimodal condition and *N*=19 in the flat condition). Vocabulary comprehension did not differ significantly between infants in the two conditions [$F(1,30)=2.77$, *ns*; $\eta_p^2=0.084$; bimodal infants $M=55$, $SD=23$; flat infants $M=42$, $SD=23$], nor between the high and low attending groups [$F(1, 30)=2.28$, *ns*; $\eta_p^2=0.00$; high attending infants $M=48$, $SD=22$; low attending infants $M=47$, $SD=22$], and the interaction was not significant [$F(1, 30)=0.20$, *ns*; $\eta_p^2=0.007$].

Vocabulary production did not differ between condition [$F(1, 30)=0.077$, *ns*; $\eta_p^2=0.003$; bimodal $M=10$ words, $SD=3$; flat infants $M=20$ words, $SD=6$] or attention level [$F(1, 30)=1.50$, *ns*; $\eta_p^2=0.048$; high attending $M=19$ words, $SD=7$, low attending $M=11$ words, $SD=3$], and there was no interaction between the two [$F(1,30)=0.60$, *ns*; $\eta_p^2=0.020$].

Discussion

This experiment followed up on the null results of Experiment 1, using different sound tokens to once again test infants of 10 months on their ability to use distributional information to change their perception of a Hindi contrast that is no longer discriminated at this age. A bimodal distribution was presented to mimic exposure to two categories. In contrast to infants of 6-8 months who use distributional information to guide speech sound category discrimination (Maye et al., 2002, 2008), the results of Experiment 2 (like those of Experiment 1) suggest that overall, distributional exposure does not affect the perception of these older infants, who did not discriminate the sounds after exposure to the bimodal distribution.

Interestingly, although looking times were equivalent between the infants in the bimodal and the flat conditions in the familiarization phase, the bimodal infants looked for less time in the test trials. This suggests that the perceptual difference between the bimodal and flat distributions may be recognized, even if discrimination remains unchanged. This test trial disparity is analogous to that found by Maye et al. (2002), where the infants receiving the distribution that conflicted with a priori discrimination (in that case, the unimodal condition) also looked for less time.

To follow up on this hint of sensitivity, infants were split into groups of high and low attenders, and in fact, the high attending (but not low attending) bimodal infants successfully discriminated the category endpoints. As expected, neither the high nor the low attending infants in the flat condition revealed discrimination. These results suggest that infants' level of attention during learning plays a role in perceptual change. Interestingly, there was no overall correlation between looking time and discrimination, indicating that it is not simply that greater attending leads to more learning, but rather that there may be a threshold level of attended information necessary to affect perception.

The comparison between Experiment 1 (where infants did not modify perception following the distributional information) and Experiment 2 (where infants who paid attention did change perception) reveals that phonetic change depends at least in part on the specific contrasts in question (see also Burnham, 1986). One possibility is that some contrasts are inherently easier to learn. Supportive evidence is newborn infants' initial perceptual organization (Eimas et al., 1971) and the fact that the same contrasts discriminated by infants are also discriminated by non-human animals (e.g., Kuhl & Padden, 1982, 1983). This hypothesis of inherently easier contrasts could be tested by comparing the current results with those of infants from different language environments. If some contrasts are easier, this would predict no role or a lesser role of ambient language environment in category change. On this account, experience with a hypothetical language environment with a reversed profile of speech sounds, where the more frequent sounds are less frequent in English and vice versa, would lead to the same pattern of phonetic learning (easier learning in Experiment 2). This is unlikely in this case, however, as the boundary at which infants failed in Experiment 1 is in precisely that location where naïve, 6-8-month-old English infants perform best (Werker & Lalonde, 1988).

An experiential hypothesis would account for the difference in learning between the Experiment 1 and Experiment 2 contrasts by appeal to the different relationships of the respective sounds to the native English category. That is, the greater difference of Experiment 2 sounds from the native category centroid would be the factor allowing easier phonetic change. This account would predict that the hypothetical group of infants with more experience on the opposite end of the speech sound continuum would experience the inverse difficulty to the English infants, with success more easily revealed on the contrast from Experiment 1. Two specific possibilities underlie this experiential account. The less-frequently heard sounds may be easier to learn because of the greater perceptual distance between the already-formed native categories and the new sounds. Theories of speech perception suggest that nonnative perception is eased when the new sounds have a differential degree of goodness to the native category (PAM; Best, 1995); when the sounds are more different from the native category (SLM; Flege, 1995); and when the sounds are perceptually farther from the category prototype (Native Language Magnet; Kuhl, 1993). These theories assert that plasticity depends on the specific sounds being learned and their relationship to the relevant L1 categories; specifically the farther the sounds are from the native category centroid, the better the learning. Another experiential explanation is a distributional frequency hypothesis, which would account for the greater ease in phonetic learning in Experiment 2 by appeal to the fact that infants have less experience with the sounds in that experiment. Because the sounds are heard less frequently, there is less preexisting frequency information to be overcome, leading to easier change. The distributional hypothesis would predict that induction of discrimination of the more-frequently heard sounds in Experiment 1 would be possible, if infants are provided an increase in the amount of exposure roughly corresponding to the overall prevalence of the sounds in the native language. However, PAM and SLM would argue that perceptual plasticity should be independent of length of exposure.

Even in this experiment, it was not the entire group of infants who succeeded, but rather, only those who paid the most attention. This suggests that attention may be necessary in phonetic change by this age, and that may be a criterial component of social interaction. An alternative to the attention explanation is that some infants are better or faster learners, and these infants also tend to pay more attention. This explanation predicts that infants in the high attending group would have more advanced vocabularies at 15 months of age, but this was not in fact the case.

The next study presses forward with the attentional hypothesis and attempts to manipulate attention by increasing the length of the familiarization phase, allowing all infants to accumulate the threshold amount of attended exposure suggested by the second experiment. This third experiment is identical to Experiment 2, except that the familiarization period was doubled with the goal of allowing infants to garner sufficient amounts of attended exposure.

EXPERIMENT 3

Experiment 2 familiarized infants to a slightly different continuum of speech sounds from that of Experiment 1. The results revealed that infants were able to change speech sound category discrimination in accordance with briefly-presented distributional information, if they were paying sufficient amounts of attention. The high attending and low attending groups were determined by a post-hoc median split based on attending during the familiarization period. The goal of Experiment 3 was to follow up on the attentional results of Experiment 2 by ascertaining whether all infants might succeed if allowed to attain the identified threshold level of attention. This was operationalized by increasing the familiarization period to increase the likelihood that all infants would attend for a sufficient amount of time.

In Experiment 2 there was no correlation between overall amount of attention and discrimination, indicating that the relationship between the two is thresholded. The threshold suggested by the results was about 1 minute, which was attained by half of the infants. The current experiment doubled the familiarization period to allow all infants to reach 1 minute of attended familiarization.

Method

Participants.

Forty-eight healthy, monolingual English-learning infants participated in this experiment. Twenty-four were in the bimodal condition (12 females; *M* age=10 months 14 days, range=10 months 0 days – 11 months 2 days) and 24 were in the flat condition (12 females; *M* age=10 months 13 days, range=10 months 0 days – 11 months 1 day).

An additional 23 infants were tested and excluded for the following reasons: fussiness (19), parental interference (2), and ear infection (1).

Procedure.

The procedure was identical to that of the first experiment, except that there were two familiarization phases instead of one, and there was no MCDI follow-up.

Stimuli.

The first familiarization phase was identical to that of Experiment 2. The second familiarization used a static image of tulips (Maye et al., 2002) as a change from the moving digital flower image used in the first familiarization phase. All other parameters were identical to Experiment 2.

Results

Familiarization.

Infants accumulated more looking time than in Experiment 2, with the bimodal infants averaging 117.8 s (*SD*=32 s) and the flat infants averaging 113.1 s (27.8) (see Table 3). Forty-six of the 48 infants attended for longer than 1 minute (1 infant in each condition failed to attend for longer than a minute). There was no difference between the two conditions [$t(46)=0.55$, *ns*; Cohen's $d=0.16$].

Infants were separated into groups of high and low attenders (n=12 each) based on whether they looked for more or less than the median of their condition. Of the 3 minute 44 second familiarization, the median looking times were 113.9 s (bimodal) and 108.7 s (flat). Within each condition, the high attending group looked for longer than the low attending group [bimodal group: $t(22)=-5.29$, $p<0.001$; Cohen's $d=-1.84$; flat group: $t(22)=-4.90$, $p<0.001$; Cohen's $d=-2.02$].

Test trials.

A 2 (condition: bimodal versus flat) x 2 (test trial type: alternating versus nonalternating) between-within ANOVA was calculated. There was no main effect of condition [$F(1, 46)=0.48$, *ns*; $\eta_p^2=0.010$] or test trial [$F(1, 46)=0.21$, *ns;* $\eta_p^2=0.034$], but the interaction was significant [$F(1, 46)=4.69$, $p=0.03$; $\eta_p^2=0.092$]. Follow-up *t*-tests reveal that the bimodal group [$t(23)=-3.21$, $p=0.004$; Cohen's $d=-0.24$] but not the flat group [$t(23)=0.53$, *ns*; Cohen's $d=0.086$] distinguished between the two types of test trials.

Paired-samples *t*-tests were performed on the high- and low attending groups, and revealed that the both the high attending and low attending bimodal groups discriminated the sounds [high attending: $t(11)=1.82$, $p=0.048$[5]; Cohen's $d=0.12$; low attending: $t(11)=2.90$, $p=0.008$; Cohen's $d=0.35$] whereas the high attending and low attending flat groups did not [high attending $t(11)=-.77$, *ns*; Cohen's $d=-0.21$; low attending: $t(11)=0.053$, *ns*; Cohen's $d=0.013$]. A "nonalternating bias" score was created by subtracting looking time to alternating trials from looking time to nonalternating trials, and its correlation with familiarization looking time was not significant in either the bimodal [$r=-0.282$, $p=0.18$] or flat conditions [$r=0.056$, *ns*].

---

[5] Looking times of the bimodal high and low attending groups were analyzed using 1-tailed *t*-tests due to the directional a priori hypothesis based on the results of Experiment 2; all others were 2-tailed.

Table 3. Experiment 3 Results: Looking Time (s)

	Familiarization	Test Trials	
		Alternating Trials	Nonalternating Trials
All Infants	mean (SD)	mean (SD)	mean (SD)
Bimodal Condition	117.83 (32.01)	4.68 (1.96)	5.15 (1.96)
Flat Condition	113.11 (27.75)	5.30 (1.47)	5.18 (1.32)
By Attention Level			
Bimodal (high)	141.28 (23.04)	4.40 (1.73)	4.61 (1.66)
Bimodal (low)	94.38 (20.31)	5.08 (2.36)	5.90 (2.29)
Flat (high)	132.74 (22.77)	5.63 (1.24)	5.36 (1.28)
Flat (low)	93.49 (15.87)	4.98 (1.59)	5.00 (1.42)

Discussion

In this final experiment, the familiarization period was twice as long as in Experiment 2. This increased the likelihood that all infants would surpass the 1 min threshold of attended exposure suggested by Experiment 2. In fact, all but two of the 48 the infants (1 in each condition) exceeded the threshold. Following familiarization, infants in the bimodal condition (as a group), but not those in the flat condition, discriminated the sounds. The results of this experiment and Experiment 2 together reveal that it is possible to change infants' perceptual discrimination at 10 months of age, after the emergence of native phonetic categories. Attention may play a key role in infants' ability to change phonetic perception after perceptual reorganization.

GENERAL DISCUSSION

Summary of Experiments

This thesis explored perceptual plasticity in phonetic development. At 6-8 months, infants modify their phonetic discrimination based on the distributional characteristics of a briefly-presented stream of tokens (Maye et al., 2002; 2008). However, perception of speech sound categories develops from language-general to language-specific by 10 months of age, and may begin to become more stable by this age. The age-related change was expected to reduce plasticity in perceptual learning at 10 months as compared to younger ages, potentially rendering distributional learning less effective in modifying discrimination. Change may be most effective in the presence of social interaction (Kuhl et al., 2003), which could be necessary by 10 months of age in light of the decreased perceptual plasticity. Attention was hypothesized to be an important component of learning, potentially a factor of social interaction mediating phonetic change.

Three experiments tested whether 10-month-old infants' perception, after perceptual reorganization has occurred, could be affected by distributional frequency information. The sounds were presented in a non-social paradigm. Attention during this learning phase was monitored to test whether infants who paid more attention would reveal stronger evidence of learning.

The stimuli used in Experiment 1 were drawn from a continuum of speech sounds that are closer to the English /d/ category centroid and are heard relatively frequently in English (with which infants would have more experience). The results revealed that 10-month-old infants' perception was not affected by the brief exposure to the nonnative sounds. The second experiment replicated the first but familiarized infants to a continuum of speech sounds that was slightly different from those of Experiment 1, using sounds that are farther from the English category centroid and are heard less frequently in English. The results of Experiment 2 revealed that perception of these sounds was more easily changeable. Although there was no evidence that as a group infants modified their perception following the distributional learning manipulation, infants who paid higher than median levels of attention during the learning phase did change their perception.

To follow up on the attention results of Experiment 2, and to corroborate phonetic learning in a non-social paradigm, the third and final experiment increased the familiarization length to make all infants "high attending". The results confirmed that with sufficient amounts of attention, infants can change perception after phonetic reorganization.

Perceptual Plasticity

The current study exposed young infants from monolingual English environments to Hindi sounds. Insofar as exposure to a new language at this age would constitute learning a true L2 (as opposed to another native language; i.e., bilingual first language acquisition), the literature on L2 speech perception is relevant to these experiments. Undeniably, there is a period of development during which perception is more plastic. The term *critical periods* (e.g., Lenneberg, 1967) implies that the intervals are marked by rigid onsets and offsets, outside of which learning does not result in native-like abilities, whereas the terms *sensitive periods* (e.g., Bateson, 1979) or *optimal periods* (Werker & Tees, 2005), recognizes that such onsets and offsets may be more flexible. There is clear evidence for the existence of sensitive or optimal periods in language: For example, second languages learned earlier in life appear to be localized in brain regions more consistent with the native language than second languages learned later in life (Kim, Relkin, Lee, & Hirsch, 1997; see also Mayberry & Lock, 2003).

In phonetic perception, learning a language after infancy results in perceptual abilities that are not as fluent as native abilities (Sebastián-Gallés & Kroll, 2003). However, many factors impact the level of observed plasticity in nonnative perception. Adults can display sensitivity to nonnative contrasts in sensitive behavioural testing situations; for example, English speaking adults can discriminate difficult contrasts such as the Thompson /k'/ - /q'/ contrast and the Hindi /Ta/ - /ta/ distinction in an AX (same-different judgment) task when the inter-stimulus interval (ISI) between the stimulus pairs is short (500ms) but not when it is long (1500ms) (Werker & Logan, 1985). With the longer ISI, the auditory memory trace of the first phone appears to decay by the time the second phone is presented, leaving only a categorical representation of the first (see also Carney, Widin, & Viemester, 1977). Implicit measures also uncover recognition of nonnative sounds at some level. Event-related potential (ERP) studies that measure brain wave patterns from the scalp reveal neural evidence of remaining discrimination to nonnative contrasts (Dehaene-Lambertz, 1997; Rivera-Gaxiola, Csibra, Johnson, & Karmiloff-Smith, 2000), and perception of within-category phonetic variation is also observed in lexical tasks (McMurray, Tanenhaus, & Aslin, 2002). Sensitivity to L2 contrasts can also be trained, although with varying degrees of success (Logan, Lively, & Pisoni, 1991; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Pisoni, Aslin, Perey, & Hennessy, 1982; Tees & Werker, 1984). Thus, there are a number of factors (L1 and L2 category difference, AOL, situational context) that affect whether nonnative sounds are perceived. Even though there is strong evidence for an optimal period of phonetic learning, the native learning is not absolute, and new phonetic categories can be learned.

Age of learning (AOL).

The age at which a language is learned is clearly an important variable in second language learning (e.g., Newport, Bavelier, & Neville, 2001; Werker & Tees, 2005). The SLM (Flege, 1995) includes a role for AOL, acknowledging that with an older AOL it is more difficult to develop L2 categories. However, SLM does not specify a mechanism by which AOL could affect perception.

A comparison of the current experiments suggests one potential way that AOL could affect learning of different contrasts. A difference in learning was found between the different contrasts, while the age of the participants was held constant. This could be interpreted as evidence that the L1 category inventory holds the significant role, over and above AOL, in determining level of perceptual plasticity. However, although AOL traditionally refers to the age at which individuals begin learning (being exposed to) the language, another way to conceive of AOL is to specify it individually for each phonetic category. This would result in different AOLs for different contrasts (Anderson et al., 2003). For the infant participants in the present study, the English end of the stimuli continuum (Experiment 1) would have developed at a younger age than the Hindi end of the continuum (Experiment 2). Therefore, the AOL for the contrast in Experiment 1 would be younger than the AOL for the contrast in Experiment 2. In this way these experiments could also be interpreted as having implications for AOL.

A distributional version of an AOL account could also explain the greater ease of infants in Experiment 2. Sounds that are heard less frequently in the native language develop slightly later (Anderson et al., 2003), and therefore the less frequently-heard sounds in Experiment 2 would have been collapsed later than those in Experiment 1. A distributional AOL account could explain the Experiment 1 and Experiment 2 results by appeal to the shorter period of time from category formation in the latter stimuli, suggesting AOL as a factor that results in easier category change.

Attention

These experiments used the (non-contingent, non-social) paradigm of Maye et al. (2002) to expose the 10-month-olds to distributions of Hindi sounds. In the second experiment infants did not use this information to change perception, failing as a group to discriminate the sounds. However, those infants who paid higher-than-median levels of attention showed evidence of perceptual change. The third experiment followed up on this result by increasing the length of exposure to allow all infants to accumulate a sufficiently high amount of attended exposure. In this study, infants as a group modified their discrimination in accordance with the presented distributional information, suggesting that it was the amount of accumulated attention that mediated whether infants' discrimination was amenable to change.

The distributional learning of infants of 6-8 months who have not yet developed native categories does not appear to be mediated by attention (Maye et al., 2002; 2008). However, it is possible that the method used is not sensitive enough to reveal these effects in the younger infants. For example, there could be a critical threshold of attention necessary to affect perception at the younger age, but it could be significantly lower than that of the older infants, so that even with only the single familiarization phase, younger infants all have the opportunity to become sufficiently high attenders. Relatedly, younger infants generally display higher levels of attending than older infants, and although the data from Maye et al. (2002) are not available, it is possible that the majority of the 6-8-month-old infants surpassed the 1 minute threshold in the single familiarization, resulting in evidence of overall group learning.

Conversely, it is possible that the role of attention in the distributional learning paradigm is one that changes with age. This could be observed in at least two ways. First, it may be that attention is not required at the younger age. Given that infants' native perceptual categories have not yet developed, it is possible that phonetic change may be relatively easily affected, with or without attention. Once native categories have developed, change would be more difficult, and attention may be required to overcome the native perceptual sensitivities. Second, it is possible that attention is important at all ages, but what may change is the ease with which attention is attracted to the stimulus. In early infancy, the bias to attend to speech (Vouloumanos & Werker, 2004; 2007a; 2007b) and to the native language in particular (Mehler et al., 1988; Moon et al., 1993) may be sufficient to direct infants' attention towards language. However, by 10 months of age infants may require elements such as social interaction to draw attention to language, perhaps because infants have begun to learn that language is used for interpersonal interactions. The present experiments tested infants of 10 months of age only and cannot speak to whether attention may play a role at younger ages. To investigate this possibility, infants of 6-8 months would need to be tested.

Social Interaction

Experiments 2 and 3 demonstrate phonetic change after the development of phonetic categories at 10 months. These results contrast with those of Kuhl et al. (2003) where it appeared that phonetic learning did not take place in the absence of live social interaction, even at a slightly younger age. The current experiments argue that attention is the responsible factor. However, many other factors differed between the social condition of Kuhl et al. (2003) and the present non-social experiments that could account for the difference in findings.

Learning environment.

Infants in Kuhl et al. (2003) were exposed to language in the form of a speaker reading books and playing with toys, closely mimicking a natural learning environment. In the present distributional learning studies, infants viewed a television screen with a digital flower image and heard a disembodied, synthesized voice pronouncing isolated syllables. This is clearly not a natural learning environment, but is nonetheless an important experimental learning paradigm as it allows for control over most of the variables that impact learning. Perhaps social interaction is necessary in natural learning environments.

Contingency

Contingency is an important factor in infants' social interactions. Twelve-month-old infants will follow the gaze of a non-human entity if it interacts in a contingent manner (Johnson, Slaughter, & Carey, 1998). A lack of contingency in interactions can impact learning. Infants exposed to non-contingent interactions exhibit negative affect and expose themselves to lower levels of stimulation (Legerstee, 1997). It is possible that the lack of contingency in Kuhl et al.'s non-social conditions may have impaired infants' ability to learn the Mandarin sounds. In the non-interactive distributional learning paradigm, contingency may be less expected by infants, and would therefore be less likely to have any negative impact.

Phonetic salience.

In the present experiments, the contrast of interest was isolate, presented in consonant-vowel syllables rather than embedded in a variety of different lexical contexts. The target syllables comprised 67% of the presented stream. In Kuhl et al. (2003), the relevant information was embedded in different contexts (words), and distributed unevenly across the exposure. Further, the target sounds were estimated to comprise only about 6.5% of the language exposure. This may have resulted in a more difficult learning situation, which may require social interaction to effect phonetic learning. Even if this is the case, it is still an unanswered question whether social interaction would be necessary because it involves some variable (such as contingency) in addition to attention that facilitates phonetic learning, or whether social interaction would be necessary because it is especially effective in garnering attention.

Meaning.

Another difference between the methods is that in Kuhl et al. (2003) the phonetic sounds were distinguished with respect to both the distributional frequencies and attachment to different semantic meanings. Meaning affects phonetic learning in infants of 9-months: When different sounds are consistently attached to the same objects (i.e., sound-A is always paired with object-A; sound-B with object-B), infants discriminate the sounds. But when the pairing is inconsistent (sound-A is sometimes with object-A and sometimes with object-B), the infants fail to distinguish the sounds even though the same distributional information is present (Yeung & Werker, 2006; under review). There was no meaning linkage in the current experiments, with infants having to rely exclusively on distributional information. This learning mechanism is an additional source of information that was present in the Kuhl et al. (2003) experiments, and is one that was likely affected by the lack of attending by infants in the non-social conditions.

Future Directions

Contingency in social interaction.

A variant on the Kuhl et al. (2003) study that could help ascertain whether contingency is a facilitative aspect in phonetic learning would be to run the audio-visual condition of their study (interactive individual presented on a television screen). Rather than the interaction being pre-taped, the individual could maintain contingency by being broadcast from a remote room (from which they could observe the infants' actions, thereby being able to act and react accordingly). If contingency is important, greater learning may be observed than in the original Kuhl et al. (2003) audio-video condition. Of course, in this case, this experiment may reveal that the contingency leads to greater attending by the infants, in which case attention would still not be ruled out as the underlying mediating factor.

Attention.

Experiment 3 followed up on the attention result found in Experiment 2 by allowing all infants to become high attending, and tested whether that would result in better learning. A complementary way to test the attentional hypothesis that would be more consistent with the adult literature on attention would be to divert attention away from the phonetic information being presented, and test whether that would negatively affect learning. One way this could be tested is by drawing infants' attention away from the distributional speech stimuli by concurrently exposing infants to live interactions with a silent individual. These interactions could be expected to draw infants' attention away from the speech stream. The hypothesis is that with attention diverted, infants may not display evidence of phonetic learning.

Conclusion

Together, these data reveal that phonetic perception can be changed even after language-specific phonetic categories have developed at 10 months of age. Changing native language perceptual categories appears to be more difficult than changing the initial broad-based sensitivities in place at a younger age, but discrimination was still impacted following a brief exposure period. Even once language-specific categories have formed and phonetic development has become native-like, the perceptual system retains a significant degree of plasticity. Crucially, the change was induced in a non-social paradigm, indicating that a social context is not necessary to affect phonetic perception. These results instead pinpoint attention as a salient facilitating factor.

REFERENCES

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behavior, 46*, 149-159.

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*, 155-182.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L. & Perey, A. J. (1981). Discrimination of Voice Onset Time by Human Infants: New Findings and Implications for the Effects of Early Experience. *Child Development, 52,* 1135-1145.

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

Baker, S. A., Golinkoff, R. M., & Petitto, L. A. (2006). New insights into old puzzles from infants' categorical discrimination of soundless phonetic units. *Language Learning and Development, 2*, 147-162.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 11, pp. 101-159). New York: Academic Press.

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press/Bradford Books.

Baptista, L. F. & Petrinovich, L. (1984). Social interaction, sensitive phases and the song template hypothesis in the White-crowned Sparrow. *Animal Behavior*, *32*, 172-181.

Bateson, P. (1979). How do sensitive periods arise and what are they for? *Animal Behaviour, 27*, 470–486.

Berent, I., Marcus, G.F., Shimron, J., & Gafos, A.I. (2002). The scope of linguistic generalizations: Evidence from Hebrew word formation. *Cognition, 83*, 113–139.

Best, C., & Jones, C. (1998). Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behavior and Development, 4,* 295.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp. 171-206). Baltimore: York Press.

Best, C. T., McRoberts, G. W., LaFleur, R. & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, *18*, 339-350.

Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 345-360.

Birnholz, J. C. & Benacerraf, B. R. (1983). The development of human fetal hearing. *Science, 222,* 517-519.

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

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science*, *16*, 298-304.

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics, 7*, 207-239.

Burns, T.C. Yoshida, K.A., Hill, K., Werker, J.F. (2007). Bilingual and Monolingual infant phonetic development. *Applied Psycholinguistics, 28*, 455-474.

Carney, A. E., Widin, G. P. & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, *62*, 961-970.

Cech, C. G., & Shoben, E. J. (2001). Categorization processes in mental comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 800-816.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K. & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*, 351-353.

Cohen, L. B., Atkison, 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). Austin: University of Texas.

Colombo, J., Richman, W.A., Shaddy, D.J., Follmer Greenhoot, A., & Maikranz, J.M. (2001). Heart-rate defined phases of attention, look duration, and infant performance in the paired-comparison paradigm. *Child Development*, *72*, 1605–1616.

Conboy, B., Brooks, R., Taylor, R., Meltzoff, A., & Kuhl, P. (2008, March). Joint engagement with language tutors predicts brain and behavioral responses to second-language phonetic stimuli. *Poster presented at the XVIth International Conference on Infant Studies*, Vancouver, Canada.

DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C. & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development, 17*, 159-164.

DeCasper, A. J. & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*, 133-150.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8*, 919-924.

Dehaene-Lambertz, G. & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport*, *9*, 1885—1888.

Dehaene-Lambertz, G. & Houston, D. (1998). Language discrimination response latencies in two-month-old infants. *Language and Speech, 41*, 21-43.

Dehaene-Lambertz, G. & Peña, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport, 12*, 3155-3158.

Dooling, R. & Searcy, M. (1980). Early perceptual selectivity in the swamp sparrow. *Developmental Psychogiology, 13*, 499-5006.

Duffy, S., Huttenlocher, J., & Crawford, L.E. (2006). Children use categories to maximize accuracy in estimation. *Developmental Science, 9,* 598-604.

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behavior, 37*, 507-508.

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

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146*, 668-670.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99*, 9602-9605.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (5, Serial No. 242).

Fiser, J. & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99,* 15822-15826.

Flege, J. E. (1995). In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, (pp. 233-277). Baltimore, MD: York Press.

Flege, J. E., Takagi, N. & Mann, V. (1995). Japanese adults can learn to produce /r/ and /l/ accurately. *Language and Speech, 38*, 25-55.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics, 62,* 1668-1680.

Francis, A. L. & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 349-366.

Gomez, R. L., Bootzin, R. R. & Nadel, L. (2006). Naps promote abstraction in language learning infants. *Psychological Science, 17*, 670-674.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (in press, 2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431-436

Gómez, R. L. & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition, 70*, 109-135.

Gomez, R. L. & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7*, 183-206

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*, 544-549.

Guion, S.G. & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. Munro (Eds.) *Language Experience in Second Language Speech Learning*, (pp. 57-77). Amsterdam: John Benjamins.

Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior, 16,* 173-184.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356-388.

Hasher, L., & Zacks, R. T. (1984) Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372-1388.

Hasselmo, M. E., & Bower, J. M. (1993). Acetylcholine and memory. *Trends in Neuroscience*, *16*, 218–222.

Hayes, J. R., & Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the Development of Language*, (pp. 221–233), New York: Wiley.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23*, 65-94.

Hockett, C. (1960). The Origin of Speech. *Scientific American, 203*, 88-96.

Hunter, M. A. & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier and L. Lipson (Eds.), *Advances in Infancy Research* (Vol. 5, pp. 69-93). New Jersey: Ablex.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. In R.A. Hindle (Ed.), *Bird Vocalizations*, (pp. 61-74). London: Cambridge University Press.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1-19.

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

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1*, 233-238.

Jusczyk, P.W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics, 21*, 3–28.

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

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

Kim, K. H. S., Relkin, N. R., Lee, K., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*, 171–174.

Kirkham, N.Z., Slemmer, J.A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence of a domain general learning mechanism. *Cognition, 83*, B35-B42.

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America, 104,* 3568-3582.

Krentz, U. C. & Corina, D. P. (2008). Preference for language in early infancy: the human language bias is not speech specific. *Developmental Science, 11*, 1-9.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. deBoysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, (pp. 259-274). Dordrecht, Netherlands: Kluwer Academic Publishers.

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science, 8*, F1-F12.

Kuhl, P. K. & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*, 1138-1141.

Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America, 73*, 1003-1010.

Kuhl, P. K., & Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics, 32*, 542-550.

Kuhl, P. K., Stevens, E., Hayashi, A., Degushi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*, F13-F21.

Kuhl, P. K., Tsao, F. M. & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*, 9096-9101.

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

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology, 20*, 215-225.

Legerstee, M. (1997). Contingency effects of people and objects on subsequent cognitive functioning in three-month-old infants. *Social Development, 6*, 307-321.

Lenneberg, E. H. (1967). *The Biological Foundations of Language*. New York: Wiley.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358-368.

Liu, H. –M., Kuhl, P. K., & Tsao, F. –M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*, F1-F10.

Logan, J. S., Lively, S. E. & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, *89*, 874-886.

Marcus, G. F., Fernandes, K., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*, 387-391.

Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule-learning in seven-month-old infants. *Science, 283*, 77-80.

Marler, P. (1990). Innate learning preferences: Signals for communication. *Developmental Psychobiology, 23,* 557-568.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*, 241-265.

Mayberry, R. I. & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language*, *87*, 369-383

Maye, J., & Gerken, L. A. (2000). Learning phoneme categories without minimal pairs. *Proceedings of the 24th Annual Boston University Conference on Language Development*.

Maye, J., & Gerken, L. A. (2001). Learning phonemes: How far can the input take us? *Proceedings of the 25th Annual Boston University Conference on Language Development*.

Maye, J., Weiss, D. J., and Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*, 122-134.

Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101-B111.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the r-l contrast to Japanese adults: predictions of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience, 2*, 89-108.

McMurray, B., Tanenhaus, M.K., & Aslin, R.N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86,* B33-B42.

Mednick, S. C., Nakayama, K., Cantero, J. L., Atienza, M., Levin, A. A., Pathak, N. & Stickgold, R. (2002). The restorative effect of naps on perceptual deterioration. *Nature Neuroscience, 5*, 677-681.

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

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development, 16*, 495-500.

Moyer, R. S., & Dumais, S. T. (1978). Mental comparison. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*, (Vol. 12, pp. 117-155). New York:

Academic Press.

Narayan, C. (2006). Acoustic-perceptual salience and developmental speech perception. *Unpublished doctoral dissertation*.

Narayan, C., Werker, J. F., & Beddor, P. (under review). *The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination*.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology : Human Perception and Performance, 24*, 756-766.

Nazzi, T., Jusczyk, P.W., & Johnson, E.K. (2000). Language discrimination by English learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43*, 1-19.

Nelson, D. A. & Marler, P. (1994). Selection-based learning in bird song development. *Proceedings of the National Academy of Sciences, 91*, 10498-10501.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *ii*, 201–227.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127–162.

Newport, E.L., Bavelier, D., & Neville, H.J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Dupoux (Ed.), *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge, MA: MIT Press.

Nordby, J. C., Campbell, S. E., & Beecher, M. D. (2001). Late song learning in song sparrows. *Animal Behaviour, 61*, 835-846.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115, 39-57.

Panneton, R., & Aslin, R. N. (in preparation). *Developmental changes in infants' attention to static-visual and dynamic-visual events*.

Patterson, M., & Werker, J.F. (1999). Matching phonetic information in face and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, *22*, 237-247.

Pegg, J. E. & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America, 102*, 3742 – 3753.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604-607.

Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F. (2007). "Listen my children and you shall hear": Auditory preferences in toddlers with Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research, 50*, 1350-1364.

Pisoni, D. B., Aslin, R. N., Perey, A. J. & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 297-*314*.

Polka, L., Colantonio, C. & Sundara, M. (2001). A cross-language comparison of /d/-/D/ perception: Evidence for a new developmental pattern. *Journal of the*

Acoustical Society of America, *109*, 2190-2201.

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

Pons, F. (2006). The effects of distributional learning on rats' sensitivity to phonetic information. *Journal of Experimental Psychology, 32,* 97-101.

Reynolds, G.D., & Richards, J.E. (2005). Familiarization, attention, and recognition memory in infancy: an ERP and cortical source localization study. *Developmental Psychology, 41*, 598–615.

Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in adults. *Behavioural Brain Research, 111*, 11-23.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and nonnative speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*, 162-172.

Rose, K. C. & Rowe, E. J. (1976). Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 142-152.

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics, 59*, 347-357.

Rosenthal, O., Fusi, S., & Hochstein, S. (2001). Forming classes by stimulus frequency: Behavior and theory. *Proceedings of the National Academy of Sciences, 98*, 4265-4270.

Rowe, E. J. (1974). Depth of processing in a frequency judgment task. *Journal of Verbal Learning and Verbal Behavior, 13,* 638-643.

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

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606-621.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*, 101-105.

Saffran, J. R., Pollak, S. D., Seibel, R.L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*, 669-680.

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

Sanders, R. E., Gonzalez, E. G., Murphy, M. D., Liddle, C. L., & Vitina, J. R. (1987). Frequency of occurrence and the criteria for automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 241-250.

Sebastián-Gallés, N. & Kroll, J. (2003). Phonology in bilingual language processing: Acquisition, perception, and production. In A. Meyer & N. Schiller (Eds.), Phonetics and phonology in language comprehension and production: Differences and similarities (pp. 279–318). Berlin: Mouton de Gruyter.

Sekiyama, K. & Tohkura, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics*, *21*, 427-444.

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. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, *12*, 70-75.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*, B11-B21.

Southgate, V., Csibra, G., Kaufman, J., & Johnson, M. H. (2008). Distinct processing of objects and faces in the infant brain. *Journal of Cognitive Neuroscience, 20*, 741-749.

Streeter, L. A. (1976). Language perception of two-month old infants shows effects of both innate mechanisms and experience. *Nature, 250*, 39-41.

Takagi, N. (1993). Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view. Unpublished Ph.D dissertation*, University of California – Irvine*.

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate nonnative speech sounds. *Canadian Journal of Psychology*, *38*, 579-590.

Teinonen, T., Aslin, R. N., Asku, P., Csibra, G. (in press). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*, 49–67.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Statistical and stress cues in infant word segmentation. *Developmental Psychology, 39*, 706-716.

Thorpe, W. H. (1958). The learning of song patterns by birds with especial reference to the song of the chaffinch, *Fringilla coelebs*. *Ibis, 100*, 535-570.

Toro, J.M., Nespor, M., Mehler, J., & Bonatti, L. (2008). Finding words and rules in a speech stream: functional differences between vowels and consonants. *Psychological Science, 19,* 137-144.

Toro, J.M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*, B25-B34.

Toro, J.M. & Trobalón, J.B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics, 67*, 867-875.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47,* 466-472.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. T. (1996). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: Language-specific developmental changes in speech perception abilities. *The Emergence of Human Cognition and Language, 3*, 57-61.

Vallabha, G. K., McClelland, J.L., Pons, F., Werker, J.F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences, 104*, 13273-13278.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition, 107,* 729-742.

Vouloumanos, A. & Werker, J. F. (under review). Infants' learning of novel words in a stochastic environment.

Vouloumanos, A. & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science, 7*, 270 - 276.

Vouloumanos, A. & Werker, J. F. (2007b). Why voice melody alone cannot explain neonates' preference for speech. *Developmental Science, 10,* 170-172.

Vouloumanos, A. & Werker, J. F. (2007a). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science, 10,* 159-164.

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

Werker, J. F., Gilbert, J. H. V., Humphrey, K. & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349-353.

Werker, J. F. & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*, 672-683.

Werker, J. F., Pan, V. Y., & Baker, S. (in preparation). *Distributional learning of handshape discrimination*.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*, 147-62.

Werker, J.F., & Tees, R.C. (1983). Developmental changes across childhood in the perception of nonnative speech sounds. *Canadian Journal of Psychology*, *37*, 278-286.

Werker, J. F. & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and*

Development, 7, 49-63.

Werker, J.F., & Tees, R.C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866-1878.

Werker, J.F. & Tees, R.C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*, 233-251.

Werker, J.F. & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-35.

Yeung, H. H. & Werker, J. F. (under review). Learning words' sounds before learning how words sound: 9-month-old infants use distinct objects as cues to categorize speech information.

Yeung, H. H. & Werker, J. F. (2006, April). Conceptual and statistical learning in infant speech perception: objects cueing distinct words. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Boston, USA.

APPENDIX

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL- MINIMAL RISK RENEWAL

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC BREB NUMBER:
Janet F. Werker		H95-80023

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Point Grey Site

Other locations where the research will be conducted:
N/A

CO-INVESTIGATOR(S):
Laurel Fais
Whitney M. Weikum
Kristen McFee
Katherine Yoshida
Krista Newbigging
Henny Yeung
Krista Byers-Heinlein
Stephanie Helm
Vivian Yi Pan
Stefan Wittman
Tania Zamuner
Naz Akmal
Ladan G. G. Hamadani

SPONSORING AGENCIES:
Natural Science Engineering Research Council

PROJECT TITLE:
Linking Speech Perception to Language Acquisition: Biases, Mechanisms and Products

EXPIRY DATE OF THIS APPROVAL: March 15, 2008

APPROVAL DATE: March 15, 2007

The Annual Renewal for Study have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. Peter Suedfeld, Chair
Dr. Jim Rupert, Associate Chair
Dr. Arminee Kazanjian, Associate Chair
Dr. M. Judith Lynam, Associate Chair