DEVELOPMENT OF VOICE DISCRIMINATION

THE DEVELOPMENT OF VOICE DISCRIMINATION DURING INFANCY

By RAYNA H. FRIENDLY, B.Sc. (HONOURS)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2013) McMaster University
(Psychology) Hamilton, Ontario

TITLE: The Development of Voice Discrimination During Infancy

AUTHOR: Rayna H. Friendly, B.Sc. (Honours, McMaster University)

SUPERVISOR: Professor Laurel J. Trainor

NUMBER OF PAGES: xii, 185

Abstract

Infants must learn to discriminate between individuals in order to determine who is familiar (and likely to provide their basic needs) from those who are unfamiliar (and possibly a threat to their survival). One cue that humans use to discriminate others is the unique sound of each individual's voice. Until the present thesis, little research existed on the topic of voice discrimination development during infancy. I conducted the first set of studies to investigate whether voice discrimination develops through a process of *perceptual narrowing*. Perceptual narrowing is defined as an experience-dependent increase in sensitivity to distinctions important in the native environment and a decrease in sensitivity to distinctions not important (often foreign to) the native environment across the first year. It has been described in previous research for the processing of a number of socially-relevant stimuli in the auditory (e.g., musical rhythms and pitches, linguistic phonemes) and visual (e.g., faces) domains. In Chapter 2, I provide the first evidence that narrowing occurs for voice discrimination, with infants specializing for the discrimination of native (human, English-speaking) over foreign (rhesus monkey) vocalizations between 6 and 12 months. In Chapter 3, I establish that the specialization demonstrated in Chapter 2 resulted primarily from familiarity with the vocalization from the human species, rather than the particular language spoken. In Chapter 4, I show that sensitivity to distinctions between monkey voices can be reinstated at 12 months of age, after narrowing has taken place, with two weeks of exposure to monkey voices. Together, these findings indicate that infants become attuned to individual distinctions between human voices by the end of their first year, but that plasticity remains such that sensitivity to distinctions between voices from rarely-heard species can be reinstated with exposure, at least until the end of the first year.

Acknowledgements

Mom, my first thank you belongs to you because you have always helped to support me during my academic career, both emotionally and financially, and frequently reminded me that I was intelligent, capable and loved. Buby, you have also always been there provide support and I cannot tell you how much I appreciate that. Thank you and I love you both very much!

Next, Laurel, I want to thank you for your guidance and commitment to my success throughout my degree. I am not sure if you realize how in awe I (as well as many others) am of your hard work and dedication, as well as your general cleverness. You have been an invaluable role model to me, always inspiring me to work harder and not to settle for anything less than perfection. Thank you for teaching me through example, for helping me see various situations in a different light, and for having a sense of humour when I make embarrassing mistakes, like inviting an entire journal group to your house for our lab holiday party!

To my committee, thank you for all of your help preparing me for the completion of my degree. Daphne Maurer, thank you for always taking the time to give me thoughtful and valuable comments on my papers and presentations. I particularly appreciate your efforts to always include some positive feedback. David Feinberg and Steven Brown, thank you for de-mystifying the world of

vocal anatomy and acoustics for me. Without you, I would still shudder every time I saw the word “formant”.

To the second authors on my papers, Drew Rendall and Hongting Li, thank you for your important contributions to my research. Drew, thank you for all of the monkey vocalizations, spectrograms, and your thoughtful (and always-speedy) input. You have no idea how well I can now imitate a rhesus coo call! Hongting, thank you for bringing my project all the way to China. I could not have tested a non-English-speaking population without your help.

I feel so lucky to have such a great lab, and so many people who have helped me along the way to complete my thesis. Katie, you were a mentor and confidant, which is something I always try to embody when interacting with others in the department. Elaine, Andrea and Lisa, you are my favourite research assistants in the lab! Special thanks to Kate, Laura, and Chris who were extra helpful these past couple of years. Lastly, thanks to all the undergraduate students who contributed to my thesis projects: Lauren, Ren Ee and Aliyah. This lab has been my family for the last 6 years, I loved all of our chats, laughs and outings, and I am going to miss you all so much!

Finally, a huge thank you to all of my not-yet-mentioned friends and family. Maxie and Sunee, you both have such kind hearts and a high regard for the people in your lives, and I am very proud to call you sisters. Special thanks to Howard and Marlyne who were always there to offer advice and guidance when it came to planning my academic future, as well as a place to stay when I was in town. Mike, ever since we met...at the coffee shop...those years ago, I have not been happier! You are my comforter, my role model and my best friend. And to all others who have been there to support me along the way, (you know who you are!), I could not be more appreciative.

Table of Contents

List of Figures

CHAPTER 2

CHAPTER 3

CHAPTER 4

Preface

This thesis consists of one study that is under review for publication in a scientific journal (Chapter 2) and two studies that are prepared for submission for publication in scientific journals (Chapters 3 & 4).

The author of the present thesis is the primary author of all three manuscripts and was chiefly responsible for the experimental design, stimulus generation, data collection and analysis, and manuscript preparation. Drew Rendall (University of Lethbridge) was second author on Chapters 2 and 4, and provided all rhesus monkey vocalizations as well as all spectrograms. Hongting Li, (Zhejiang Sci-Tech University) was second author on Chapter 3, and collected data from participants in China. The thesis supervisor is the third author on all three manuscripts.

Chapters 3 and 4 present extensions to the empirical findings of Chapter 2 and thus overlap in stimuli and/or data presented.

CHAPTER 1: Introduction

It is important for infants to be able to discriminate among (i.e., tell apart or identify) individuals in order to determine for example, which individuals are familiar (and likely to provide items essential to their survival, such as food, shelter and protection), and which are unfamiliar and potentially dangerous or unresponsive to their needs. Such evaluations of identity enable infants to decide whom to approach and from whom to withdrawal, both in an emotional (e.g., crying versus smiling) or physical (movement towards versus away from) sense. It is therefore important that infants learn about the characteristics that uniquely identify individual people. One cue that is used by humans, as well as some non-human primates such as rhesus monkeys (Rendall, Rodman & Emond, 1996) and baboons (Rendall, Notman & Owren, 2009), to differentiate and identify group members is the unique sound of an individual's voice. This is a particularly important cue when discriminating between individuals who cannot be seen by the listener (e.g., are behind the listener or are on the telephone). Indeed, the human voice contains many perceptual cues to a speaker's identity, such as their age, size, and gender (Smith & Patterson, 2005). Many of these cues arise from formants, or vocal tract resonances, which reflect differences across individuals in vocal tract length and shape (e.g., Bachorowskia & Owren, 1999; Fitch, 2000). However, the range of such cues differs for human and non-human voices because of fundamental differences across species in vocal tract length and shape (e.g., Fitch, 2000). If discriminating humans by voice is important for infants, one might expect that with increasing experience with human voices, infants become specialized for differentiating human voices.

In the present thesis, I examined the role of learning and experience in the development of the ability to individuate human compared to rhesus macaque (*Macaca mulatta*) vocalizations. I tested 6-month-old, 12-month-old, and adult humans in order to address three main questions for which the answers were unknown previously. First, does the development of voice discrimination follow a pattern of *perceptual narrowing* (i.e., a process through which perceptual discrimination becomes better for discriminating stimuli in one's environment, and worse at discriminating stimuli from foreign environments) during the first year after birth? Second, if perceptual narrowing does indeed occur, what is the timeline of this narrowing? And third, can experience with voices from foreign species improve the ability to discriminate foreign-species voices once narrowing has occurred? My results provide evidence that perceptual narrowing does occur in infants' ability to discriminate voices, and that this narrowing occurs largely between 6 and 12 months of age. Specifically, during this time period, infants become significantly less sensitive to individuating characteristics of rhesus monkey voices and marginally more sensitive to individuating characteristics of human voices, such that at 12 months there is a robust superiority for discriminating human over rhesus monkey voices. Furthermore, I provide evidence that indicates that the voice-processing system remains plastic at 12 months such that two weeks of exposure to rhesus monkey voices can reinstate sensitivity to characteristics of individual monkey voices.

Perceptual Narrowing During the First Year

There is a great deal of evidence demonstrating that, during the first year after birth, infants' processing of important distinctions in socially-relevant stimuli, such as faces, language and music, follows an experientially-driven process of *perceptual narrowing* (for reviews see Lewkowicz & Ghazanfar, 2009; Scott, Pascalis & Nelson, 2007). Perceptual narrowing refers to a process by which, during the first months after birth, young infants are sensitive to important distinctions in both native and foreign stimuli, but become increasingly sensitive to native distinctions and decreasingly sensitive to foreign distinctions as they progress toward the end of their first year. For example, 6-month-old infants discriminate between two faces from a foreign- (monkey) species as easily as between two faces from their native- (human) species, but 9-month-old infants, like adults, are much better at human compared to monkey face discrimination (Pascalis, de Haan & Nelson, 2002). Trehub (1976) was the first researcher to show that young infants (1 to 3 months of age) show equal ability to discriminate foreign and native stimuli, in this case phonemes, and hypothesize that specialization for native (phonetic) stimuli in adulthood might result from a process of perceptual narrowing, where exposure to one's native language leads to a loss in ability to process distinctions in a foreign language. This idea was popularized in a classic study by Werker & Tees (1984). Over the last four decades, perceptual narrowing has been observed in the auditory and visual domains during the processing of a number of socially-relevant stimuli, including faces (Kelly et al., 2007; Pascalis et al., 2002; Simpson, Varga, Frick, & Fragaszy, 2010), intersensory (face/voice) stimuli (Lewkowicz & Ghazanfar, 2006; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009), linguistic phonemes, or speech-sound categories (Palmer, Fais, Michnick Golinkoff & Werker, 2012; Kuhl at al., 2006; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Polka & Werker, 1994; Tsao, Liu, Kuhl & Tseng, 2000; Werker & Tees, 2005; for reviews see Curtin & Werker, 2007; Kuhl, 2004, 2008), musical rhythms and scales (Hannon & Trainor, 2007; Hannon & Trehub, 2005a, 2005b; Lynch, Eilers, Oller, & Urbano, 1990; Trainor 2005; Trehub & Hannon, 2006), and, more recently, the processing of American sign language (Palmer et al., 2012) and dynamic human actions (Loucks & Sommerville, 2012). For this reason, perceptual narrowing has been proposed to be a domain-general mechanism of infant learning (Scott et al., 2007).

Despite the importance of identifying and discriminating individuals by voice, it is unknown whether the development of voice discrimination follows a pattern of perceptual narrowing similar to that found during the processing of other socially-relevant stimuli and the degree to which sensitivity to individual distinctions in foreign voice types remains plastic during the first year after birth. This issue is the focus of the present thesis. As background, here I describe evidence of perceptual narrowing during the discrimination of stimuli from the visual (e.g., faces) and auditory (e.g., language and music) domains. I then discuss some of the proposed mechanisms supporting perceptual narrowing during development. Next, I review what is known about infants' ability to process distinctions in voices. Finally, I summarize the literature pertaining to the effects of experience at different ages on the narrowing process.

a) Perceptual Narrowing During Face Perception

As with voices, the unique characteristics of faces (e.g., eye shape, featural configuration, skin texture and colour) can be used to differentiate individuals (Carroo, 1987; Haxby, Hoffman, & Gobbini, 2002; Kaminski, Dridi, Graff, & Gentaz, 2009; Nelson & Ludemann, 1989; Schultz, 2005). Numerous studies provide evidence that adults differentiate between faces of individuals from their own race more easily than between individuals from another race, a phenomenon termed the ”other race effect” (ORE) (for a review see Meissner & Brigham, 2001). For example, Caucasian adults demonstrate an ORE, in that they are more accurate at recognizing Caucasian, compared to African American or Japanese faces (Chance, Goldstein & McBride, 1975). African American adults also demonstrate an ORE, in that they are more accurate at recognizing African American than Caucasian or Japanese faces (Chance, Goldstein & McBride, 1975). Adults have also been shown to differentiate between faces of individuals from their own species more easily than between faces of individuals from another species. For example, Pascalis and Bachevalier (1998) found that human adults are superior at discriminating between human compared to rhesus monkey faces. However, this advantage for own-race and own-species face individuation is not present early in infancy. Infant studies using behavioural (e.g., Kelly et al., 2007, 2009; Pascalis et al., 2002; Simpson et al., 2010) and EEG (e.g., de Haan, Pascalis & Johnson, 2002; Scott, Shannon & Nelson, 2006; Scott & Monesson, 2010) methodologies reveal that 3-month-olds discriminate between foreign faces as easily as between native faces, but that an advantage for native face processing is developing noticeably at 6 months and is largely in place by 9 months of age. For example, Kelly et al. (2007) used visual paired-comparison (VPC) (a task that capitalizes on an infant's preference for novel stimuli by comparing the amount of time they spend looking at a novel object following familiarization with the object with which it is paired) to test 3-, 6-, and 9-month-old Caucasian infants' ability to discriminate faces. The faces were presented at various vantage points ranging from left to right and fell within other-race (African, Chinese, and Middle Eastern) or own-race (Caucasian) groups. Although the 3-month-olds discriminated the faces successfully in all race groups, the 6-month-olds provided evidence of discriminating only Chinese and Caucasian faces, and the 9-month-olds only Caucasian faces. A similar effect was found when Chinese infants were tested on discrimination of Chinese, African and Caucasian faces. Three-month-olds provided evidence of discriminating faces from all races, 6-month-olds only Chinese and Caucasian faces, and by 9 months of age, evidence of discrimination was limited to faces from their own race (Kelly et al., 2009). A comparable pattern of narrowing has been found for the processing of own- and other-species faces during infancy. Pascalis et al. (2002), also using the VPC task, tested infants and adults on their ability to distinguish between individual human or individual monkey faces that were in a forward-facing orientation. Six-month-olds discriminated faces within both species with equal ability, but 9-month-olds and adults demonstrated a clear advantage for human face discrimination. Face processing abilities during infancy have also been examined in an intersensory context and, interestingly, the ability to match silently-articulating dynamic faces to vocalizations also follows a pattern of perceptual narrowing. Specifically, at 6 months of age, intersensory (face/voice) matching is language- and species-general, but becomes specialized for one's native language and species by 8 to 11 months of age (Lewkowicz & Ghazanfar, 2006; Pons et al., 2009). Lewkowicz and Ghazanfar (2006) tested the ability of 4-, 6-, 8- and 10-month-old infants to match foreign-species (rhesus monkey) vocalizations to videos of monkeys making these calls. Infants were shown two silent videos side-by-side (one of a monkey producing a “coo” call and one of the same monkey producing a “grunt” call). They were then played either a “coo” or “grunt” vocalization and their looking time to each face was measured. At 4 and 6 months of age, but not at 8 and 10 months, infants looked longer at the monkey face that matched the vocalization being played. A similar result was found between 6 and 11 months of age for matching of foreign-language vocalizations to a silently-articulating dynamic human face (Pons et al., 2009). Pons and colleagues (2009) measured English- and Spanish-learning infants’, and well as Spanish adults’, ability to match the auditory English phonemes /ba/ and /va/ to a silent dynamic face articulating either the /ba/ or /va/ phoneme. Results showed evidence of narrowing in that the Spanish 6-month-olds were like the English infants at 6 and 11 months in that they could match the audible and visible English phonemes successfully. However, at 11 months and extending into adulthood, the Spanish participants no longer demonstrated intersensory matching for the foreign phonemes. In sum, perceptual narrowing for the processing of facial stimuli appears to begin between 3 and 6 months of age, and the specialization for native face-types is fully present by 9 to 11 months of age.

b) Perceptual Narrowing During Language Perception

Perceptual narrowing for the ability to discriminate individuals by voice has not yet been examined, but much research has investigated perceptual narrowing for the speech-sound (phonemic) categories used in different languages around the world. Different languages use somewhat different phonemic categories. For example, the English language contrasts /r/ and /l/ categories, while in Japanese these sounds form one category. On the other hand, English does not distinguish between the retroflex /T/ and dental /t/ that are used by Hindi and other South Asian languages (Werker & Tees, 2005). Adults demonstrate a well-developed advantage for native phoneme processing. They have difficulty discriminating between phonemes that fall into two categories in a foreign language but only one category in their native language. At the same time, they discriminate with ease phonemic contrasts that fall into two categories in their native language (Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Kuhl, 1998; Kuhl et al., 2006; Werker & Lalonde, 1988; Werker & Tees, 1984, 2005; Werker et al., 2007). For example, Werker, Gilbert, Humphrey, and Tees (1981) demonstrated that English-speaking adults have difficulty discriminating between the Hindi retroflex /T/ and dental /t/ contrast, as sounds in both of these categories fall into the /t/ category in the English language; yet the adults were successful in discriminating between the /b/ - /d/ contrast, which falls into two separate categories in the English language. A large body of evidence shows that the advantage for distinguishing native, compared to foreign, phonemic contrasts is developed through a process of perceptual narrowing (for reviews see Curtin & Werker, 2007; Kuhl, 2004, 2008). For consonant contrasts, an advantage for distinguishing native compared to foreign phonemic categories is absent at 6 months of age, but in place by 10 to 12 months of age (e.g., Best et al., 1995; Kuhl, 1998; Kuhl at al., 2006; Tsao et al., 2000; Werker & Lalonde, 1988; Werker & Tees, 1984, 2005; see Werker & Tees, 1999; Lewkowicz & Ghazanfar, 2009 for review). For example, Werker and Tees (1984) examined infants' ability to discriminate consonant sound categories using a Conditioned Head Turn (CHT) procedure. In this procedure, infants are taught to turn their head in response to a change from one category to another. The authors found that although 6- to 8-month-old English-learning infants could discriminate between the Hindi retroflex /Ta/ and dental /ta/, by 10 to 12 months of age infants no longer provided evidence of doing so. Similarly, declines in the ability to perceive distinctions such as those between the Hindi retroflex-dental /Da/-/da/ (Werker & Lalonde, 1988) and the Nthlakampx (an interior Salish language spoken by indigenous First Nations [i.e., aboriginal] people of Canada in the province of British Columbia) ejective categories /k'*æ*/ - /q'*æ*/ (Best et al., 1995) have been reported for English-learning infants, and a decline in the ability to perceive the English /ra/-/la/ distinction has been reported in Japanese-learning infants (Kuhl, 1998). On the other hand, the ability to discriminate vowel contrasts appears to narrow considerably earlier in development than for consonant contrasts, and is already evident by 6 months of age (Kuhl, 2004; Kuhl et al., 1992; Polka & Werker, 1994). For example, using the CHT procedure, Polka & Werker (1994) found that although 4-month-old English-learning infants can discriminate between vowel contrasts used in German (dYt/DuT) and English (dit/dat), 6- to 12-month olds provide no evidence of being able to distinguish the foreign contrast. Also, a "perceptual magnet effect" (treating exemplars within a category differently depending on how similar they are to the prototype of that category) is demonstrated at 6 months of age for native-, but not foreign-, language vowel categories (Kuhl et al., 1992). In sum, perceptual narrowing is well in progress by 6 months of age for vowels, but not clearly observed until 10 to 12 months of age for consonants.

c) Enculturation to Musical Rhythm and Pitch

Music is another human communication system for which perceptual narrowing has been examined. The structure of musical rhythm (patterns of beat durations) and meter (perceptually derived regular patterns of strong and weak beats) differs across cultures (see Geiser, Ziegler, Jancke & Meyer, 2009 for further definitions of rhythm and meter). North American traditional classical and popular music typically has a simple metrical structure. For example, a march rhythm has alternating accented and unaccented beats that produces a simple time ratio of 2:1. A waltz has every third beat accented to produce a ratio of 3:1. On the other hand, Bulgarian folk music often contains complex ratio durations, such that a repeating 5-beat grouping would be divided into smaller groups of 3 and 2 beats, resulting in ratios of 3:2 or 2:3 (Hannon & Trehub, 2005a). Research by Hannon and Trehub (2005a, 2005b) suggests that perceptual narrowing occurs during the development of sensitivity to musical rhythms, with infants becoming specialized for their native musical system by 12 months of age. For instance, Hannon & Trehub (2005a, 2005b) used a familiarize-preference procedure (where infants controlled how long they listened to structure - preserving verses structure-violating rhythms) to demonstrate that 6-month-old North American infants distinguish violations in both simple (American-typical) and complex (Bulgarian-typical) rhythms In both cases, the infants demonstrated a novelty preference, choosing to listen to the structure-violating versions of the rhythms for longer durations than the structure-preserving rhythms. However, 12-months-olds and adults from North America detected successfully only the violations in the simple, American-typical rhythms.

Similar to the rhythmic structures of music, musical scales also differ across cultures. Different musical systems divide the octave into different intervals (pitch distances between notes). These notes form scales from which music is composed. In Western traditional classical musical structure, scales contain a subset of all available musical notes (e.g., of all 12 available notes per octave in Western classical music: A, A sharp (#), B, C, C#, D, D#, E, F, F#, G, and G#, the C major scale contains only the notes C, D, E, F, G, A, and B; Aldwell, Schachter & Cadwaller, 2011). Specialization in the form of perceptual narrowing for native musical scale structure appears to be established somewhat later than for the processing of rhythmic structure (Lynch et al., 1990; Trainor & Trehub, 1992), although definitive studies have yet to be done. For example, Trainor and Trehub (1992) showed that 8-month-olds were equally good at detecting changes in melodies that violated Western music scale structure as changes that did not, but adults were much better at detecting changes that violated Western structure. In some conditions, infants actually performed better than adults. Similarly, Lynch, Eilers, Oller and Urbano (1990) found that musically untrained Western adults were better at detecting mistunings in sequences based on the native major scale compared to the foreign pelog scale, where Western 6-month-olds were equally good at both, although with very simple stimuli. Thus, some processing advantages for major over pelog scales have been reported at 6 months of age (Lynch & Eilers, 1992). However, because a crossover design that tests both infants exposed only to Western major scales and infants exposed only to pelog scales has not yet been done, it remains unclear as to whether this is due to learning or intrinsic differences in the difficulty of processing major and pelog scale structure regardless of familiarity. That learning is involved in acquiring sensitivity to musical tonality comes from a study showing that when participating in active musical training, 12-month-old infants from North America demonstrate knowledge of Western tonality whereas 12-month-old infants, also from North America but not receiving this training, do not (Gerry, Unrau, & Trainor, 2012; Trainor, Marie, Gerry, Whiskin, & Unrau, 2012). In any case, the advantage for native musical scale is present clearly for most children regardless of musical training by preschool age (Corrigall & Trainor, 2010; Trainor & Trehub, 1994, for reviews see Hannon & Trainor, 2007; Trainor & Corrigall, 2010; Trainor & Unrau, 2012). In sum, there is no evidence to date of perceptual narrowing for musical stimuli until after 6 months of age. However, substantially fewer studies have been done in the domain of music compared to the domains of language and face processing. While narrowing appears to occur between 8 and 12 months of age for the processing of musical rhythmic structure, narrowing for the processing of musical scale structure does not appear to occur until after 8 months of age. At 12 months, infants receiving music training do exhibit a bias for their native scale. However, without such training, this bias generally has not been documented until the preschool years. These findings may speak to the role of the amount and type of exposure in facilitating the perceptual narrowing process, as many infants are exposed to music less regularly than to language and faces in their native environment.

d) Mechanism of Perceptual Narrowing and Summary

Perceptual narrowing occurs across a number of important auditory and visual domains, leading to the hypothesis that it reflects a general developmental learning mechanism that enables the infant brain to capture distinctions that are critical in its environment. Three neural models have been proposed to explain perceptual narrowing (see Lewkowicz & Ghazanfar, 2009, Box 2). The first model proposes that perceptual narrowing is driven chiefly by a process of synaptic ‘pruning’, whereby an initial overproduction of neural connections is followed by selective elimination of connections that receive insufficient stimulation (Black, Jones, Nelson & Greenough, 1998; Changeux & Danchin, 1976; Cowan, Fawcett, O'Leary & Stanfield, 1984; Goldman-Rakic, 1987; Huttenlocher, 1979; Low & Chen, 2006; Scott et al., 2007; Webb, Monk, & Nelson, 2001).

The second model postulates that neural elaboration (i.e., creation of new neural connections), rather than attenuation, drives perceptual narrowing. Support for the neural elaboration model comes from animal studies showing that there is continued growth of the non-human rhesus monkey brain long after birth (Bourgeois & Rakic, 1993; Purves, White & Riddle, 1996), and that there is no evidence of an overproduction of neural connections during early development in certain areas of the ferret (e.g., visual cortex) and rodent (e.g., somatosensory cortex) brains (Agmon, Yang, O'Dowd & Jones, 1993; Quartz & Sejnowski, 1997). Note that these two models do not appear to be mutually exclusive. For a number of mammalian species (including humans) there is evidence suggesting that pruning follows elaboration during neural development (e.g., Bourgeois, 1997; Sowell et al., 2003). However, observed changes in grey matter (GM) and white matter (WM) volumes during development suggest that elaboration might play a larger role than pruning during the first year after birth, when perceptual narrowing generally occurs (e.g., Giedd et al., 1999; Gilmore et al. 2007, 2011; Knickmeyer et al., 2008; see Chapter 5 for further discussion on this point).

The third model suggests that perceptual narrowing could be driven by a shift in the brain regions that are activated during processing during the first year after birth (Vogel, Monesson & Scott, 2012). This idea is supported by a study in which event-related potentials (ERPs) were recorded during 5- and 9-month-olds' processing of emotional native- and foreign-race faces, as well as emotional, non-verbal voices (Vogel et al., 2012). Results provide evidence of a shift, for the most part, from brain areas related to attention to those related to perception between 5 and 9 months of age, a time frame during which infants also develop advantages for processing matches between emotional native-race faces and voices. Again, it should be noted that these models are not necessarily mutually exclusive, but this short review makes it clear that little is known about the brain processes underlying perceptual narrowing.

In summary, perceptual narrowing during the first year after birth has been documented for a number of socially-related stimuli within the auditory and visual domains, including faces, language and music. Generally, specialization for processing native distinctions has been found to occur between 6 to 12 months of age, though earlier narrowing has been documented for faces and vowels than for consonants or for musical pitch and rhythm. However, it is unknown whether the socially-relevant ability to discriminate individuals by their voices also narrows during the first year after birth. In fact, very little is known about the development of voice discrimination abilities during infancy, even though there have been a number of studies investigating voice processing abilities in general during this time period.

Voice Processing During Infancy

Virtually nothing is known about whether the ability to distinguish individuals by voice undergoes a process of perceptual narrowing during development. However, a number of studies have investigated infants' perception of voices, in general. Many of these studies examined the ability of infants to discriminate an unfamiliar voice from a familiar (parental) voice. For example, measures of changes in fetal heart rate in response to different voices demonstrated that fetuses close to term (between 37 to 42 months gestational age) of Chinese women can discriminate the voice of their mother from that of an unfamiliar female (Kisilevsky et al., 2003). Studies examining differences in sucking rates on a nonnutritive nipple have demonstrated that newborn infants can discriminate infant-directed speech produced by their mother from that produced by a female stranger, showing a preference for their mother's voice (DeCasper & Fifer, 1980; Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978; Spence & Freeman, 1996) and that newborns can discriminate (but do not show a preference for) their father's voice compared to the voice of an unfamiliar male (DeCasper & Prescott, 1984). Few studies have examined the identification or discrimination of unfamiliar voices in infancy.

One question concerns cues that are used to tell apart individuals by voice. Adult studies suggest that the voices of adult humans differ in a number of ways that lead to perceived differences in vocal timbre or voice quality, including vocal tract size and shape (e.g., Fant, 1973; Peterson & Barney, 1952), dialect (e.g., Byrd, 1992), articulatory habits unique to an individual such as intonational, rhythmic and intensity variations (e.g., Endres, Bambach & Flösser, 1971; Klatt, 1986), glottal characteristics dictated largely by the properties of the vocal folds (e.g., Klatt & Klatt, 1990), and voice-onset-time [VOT] (e.g., Allen, Miller & DeSteno, 2003). In contrast, much less is known about the cues that infants use to individuate human voices. There is evidence to support the idea that English-learning infants are sensitive to distinctions in VOT (at least for synthetic voices) at 6 to 11 months of age (Aslin, Pisoni, Hennessy & Perey, 1981), suggesting that they might be able to use VOT as a cue for individuating voices. However, much more research is needed to understand what cues infants might be using to individuate the voices of others.

A second question related to the perception of unfamiliar voices concerns how this ability develops during infancy. Lecanuet, Granier-Deferre, Jacquet, Capponi, and Ledru (1993) have demonstrated, by measuring changes in heart rate during the presentation of voices, that the fetus at 36 to 39 months gestational age can discriminate an unfamiliar male voice from an unfamiliar female voice. Newborn infants have also been found, using nonnutritive sucking, to *prefer* unfamiliar voices speaking their native language to those speaking foreign languages (Mehler et al., 1988; Moon, Panneton Cooper & Fifer, 1993). Recently, Johnson, Westrek, Nazzi, and Cutler (2011) conducted the first study to examine infants' *discrimination* of unfamiliar voices of the same sex. Using a visual fixation procedure, where infants' looking time to a novel voice is compared to that for a voice to which they were familiarized, they demonstrated that Dutch-learning 7-month-old infants exhibited an *own-language effect* when processing voices. In other words, these infants were able to discriminate between the voices of two unfamiliar women speaking sentences in Dutch, but they showed no evidence of being able to discriminate between unfamiliar women speaking sentences in foreign languages (Italian or Japanese). However, Johnson et al. did not examine infants younger than 7-months of age in order to determine when in infancy this native-language voice bias develops. Specifically, they did not investigate whether the specialization for native-language voice discrimination might develop through a process of perceptual narrowing. That is, they did not investigate whether discrimination is initially as good for two foreign voices as it is for two native voices, but becomes better for native than foreign voices by the end of the first year after birth. Furthermore, very little is known as to whether infants specialize solely for only own-language voices, or if they also specialize for the unique acoustic qualities of human voices in general, regardless of the language spoken. This question is the focus of the present theses.

In the following chapters, I describe the first set of studies to investigate whether perceptual narrowing occurs during the development of voice discrimination during infancy. Specifically, because my primary interest was the individuation of human voices in general, rather than of the voices of one's own language in particular, in Chapter 2, I compared human 6-month-olds', 12-month-olds' and adults' discrimination of native-species (human) voices to their discrimination of phylogenetically close foreign-species (rhesus monkey) voices. In other words, monkey voices were chosen as the foreign category because monkeys' vocal anatomy has some similar qualities to that of humans but monkey voices are never or rarely heard in our participants' native environment. Results revealed a developmental pattern similar to the perceptual narrowing observed in other domains. Between 6 and 12 months of age, I found evidence that infants' ability to discriminate between two rhesus monkey voices decreased significantly, while their ability to discriminate between two human, English-speaking voices increased marginally, leading to a robust superiority for discriminating human over monkey voices by 12 months of age.

In Chapter 3, I examined whether the findings from Chapter 2 might result in part to familiarity with the English language used in my human voice stimuli by testing the ability to discriminate among human and among rhesus monkey voices cross-culturally. In collaboration with a colleague in China, we tested how well a population of adults with limited or no English exposure would be able to tell apart the human (English) voices as well as the monkey voices. The results suggest that the effects were driven by familiarity with human species voices in general, although there may also be some effect of language familiarity on voice discrimination ability in adults.

In Chapter 4 of the present thesis, I examined the effects of exposure on the development of voice discrimination around 12 months of age, after perceptual narrowing has occurred. This study stemmed from a substantial body of research demonstrating that there is a great degree of plasticity at this age in other domains. The literature on exposure effects is reviewed below.

Effects of Exposure for Foreign Stimuli on Perceptual Narrowing

A number of studies have examined the effects of exposure on perceptual narrowing across several domains (e.g., Anzures et al., 2012; Burns, Werker & McVie, 2003; Hannon & Trehub, 2005b; Kuhl, Tsao & Liu, 2003; Pascalis et al., 2005; Scott & Monesson, 2009, 2010). These studies reveal that sensitivity to foreign distinctions is greatly influenced by the timing and type of exposure to foreign stimuli. Nearly 30 years ago, Brigham and Malpass (1985) proposed the *contact hypothesis*, the idea that experience might play a crucial role in the specialization observed for own-race faces in adults. More recently, research on infants has demonstrated that a brief period of exposure to foreign stimuli during the first year can *maintain*, or prevent the loss in, sensitivity to distinctions in these foreign stimuli that would otherwise be lost (e.g., Burns et al., 2003; Pascalis et al., 2005; Scott & Monesson, 2009). Exposure to foreign stimuli at the end of the first year can *reinstate* sensitivity previously seemingly lost (e.g., Anzures et al., 2012; Hannon & Trehub, 2005b; Kuhl et al., 2003). Pascalis et al. (2005) provide evidence that 9-month-old infants maintain the ability to discriminate novel monkey faces after receiving regular exposure to a set of monkey faces for a 3-month period (2 minutes every day for 2 weeks, then less frequently the following 2.5 months) when the infants were between 6 and 9 months of age. On the other hand, Kuhl et al. (2003) provide evidence of reinstatement, showing that, even though English-learning infants typically lose their sensitivity to Mandarin phonemic contrasts that are not found in the English language by 10 months of age, their sensitivity to these contrasts can be regained after a 4-week period (which included 12 sessions, each 25 minutes in duration) of interaction with Mandarin-speaking adults. Sensitivity to foreign musical rhythmic structure also has been successfully reinstated in 12-month-olds after 2 weeks of twice-daily exposure to a 10-minute CD (Hannon & Trehub, 2005b), and sensitivity to distinctions in foreign-race faces was regained in 8- to 10-month-olds after 3 weeks of daily 4-minute DVD exposure to other-race faces (Anzures et al., 2012).

It is important to note that there is evidence supporting the idea that the timing of exposure plays an important role in the plasticity of the ability to process foreign distinctions. A number of studies suggest that exposure to foreign stimuli in older childhood can increase sensitivity to distinctions in these foreign stimuli (e.g., Feinman & Entwisle, 1976; Macchi Cassia, Kuefner, Picozzi & Vescovo, 2009; Sangrigoli, Pallier, Argenti, Ventureyra & de Schonen, 2005; Wang & Kuhl, 2003) and there is similar evidence from adulthood (e.g., Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; de Heering & Rossion, 2008; Iverson, Hazan & Bannister, 2005; Kuefner, Macchi Cassia, Picozzi & Bricolo, 2008; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Pruitt, Jenkins & Strange, 2006; Scott, Tanaka, Sheinberg & Curran, 2006, 2008; Tees & Werker, 1984; Zhang et al., 2009). In adulthood, this exposure is particularly effective in people exposed previously to the stimuli during infancy or childhood (e.g., Lenneberg, 1967; Macchi Cassia et al., 2009; Newport, Bavelier & Neville, 2001; Oh, Au & Jun, 2010; Sangrigoli et al., 2005; Tees & Werker, 1984), especially if the training stimuli are numerous and highly variable (Zhang et al., 2009) and if stimulus distinctions are exaggerated (McCandliss et al., 2002). However, the effectiveness of exposure generally decreases with increasing age, such that adults rarely, if ever, demonstrate an ability to process foreign distinctions at a level similar to their native distinctions (e.g., Flege, Yeni-Komshian & Liu, 1999; Hannon & Trehub, 2005b; Iverson et al., 2005; McCandliss et al., 2002; Takagi, 2002; Takagi & Mann, 1995, for reviews see Birdsong, 2006; Hernandez & Li, 2007). However, exposure can be very effective at least up to 9 years of age under some circumstances. Sangrigoli et al. (2005) demonstrated native-like levels of Caucasian-face discrimination in Korean adults who were adopted by Caucasian (French) families between 3 to 9 years of age. On the other hand, Korean adults, who have no exposure to their first language following adoption by French families in childhood, have difficulty processing Korean in the same manner as French-speaking adults for which Korean is a foreign language (Ventureyra, Pallier & Yoo, 2004). Furthermore, theses adoptees show little advantage in the processing Korean phonemes following re-exposure during trips to Korea (Ventureyra et al., 2004).

In addition to the timing of exposure, the type of exposure that an individual receives can also have a large influence on the plasticity of perceptual narrowing. Exposure at an individual level is crucial in maintaining the ability to discriminate individual monkey faces at 9 months of age (Archambault, O'Donnell & Schyns, 1999; Scott & Monesson 2009, 2010; Scott et al., 2006, 2008; Waxman & Braun, 2005). For example, in the training regiment used by Pascalis et al. (2005), one group of parents showed infants pictures of monkey faces one at a time, verbally labelling each face with a unique name (e.g., Dario, Flora, Boris). Using the same procedure, Scott and Monesson (2009) found that only infants who received this individual-level training maintained sensitivity to the differences among monkey faces; infants who received category-level (i.e., each monkey was labelled "monkey") or exposure (i.e., monkeys were not labelled) training showed no evidence of being able to discriminate between monkey faces at 9 months of age. Scott and Monesson suggested that the individual-level training was successful in maintaining foreign sensitivity because it drew infants' attention to how the monkey faces differed, whereas category-level training drew their attention to what the faces had in common. Indeed, Anzures et al. (2012) demonstrated sensitivity to foreign-race faces after exposing 8- to 10-month-olds daily to videos of foreign-race women who introduced themselves by name. The authors claim that this was done in order to "maximize individuation between the faces" (pg. 486); however, they did not compare this to a condition where the females did not introduce themselves by name, so it is unclear if individual-level labelling had an effect on reinstatement in this case. To date, there are no similar studies, to my knowledge, in the auditory domain that compare the effects of individual- and category-level training on the ability of humans to process distinctions in speech or voice stimuli.

Social interaction has also been implicated as a potentially important factor for reinstating sensitivity to foreign distinctions in auditory, linguistic stimuli. For example, in the aforementioned study by Kuhl et al. (2003), sensitivity to Mandarin phonemes was reinstated only in a group of 10-month-olds who interacted with male and female Mandarin speakers. When recordings of the Mandarin speakers were viewed on a monitor (audio-visual condition) or heard over speakers (audio-alone condition), 10-month-olds did not learn to discriminate the Mandarin phonemes, preforming similarly to their untrained counterparts. In contrast, sensitivity to foreign-race faces was reinstated at 8 to 10 months from audio-visual video exposure (Anzures et al., 2012) and to foreign musical rhythms at 12-month-old infants simply from audio-alone exposure (Hannon & Trehub, 2005b).

One caveat in this line of research is a study demonstrating that sensitivity to foreign-species (monkey) faces actually can be reinstated without a relatively long period of at home training, but rather with a longer familiarization period at the time of testing (Fair, Flom, Jones & Martin, 2012). Fair et al. (2012)

discovered that, although infants at 12 month of age failed to show evidence of discrimination of unfamiliar monkey faces after a 20s familiarization period, the 12-month-olds did show evidence of discrimination after a 40s familiarization period. Although this extra familiarization time could be considered a relatively brief and non-repeated form of training, it certainly raises the question of whether this extra familiarization time would be sufficient for reinstatement in domains other than faces. This question requires further exploration.

Although past literature on exposure-effects reveals that sensitivity to foreign phonetic contrasts, musical rhythms, and foreign-race faces can be reinstated after perceptual narrowing has occurred, it remains unknown whether the ability to discriminate foreign *voices* could also be reinstated. Chapter 4 of the present thesis begins to address this question by examining whether discrimination of voices from a foreign species (rhesus monkey) can be re-established through exposure to monkey voices at around 12 months of age, after perceptual narrowing has occurred (as shown in Chapter 2). Because of the potential importance of individual-level encoding demonstrated for training with monkey faces (Scott & Monesson, 2009), and of the social-interaction context demonstrated for training with foreign phonemes (Kuhl et al., 2003), parents were instructed to interact with their infant during training with the monkey voices, and each individual monkey's voice was labelled with a unique name. Findings revealed plasticity after perceptual narrowing was established, as two weeks of regular monkey-voice exposure reinstated sensitivity to rhesus monkey voices.

Summary: Thesis Outline and Contributions

My thesis is composed of three studies that examine the development of native- and foreign-species voice discrimination abilities in human infants and adults. Previous literature has found an own-language voice-discrimination advantage for infants at 7 months of age, but the discrimination of native compared to foreign voices has not yet been tested in infants younger than 7-months of age. Furthermore, the discrimination of human compared to nonhuman voice types by human participants has not yet been tested.

In Chapter 2, I address this gap in the literature by investigating the ability of 6-month-olds, 12-month-olds and adults to discriminate native-species (human) speech and foreign-species (rhesus monkey) vocalizations. My objective was to determine whether perceptual narrowing occurs for voice discrimination, as demonstrated by the development of a processing advantage in native-voice, and a decrement in foreign-voice, discrimination. Foreign distinctions were made at the species, rather than linguistic, level in order to assess the ability to distinguish qualities in voices that are unique to humans in general, rather than to speakers of one's native language in particular. The results provide evidence that the discrimination of voices develops through a process of perceptual narrowing similar to that observed in other domains. Between 6 and 12 months of age, infants develop a robust superiority for discriminating human compared to monkey voices involving a significant decrease in ability to discriminate among rhesus monkey voices. This is the first study to investigate developmental changes in the human ability to individuate the voices of two unfamiliar human or monkey females and to find evidence for perceptual narrowing of voice discrimination during infant development. Whether this generalizes to male voices remains to be determined.

In Chapter 3 of the thesis, I address the possible role of language familiarity for the results presented in Chapter 2. Since the human voice stimuli in Chapter 2 consisted of English-speaking female voices uttering the English word “balloon”, the human vocal stimuli were native to the monolingual English-speaking participants, both in terms of language spoken and species type. My objective in Chapter 3 was to investigate whether this advantage was driven at least in part by the species of the vocalizations. To test this, I collaborated with co-author Hongting Li in China to test the ability of monolingual Mandarin-speaking adults to discriminate the same human and rhesus monkey vocalizations that were used in Chapter 2. The results demonstrated that the Mandarin-speaking adults discriminated individual human voices much more easily than individual rhesus monkey voices, despite the fact that the human voices were producing utterances in a language foreign to them. A secondary finding was the suggestive evidence for the co-existence of an *own language effect* during voice individuation, consistent with Johnson et al. (2011). English-speaking adults were significantly better than the Mandarin-speaking adults at discriminating the human (English) voices. Although this poorer performance might be explained by other differences between the groups such as education level, it suggests that language might also have an effect. This would be consistent with the face literature, which shows both native-species and native-race biases in adults for face processing (see Lewkowicz & Ghazanfar, 2009 for a review). The results of Chapter 3 suggest that the human-voice processing advantage observed in Chapter 2 by English-speaking adults was driven in large part by familiarity with voices from one's native species. These findings are the first to show that humans become specialized for processing the acoustic structure of the human voice.

In Chapter 4 of the thesis, I investigated the influence of exposure on the process of perceptual narrowing for voice discrimination. Specifically, I examined whether a fixed period of exposure to foreign-species (rhesus monkey) voices at the end of the first year after birth can reinstate 12-month-old infants' sensitivity to individual distinctions in this foreign voice type. The results revealed that 12-month-olds who received two weeks of twice-daily monkey-voice training, in the form of a CD-narrated storybook containing monkey vocalizations, were much better at discriminating individual monkey voices than 12-month-olds who did not receive training. Furthermore, trained 12-month-olds' performance was not significantly different from that of untrained 6-month-olds, suggesting that their sensitivity to differences in monkey voices was reinstated to a 6-month-old level. These results highlight the sizeable influence that exposure has on the process of perceptual narrowing for voice discrimination, even after perceptual narrowing appears to have occurred. These results also provide evidence that the window of plasticity for sensitivity to foreign-species voices stays open at least until the end of the first year. This study is the first to investigate the effects of exposure on the discrimination of foreign voices during development and the findings add considerably to our understanding of how people become specialized for processing differences in the individual voices around them. Importantly, the results indicate that a period of plasticity remains even after narrowing has occurred. It remains for future studies to investigate the duration of this period.

REFERENCES

Agmon, A., Yang, L. T., O'Dowd, D. K., & Jones, E. G., (1993). Organized growth of thalamocortical axons from the deep tier of terminations into layer IV of developing mouse barrel cortex. *The Journal of Neuroscience, 13*(12), 5365-5382.

Aldwell, E., Schachter, C. & Cadwaller, A. (2011). Key, scales, and modes. In *Harmony & Voice Leading (4th edition)* (pp. 3-20). Boston, MA: Schirmer, Cengage Learning.

Allen, J. S., Miller, J. L, & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 113*(1), 544-552.

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology, 112,* 484-495.

Archambault, A., O'Donnell, C. & Schyns, P. G. (1999) Blind to object changes: When learning the same object at different levels of categorization modifies its perception. *Psychological Science 10*(3), 249-55.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development, 52*(4), 1135-1145.

Bachorowskia, J & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America, 106*(2), 1054-1063.

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*(3), 339-350.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*, 9-49.

Black, J., Jones, T., Nelson, C., & Greenough, W. (1998). Neuronal plasticity and the developing brain. In J. Noshpitz (Ed.), *Handbook of child and adolescent psychiatry: Vol. 6. Basic psychiatric science and treatment* (pp. 31-53). New York: Wiley.

Bourgeois, J. P. (1997). Synaptogenesis, heterochrony and epigenesis in the mammalian neocortex. *Acta Paediatrica*, *86*(S422), 27-33.

Bourgeois, J. P. & Rakic, P. (1993). Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. The *Journal of Neuroscience, 13*(7), 2801–2820.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999).

Training Japanese listeners to identify English /r/ and / l/: Long-term retention of learning in speech perception and production. *Perception & Psychophysics*, *61*(5), 977-985.

Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own- and other-race persons. *Journal of Social Issues, 41*, 139-155.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. In B. Beachley, A. Brown & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference of Language Development* (pp. 173-184). Cascadilla, Sommerville, MA.

Byrd, D. (1992). Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America, 92*, 593-596.

Carroo, A. W. (1987). Recognition of faces as a function of race, attitudes and reported cross-racial friendships. *Perceptual and Motor Skills,* 64, 319-325.

Chance, J., Goldstein, A. G., & McBride, L. (1975). Differential experience and recognition memory for faces. *Journal of Social Psychology, 97*, 243-253.

Changeux, J. P., & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature, 264*, 705-712.

Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. *Music Perception, 28*(2), 195-200.

Cowan, W. M., Fawcett, J. W., O`Leary, D. D. M., & Stanfield, B. B. (1984). Regressive events in neurogenesis. *Science, 225*(4668), 1258-1265.

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. In M. G. Gaskell (Ed.), *The oxford handbook of psycholinguistics*. Oxford, UK: Oxford University Press.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176.

DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: Preference, discrimination, and reinforcing value. *Developmental Psychobiology, 17*(5), 481-491.

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience, 14*(2), 199-209.

de Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLoS ONE, 3*(5), e2317.

Endres, W., Bambach, W. & Flösser, G. (1971). Voice spectrograms as a function of age, voice disguise and voice imitation. *Journal of the Acoustical Society of America, 49*, 1842-1848.

Fant, G. (1973). Speech sounds and features. Cambridge, MA: MIT Press.

Feinman, S. & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development, 47,* 506-510.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences, 4*(7), 258-267.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language, 41*, 78-104.

Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental Science, 15*(3), 398-407.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C. & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*(10), 861-863.

Gilmore, J. H., Lin, W., Prastawa, M. W., Looney, C. B., Vetsa, Y. S., Knickmeyer, R. C., Evans, D. D., Smith, J. K., Hamer, R. M., Lieberman, J. A., & Gerig, G. (2007). Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *Journal of Neuroscience, 27*(6), 1255–1260.

Gilmore , J. H., Shi, F., Woolson, S. L., Knickmeyer, R. C., Short, S. J., Lin, W., Zhu, H., Hamer, R. M., Styner, M., Shen, D. (2011). Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cerebral Cortex*, doi: 10.1093/cercor/bhr327.

Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development, 58*(3), 601-622.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*(11), 466-472.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48-55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, USA, 102*(35), 12639-12643.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*(1), 59-67.

Hernandez, A., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638-650.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*(5), 1002-1011.

Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers' faces: Effects of the degree of relatedness. *Proceedings of the Royal Society of Biological Sciences, 276*(3193), 3200.

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology, 104*(1), 105-114.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy. *Psychological Science, 18*(12), 1084-1089.

Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science, 14*, 220-224.

Klatt, D. H. (1986). The problem of variability in speech recognition and in models of speech perception, in J. S. Perkell and D. H. Klatt (Eds), *Invariance and Variability in Speech Processes* (pp. 300-319). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Klatt, D. H., and Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87*, 820- 857.

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., Hamer, R. M., Lin, W., Gerig, G., & Gilmore, J. H. (2008). A structural MRI study of human brain development from birth to 2 years. *The Journal of Neuroscience, 28*(47), 12176 –12182.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 811-817.

Kuhl, P. K. (1998). Effects of language experience on speech perception. *The Journal of the Acoustical Society of America, 103*(5), 2931.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

Kuhl, P. K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research directions* (pp. 213-243). New York: Erlbaum.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*(2), 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, USA, 100*(15), 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*(5044), 606-608.

Lecanuet, J. P., Granier-Deferre, C., Jacquet, A. Y., Capponi, I., & Ledru, L. (1993). Prenatal discrimination of a male and a female voice uttering the same sentences. *Early Development and Parenting, 2*(4), 217-228.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley & Sons, Inc.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences, USA, 103*(17), 6771-6774.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13*(11), 470-478.

Loucks, J., & Sommerville, J. A. (2012). Developmental changes in the discrimination of dynamic human actions in infancy. *Developmental*

Science, 15(1), 123-130.

Low, L. K., & Cheng, H. J. (2006) Axon pruning: An essential step underlying the developmental plasticity of neuronal connections. *Philosophical Transactions of the Royal Society of Biological Sciences 361*(1473)*,* 1531-1544.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Attention, Perception and Psychophysics, 52*(6), 599-608.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1*(4), 272-276.

Macchi Cassia, V., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing: evidence for the reactivation of dormant effects. *Psychological Science, 20*(7), 853-859.

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: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective & Behavioral Neuroscience, 2*(2), 89-108.

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception, 7*(5), 491-497.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public*

Policy and Law 7(1), 3-5.

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

Nelson, C. A., & Ludemann, P. (1989). Past, current and future trends in infant face perception research. *Canadian Journal of Psychology, 43*(2), 183-198.

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Doupoux (Ed.), *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler* (pp. 481-502). Cambridge, MA: MIT Press.

Oh, J. S., Au, T. K. F., & Jun, S. A. (2010). Early childhood language memory in the speech of international adoptees. *Journal of Child Language, 37*(5), 1123-1132.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development, 83*(2), 543-53.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes*, 43, 87-96.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296(5571), 1321-1323.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences, USA, 102*(14), 5297-5300.

Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America, 24*, 175- 184.

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

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences of the United States of America, 106*(26), 10598-10602.

Pruitt, J. S., Jenkins, J. J., & Strange, W. (2006). Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *The Journal of the Acoustical Society of America, 119*(3), 1684-1696.

Purves, D., White, L. E., & Riddle, D. R. (1996). Is neural development Darwinian? *Trends in Neurosciences, 19*(11), 460-464.

Quartz, S. R., & Sejnowsky, T. J. (1997) A neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences, 20*, 537-556.

Rendall, D., Notman, H. & Owren, M. J. (2009). Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *Journal of the Acoustical Society of America, 125*(3), 1792-1805.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behavior, 51*(5), 1007-1015.

Sangrigoli, S., Pallier, C., Argenti, A., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*(6), 440-444.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*(2-3), 125-141.

Scott, L. S., & Monesson, A. (2009). The origins of biases in face perception. *Psychological Science, 20*(6), 676-680.

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. *Neuropsychologia, 48*, 1857-1861.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16*(4), 197-201

Scott, L. S., Shannon, R.W., & Nelson, C. A. (2006). Neural correlates of human and monkey face processing by 9-month-old infants. *Infancy*, *10*(2), 171-186.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience, 18*(9), 1453-1465.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L. & Curran, T. (2008) The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research, 1210*, 204-215.

Simpson, E. A., Varga, K., Frick, J. E., & Fragaszy, D. (2010). Infants experience perceptual narrowing for nonprimate faces. *Infancy, 16*(3), 318-328.

Smith, D. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal tract length in judgments of speaker size, sex, and age. *Journal of the Acoustical Society of America, 118*(5), 3177-3186.

Spence, M. J., & Freeman, M. S. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development, 19*(2), 199-212.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*(6), 2887-2896.

Takagi, N., & Mann, V. A. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics, 16*(4), 379-405.

Tees, R. C., & Werker, J. F. (1984). Perpetual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 34*(4), 579-590.

Trainor, L. J. (2005). Are there critical periods for music development? *Developmental Psychobiology, 46,* 262-278.

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones, & R. R. Fay (Eds.), *Springer handbook of auditory research: Music perception* (pp. 89-128). Heidelberg: Springer.

Trainor, L. J., Marie, C., Gerry, D., Whiskin, E., & Unrau, A. (2012). Becoming musically enculturated: Effects of music classes for infants on brain and behavior. *Annals of the New York Academy of Sciences, 1252,* 129-138.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 394-402.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in western tonal music: Developmental perspectives. *Attention, Perception and Psychophysics, 56*(2), 125-132.

Trainor, L. J., & Unrau, A. J. (2012). Development of pitch and music perception. In L.Werner, R. R. Fay & A. N. Popper (Eds.), *Springer Handbook of*

Auditory Research: Human Auditory Development (pp. 223-254). New York: Springer.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, 47(2), 466-472.

Trehub, S. E. and Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100, 73-99.

Tsao, F., Liu, H., Kuhl, P. K., & Tseng, C. (2000). *Perceptual discrimination of a Mandarin fricative-affricate contrast by English-learning and Mandarin-learning infants*. Poster presented at the meeting of the International Society on Infant Studies, Brighton, England.

Ventureyra, V. A. G., Pallier, C. & Yoo, H. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79-91.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental Science, 15*(3), 359-372.

Wang, Y., & Kuhl, P. K. (2003). Evaluating the "Critical period" hypothesis: Perceptual learning of mandarin tones in American adults and American children at 6, 10 and 14 years of age. *15th ICPhS,* Barcelona. 1537-1540.

Waxman, S. R. & Braun, I. E. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition*, 95, B59-B68.

Webb, S. J., Monk, C. S., & Nelson, C. A. (2001). Mechanisms of postnatal neurobiological development: Implications for human development. *Developmental Neuropsychology, 19*(2), 147-171.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*(1), 349-355.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*(5), 672-683.

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*(1), 147-162.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*9(1), 49-63.

Werker, J. F. & Tees, R. C. (1999). Influences on infant speech processing: Toward

a new synthesis. *Annual Review of Psychology, 50*, 509-535.

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(3), 233-251.

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., Kawakatsu, M., Tohkura, Y., & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage, 46*, 226-240.

CHAPTER 2: Learning to differentiate individuals by their voices: Infants' individuation of native- and foreign- species voices

Friendly, R. F., Rendall, D. & Trainor, L. J. (2013). Learning to differentiate individuals by their voices: Infants' individuation of native- and foreign-species voices. *Developmental Psychobiology: Special Issue on Perceptual Narrowing.* Manuscript in press.

Preface

The ability to perceive important distinctions in a number of socially-relevant auditory and visual stimuli (such as faces and phonemes) has been shown to develop through a process of *perceptual narrowing*, whereby sensitivities to native distinctions increase, while sensitivities to foreign distinctions decrease as infants progress through their first year after birth (for reviews see Lewkowicz & Ghazanfar, 2009; Scott, Pascalis & Nelson, 2007). In Chapter 2, I investigate whether the ability to perceive distinctions in another type of socially-relevant stimulus, the voices of individuals, also develops though a process of perceptual narrowing during this age period. In order to test this, I compared the ability of 6-month-olds, 12-month-olds and adults to discriminate two native-species (human) voices and two foreign-species (rhesus macaque [*Macaca mulatta*]) voices. Results showed that between 6 and 12 months of age, the ability to discriminate rhesus monkey voices deceased significantly and ability to discriminate human voices increased marginally, revealing a pattern of perceptual narrowing during the development of voice discrimination abilities.

ABSTRACT

The ability to discriminate and identify people by their voice is important for social interaction in humans. In early development, learning to discriminate important differences in a number of socially-relevant stimuli, such as phonemes and faces, has been shown to follow a common pattern of experience-driven perceptual narrowing, where the discrimination of native stimuli improves, while the discrimination of foreign stimuli worsens. The aim of the present study was to investigate whether similar perceptual narrowing occurs for discriminating individuals by voice. We tested the ability of English-speaking adults and English-learning 6-month-olds and 12-month-olds to discriminate either native-species (human) or foreign-species (rhesus monkey [*Macaca mulatta*]) individuals by their vocalizations. Between 6 and 12 months of age, the ability to discriminate monkey voices decreased significantly and there was a non-significant trend for improved ability to discriminate human voices. The results support the hypothesis of experience-driven perceptual narrowing for voice individuation during the first year after birth.

INTRODUCTION

The ability to identify individual people quickly and effortlessly is important for human social interaction and likely enhanced survival in our ancestors. One way that humans distinguish individuals is by the unique sound of each person's voice. This ability to distinguish people by voice is especially important in situations where the talker is not clearly visible to the listener, including the example from recent times of talking on the telephone. Because voice recognition plays an important role in social interaction, understanding the developmental trajectory for voice discrimination is important for understanding social and communicative development. Much evidence indicates that complex perceptual processing, such as that involved in face recognition (Kelly et al., 2007; Pascalis, de Haan, & Nelson, 2002; Simpson, Varga, Frick, & Fragaszy, 2010), speech sound categorization (Kuhl et al., 2006; Werker & Tees, 2005), musical pitch and rhythmic processing (Hannon & Trehub, 2005a, 2005b; Lynch & Eliers, 1992; Lynch, Eilers, Oller, & Urbano, 1990; see Hannon & Trainor, 2007; Trainor & Corrigall, 2010; Trainor & Unrau, 2012 for reviews), intersensory (face/voice) processing (Lewkowicz & Ghazanfar, 2006; Pons, Lewkowicz, Soto-Faraco, & Sebastián-Gallés, 2009), and action perception (Loucks & Sommerville, 2012) is strongly influenced during the first year after birth by experience with the specific sounds and objects in the environment in which the infant is developing. Specifically, the ability to make perceptual discriminations relevant to the particular language, musical system and facial features experienced in the native environment improves, while the ability to make discriminations relevant to foreign (and not to native) languages, musical systems and faces diminishes (see Lewkowicz & Ghazanfar, 2009; Scott, Pascalis & Nelson, 2007, for reviews). This process is known as perceptual narrowing (Lewkowicz & Ghazanfar, 2009).

There is extensive evidence for perceptual narrowing for face processing. Adults demonstrate a bias for better discrimination of own-race and own-species faces compared to the discrimination of other-race and other-species faces (e.g., Meissner & Brigham, 2001; Pascalis & Bachevalier, 1998). Findings from behavioural (e.g., Bar-Haim, Ziv & Hodes, 2006; Hayden, Bhratt, Joseph & Tanaka, 2007; Kelly et al., 2005, 2007, 2009; Pascalis et al., 2002; Sangrigoli & De Schonen, 2004; Simpson et al., 2010) and ERP (e.g., de Haan, Pascalis & Johnson, 2002; Scott, Shannon & Nelson, 2006; Scott & Monesson, 2010) studies with infants suggest mixed evidence for such biases at 3 months, clearly developing biases by 6 months, and largely adult-like biases by 9 months of age. For example, when familiarized with one face, and then subsequently presented with the familiarized and a novel face shown from different vantage points (ranging from facing toward the left to the right) both from the same race, Kelly and colleagues found that at 3 months of age Caucasian and Chinese infants can discriminate individuals from another race as accurately as individuals from their own race (Kelly et al., 2007, 2009). However, Caucasian 6-month-olds discriminate both Caucasian and Chinese faces but fail the same tests with African or Middle Eastern faces, and Chinese 6-month-olds discriminate both Chinese and Caucasian but not African faces (they were not tested on the Middle Eastern faces). By 9 months of age, both Caucasian and Chinese infants show no evidence of discrimination for other-race faces. A similar developmental progression to the Kelly et al. studies has also been reported for own-species compared to other-species faces. In this case, 6-month-old infants have been reported to be equally good at discriminating forward-facing faces of individuals from native (human) and foreign (monkey) species but show no evidence of discrimination for monkey faces by 9 months of age (Pascalis et al., 2002). In addition, the ability to detect whether the auditory and visual aspects of vocalizing dynamic faces are synchronous is species- and language-general at 6 months of age but by 8 to 11 months, there is evidence of this intersensory ability for native but not for foreign species and languages (Lewkowicz & Ghazanfar, 2006; Pons et al., 2009).

Despite the importance of discriminating and identifying people by voice, little is known about how this ability develops. However, several studies indicate that infants recognize familiar voices in their environment. By 36 to 39 weeks gestational age, the fetus can discriminate its mother's voice from that of another female (Kisilevsky et al., 2003) and can discriminate between a male and a female voice (Lecanuet, Granier-Deferre, Jacquet, Capponi, & Ledru, 1993). Newborn infants recognize and prefer their mother's voice to that of a female stranger when the voices are producing infant-directed speech (DeCasper & Fifer, 1980; Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978; Spence & Freeman, 1996). Although newborns show no evidence of preferring sentences spoken by their father compared to those spoken by an unfamiliar male, they can tell these voices apart (DeCasper & Prescott, 1984).

When listening to voices, infants are learning to extract two different kinds of information, one related to learning the language spoken and the other to identifying people by their voice. As far as learning language, infants show early sensitivity to prosodic features, preferring to listen to their native language over languages with different prosodic structures (Mehler et al., 1988; Moon, Panneton Cooper & Fifer, 1993). Furthermore, many studies show that learning phonemic categories follows a trajectory of perceptual narrowing similar to that of face processing. Adults show well-developed specialization for native phonemic categories in that they have difficulty discriminating between two speech sounds that fall into different phonemic categories in a foreign language but within a single phonemic category in their native language (Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Kuhl, 1998; Kuhl et al., 2006; Werker & Lalonde, 1988; Werker et al., 2007; Werker & Tees, 1984, 2005). For consonant contrasts, there is no evidence for perceptual narrowing until after 6 months of age, but narrowing is established by 10 to 12 months of age (for reviews see Curtin & Werker, 2007; Kuhl, 2008). For vowel contrasts, narrowing appears to occur somewhat earlier, beginning around 4 months of age (Kuhl, 2004; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Polka & Werker, 1994). Interestingly, perceptual narrowing also occurs for sign language (Palmer, Fais, Golinkoff, & Werker, 2012). English-hearing infants can discriminate American Sign Language (ASL) hand shape distinctions at 4 months, but fail the comparable test at 14 months, of age, while their ASL-learning counterparts can discriminate the distinctions at both ages (Palmer et al., 2012).

Despite the rich literature on language learning, only one study to our knowledge has addressed whether perceptual narrowing occurs for voice identification. Johnson, Westrek, Nazzi, & Cutler (2011) examined the ability of 7-month-old infants to discriminate two different unfamiliar talkers of the same sex. They demonstrated that Dutch 7-month-olds showed a native-language processing bias, discriminating between two unfamiliar female voices speaking Dutch sentences, but failed the same test with two unfamiliar female voices speaking Japanese or Italian sentences. Although further studies are needed with younger and older infants in order to map out the developmental trajectory, this study suggests that, just as infants acquire an own-race bias for discriminating faces, they acquire an own-language bias for discriminating voices.

In the present paper, we examine the related question of whether infants also acquire an own-species bias for the ability to discriminate individuals by voice, similarly to how they acquire an own-species bias for discriminating faces. In particular, we compared the discrimination of human vocalizations to those of a phylogenetically close foreign species, rhesus monkey. Importantly, there is behavioural (Vouloumanos, Hauser, Werker, & Martin, 2010) and neural (Minagawa-Kawai et al., 2011) evidence that young infants can discriminate human and non-human primate vocalizations. Unlike in the case of better discrimination of voices in a familiar versus foreign language (Johnson et al., 2011), where basic characteristics of the human vocal apparatus remain the same across languages spoken, perceptual narrowing for human compared to rhesus monkey voice discrimination would indicate specialization for processing characteristics of sounds made by the adult human vocal tract that are unique to our species (such as a relatively larger oral cavity and descended larynx, see review by Fitch, 2000). We hypothesized that, if the development of voice discrimination occurs through a process of perceptual narrowing, then ability to discriminate native-species voices would improve, while the ability to discriminate foreign-species voices would worsen between 6 to 12 months of age. This age range was chosen because previous research has found evidence of narrowing over this period in other auditory domains, such as perceiving consonant phonemes and musical rhythms (e.g., Curtin & Werker, 2007; Hannon & Trehub, 2005a, 2005b).

METHODS

Participants

Twenty-four adults between 18 and 40 years of age (mean = 21 years, SD = 6.26 years; 12 females), each reporting normal hearing and providing informed consent, participated in this experiment. Participants were drawn from first- and second-year undergraduate psychology classes and received course credit for their participation. All reported (Canadian) English as their only spoken and understood language.

The final infant sample contained forty-eight infants between 5.5 and 6.5 months of age (mean age = 6.01 months, SD = 0.24 months; 25 females), and 48 infants between 11.5 and 12.5 months of age (mean age = 11.98 months, SD = 0.17 months; 19 females), all reported to have normal hearing. Infants were randomly assigned to listen to either human or rhesus monkey voices. All infants were reported as hearing English spoken in their home environment 98-100% of the time. In the sample of 12-month-olds, three additional infants were tested, but excluded from the final sample due to fussiness (n = 1) or failure to pass training (n = 2), all three from the human voice condition. In the sample of 6-month-olds, eleven additional infants were tested, but excluded from the final sample due to fussiness (n = 3, human condition; n = 1, monkey condition) or failure to pass training (n = 4, human; n = 3, monkey). All procedures were approved by the McMaster Research Ethics Board. Informed consent was obtained from parents of infant participants.

Stimuli and Apparatus

Human Voice Stimuli. Voice recordings were made of eight monolingual English-speaking Canadian female adults using a Neutrik AKG (Emotion D 770) microphone, and Felt Tip Sound Studio 2.1 software via a USB Audio/ Midi interface (US - 122) (sampling rate = 44.1 kHz; resolution = 16-bit) on a Macintosh computer (Power Mac G5, OS X version 10.3.9). Female human voices were chosen over male voices as the monkey voice stimuli that we had access to were also of female voices. Each female speaker produced the word *balloon* with six different intonational contours, obtained by reciting the following six sentences: 1. This is a /*balloon*/. 2. Do you want a /*balloon*/? 3. What a great /*balloon*/! 4. Have you seen a /*balloon*/ today? 5. This /*balloon*/ is very light! 6. /*Balloon*/ stands sell lots of balloons! The tokens of the word *balloon* were extracted from the six sentences using Cool Edit Pro and were normalized for intensity. The same word was used for all human vocalizations so that linguistic information would not be informative in any way. The intonational differences were created to ensure that infants were not using intonation as a cue to individuate the voices. The word *balloon* was chosen as the human voice stimulus because it is an infant-friendly word that can be spoken easily with different intonation contours in a variety of carrier sentences.

After acoustic analyses using Praat software (Boersma & Weenink, 2009), two pairs of female voices (pairs 1 and 2) were chosen for use as test stimuli such that the set of six /balloon/ tokens from each of the two voices within a voice pair were matched for mean duration (mean = .694 s, SD = .010 s; mean = .511 s, SD = .067s, for the two pairs, respectively) and minimum F0 (mean = 164.04 Hz, SD = 27.86 Hz; mean = 183.36 Hz, SD = 19.14 Hz), maximum F0 (mean = 348.61 Hz, SD = 78.18 Hz; mean = 279.95 Hz, SD = 39.23 Hz) and mean F0 (mean = 234.57 Hz, SD = 35.84 Hz; mean = 221.64 Hz, SD = 14.50 Hz). In other words, the average of the set of tokens from voice pair 1 was matched to the average of the set of tokens from voice pair 2, for example. This matching was important to ensure that participants would need to use attributes other than these characteristics to discriminate the voices.

Four conditions (1A, 1B, 2A, 2B) were constructed for testing infants. For each voice pair (1, 2), there were two conditions (A, B) such that for condition A one speaker's voice in the pair served as the "background" voice, while the other served as the "change" voice, and vice versa for condition B (see Procedure).

Rhesus Monkey Voice Stimuli. Rhesus monkey (*Macaca mulatta*) voices were chosen as the foreign-species test stimuli for the following reasons. First, the ability of non-human primates to perceive formants (an acoustic feature of voices) is similar to that of humans (Ghazanfar et al., 2007; Owren, 1990; Sommers, Moody, Prosen, & Stebbins, 1992). Second, there is evidence of vocal recognition by rhesus monkeys of both individuals and kin (Rendall, Rodman, & Emond, 1996). Third, six or more vocal samples of the same call category ("coo") from each of six female rhesus monkeys were available from author DR (for methodology on obtaining these recordings see Owren & Rendall, 2003; Rendall et al., 1996). Using these rhesus monkey stimuli, Owren & Rendall (2003) compared human adults' ability to discriminate rhesus coos and human vowel sounds, as well other rhesus vocalizations, such as screams. They showed that human adults found the rhesus coos more difficult to discriminate than human vowels, but easier to discriminate than rhesus screams (Owren & Rendall, 2003).

Example spectrograms comparing the human and rhesus monkey stimuli used in the present study are shown in Figure 1. Both voice types contain a stable fundamental frequency, F0 (the bottom-most band of energy in the spectrograms), above which are harmonic overtones. For both voice types, the amplitudes of the harmonics are filtered by the unique resonances, or formants, of each individual's vocal tract shape and length. The "coo" calls were edited using Cool Edit Pro (sampling rate = 44.1 kHz, resolution = 16-bit) and normalized for intensity. Acoustic analyses using Praat led to the choice of two pairs of primate voices (6 different tokens of the coo call for each individual monkey). Each pair was matched, as with the human voices, on mean duration (mean = .302 s, SD = .047 s; mean = .274 s, SD = .061 s, for the two pairs, respectively) and minimum F0 (mean = 282.36 Hz, SD = 66.28 Hz; mean = 502.48 Hz, SD = 18.93 Hz), maximum F0 (mean = 351.84 Hz, SD = 49.87 Hz; mean = 562.59 Hz, SD = 32.49 Hz) and mean F0 (mean = 320.50 Hz, SD = 48.72 Hz; mean = 542.04 Hz, SD = 26.30 Hz) F0. Four conditions (3A, 3B, 4A and 4B) were constructed for testing infants, as with the human voices, such that for each voice pair (3 and 4), one voice in the pair served as the “background” and the other as the “change” voice for condition A, and vice versa for condition B.

In the adult procedure, the voice samples were presented using the program Presentation on an LG (Antec, Windows XP) computer through headphones (Sennheiser, HDA 200). In the infant procedure, the voice stimuli were presented on a Macintosh computer (Power Mac G4) and the experiment was run using in-house software. The computer was connected to a NAD Stereo Integrated Amplifier (C352) and a single Westsun loudspeaker, and a button box that was used by the experimenter for signaling infant responses to the computer. The button box controller and loudspeaker were located in a sound-attenuating chamber (Industrial Acoustics Company Inc.), with the loudspeaker located on top of a Plexiglas-covered cabinet containing animated toys that served as reinforcement for infant behavioural responses. The computer controlled the animated toys as well as the lights that made them visible (see Procedure for more details).

Insert Figure 1 about here

Procedure

Adult Same-Different Task. Adults were tested using a within-subjects design that measured their discrimination of both the human and monkey vocalizations. Each adult completed two blocks of 40 same-different trials, one containing human and the other rhesus monkey voices. The blocks were presented in random order across participants. On the 20 "same" trials, two voice tokens were presented from the same individual and on the 20 "different" trials, two tokens were presented from two different individuals. On each trial, the two voices presented were always from a voice pair that was matched on duration and mean, minimum and maximum F0 (see Stimuli), and tokens were chosen randomly with the constraints that no token pair be presented more than once, no pair from a "same" trial contain the same token repeated twice, and that equal numbers of each token per voice be presented during testing. Participants were told that they would hear two voice tokens and that they were to press one button if the voices belonged to the same individual, or a second button if the voices belonged to two different individuals. For data analysis, for each adult, percent correct scores in the same-different paradigm were converted to d' separately for the human and monkey voice conditions.

Infant Conditioned Head Turn Task. Because infants are not able to perform the task used with adults, infants were tested in a conditioned head turn paradigm (Werker et al., 1998), using a between-subjects design where infants were tested on their discrimination of human or monkey vocalizations. After the experimenter obtained informed consent from parents, infants were randomly assigned to one of 8 stimulus conditions, counter-balancing species (human, rhesus), voice pair (1, 2 for human; 3, 4 for rhesus) and which voice was the background and which the change voice within the pair (A, B). Thus, at each age, 6 infants were tested in each of the 8 conditions.

During the testing phase, the six /balloon/ or /coo/ tokens from the background voice were repeated in random order with a stimulus onset asynchrony (SOA) of two seconds through the loudspeaker located 90^{o} to the infant's left. The infant was seated on his or her parent's lap across from and facing the experimenter. To eliminate parental or experimenter influence on the infant's behaviour, both parent and experimenter listened to masking music over headphones during the procedure. The background tokens were played continuously throughout the experiment. Once the infant was attentive and facing forward (toward the experimenter), the experimenter pressed a button, indicating to the computer that the infant was ready for a trial. Of the 24 trials, half (12) were change trials, on which the background voice was replaced by one of the 6 tokens of the change voice for one repetition. Each of the six tokens of the change voice was presented twice during the testing phase in random order. The other half of the trials (12) were control (no-change) trials, on which the background voice continued such that control trials were indistinguishable from the background. Change and control trials were presented in a quasi-random order with the constraint that no more than two control trials be presented in a row. The experimenter pressed a second button every time the infant turned their head at least 45 degrees to the left. Head turns on change trials (i.e.,, hits) within 1s of the onset of the change-voice were rewarded with 2s of an animated toy and light display. However, head turns occurring on control trials (i.e., false alarms) were recorded but not rewarded by the computer. For data analysis, the number of hits (head turns during change trials) and false alarms (head turns during control trials) were converted into d' scores.

Prior to the testing phase, infants completed a training phase designed to familiarize them with being rewarded with an animated toy display for turning their head when the voice changed from one individual to another (one human to another when the infant was in the human voice condition, and one monkey to another when the infant was in the monkey voice condition). The training phase did not contain control trials, and only two of the six tokens from the change voice were used. In addition, the change voice was 10 dB louder than the repeating background voice, such that this noticeable difference would attract the infant's attention towards the loudspeaker. Infants were required to make four correct head-turn responses in a row within 20 training trials in order to proceed to the testing phase, where all 6 tokens of each voice were used and change and background tokens were presented at equal intensities. Infants who did not pass this training criterion were excluded from the final data set (see Participants).

RESULTS

Adults

A paired-sample t-test revealed that adults' d' scores for human voices were significantly larger than for rhesus monkey voices, t(23) = 17.90, p < .001 (human, m = 2.76, SE = .13; rhesus, m = .37, SE = .10), suggesting that they discriminated between human voices much more easily than between rhesus monkey voices (see Figure 2). One-sample t-tests revealed that adult's d' scores were significantly above chance levels (expected d' value of 0) for both voice types (both p's < .002). Thus, although adults were able to successfully discriminate between voices for both human and rhesus monkey voice conditions, their performance was much better for voices from their native species compared to voices from the foreign species.

Infants

One-sample t-tests revealed that d' scores were significantly above chance levels ($p < .002$) for each voice type for both the 6-month-olds and 12 month olds after Bonferroni correction (alpha value = 0.0125). An analysis of variance (ANOVA) with d' score as the dependent variable and age (6, 12 months) and voice type (human, rhesus monkey) as independent variables revealed only a significant interaction between age and voice type, $F(1, 92) = 7.95$, $p = .006$, η2p = .08 (see Figure 2). Follow up independent sample t-tests conducted separately for each voice type revealed that the d' scores for rhesus monkey voices were significantly lower at 12 than at 6 months of age, $t(46) = 2.69$, $p = .01$ (6 months, mean = .79, SE = .08; 12 months, mean = .41, SE = .12), suggesting a decrease in discrimination ability. There was also a trend for d' scores for human voices to increase between 6 and 12 months of age, although this effect did not reach significance $t(46) = -1.63$, $p = .11$ (6 months, mean = 1.10, SE = .13; 12 months, mean = 1.46, SE = .18). The results indicate that infants' ability to discriminate between two foreign-species voices decreases between 6 and 12 months of age, supporting the hypothesis of experience-driven perceptual narrowing for voice discrimination during the first year after birth.

Insert Figure 2 about here

DISCUSSION

This is the first study to show that infants' ability to discriminate individuals by voice becomes increasingly specialized for human voices between 6 and 12 months of age. Specifically, between 6 and 12 months of age we found a significant decline in the ability to individuate rhesus monkeys by their vocalizations, as well as a non-significant trend for increased ability to discriminate humans by their vocalizations. The finding that perceptual narrowing occurs by the end of the first year for voice discrimination roughly parallels findings for face discrimination (Kelly et al., 2007; Pascalis et al., 2002; Simpson et al., 2010), intersensory (face/voice) matching (Lewkowicz & Ghazanfar, 2006; Pons et al., 2009), phonemic categorization (Kuhl et al., 2006; Werker & Tees, 2005), and musical pitch and rhythm processing (Hannon & Trehub, 2005a, 2005b; Lynch et al., 1990; for reviews see Hannon & Trainor, 2007; Trainor & Corrigall, 2010; Trainor & Unrau, 2012). Furthermore, it complements a previous study indicating that 7-month-old infants demonstrate an advantage for discriminating voices in their native compared to a foreign language (Johnson et al., 2011), by showing that perceptual narrowing also applies at a species level, with infant perception becoming specialized for processing unique characteristics of human voices during the first year after birth.

Although the ability to discriminate the monkey voices decreased significantly between 6 and 12 months of age, 12-month-olds remained above chance levels for discriminating the monkey voices. This is in contrast to Pascalis et al. (2002) who found that 6-month-olds, but not 9-month-olds and adults, showed evidence of discriminating between two monkey faces above chance levels. This seeming discrepancy might be explained by methodological differences. Pascalis et al. (2002) used a visual paired comparison (VPC) procedure whereas a conditioned head turn was used in the present study. In Pascalis et al.'s study, the VPC involved habituating an infant to one face for a fixed period of 20 seconds (stopping and starting the count whenever the infant looked away; adults only received 5 s of familiarization), and then comparing their looking time to either the same face or a novel face. The Conditioned Head-Turn (CHT) procedure has been reported to be better able than habituation-based methods to separate infant boredom (or disinterest) from perceptual difficulty (Werker, Polka & Pegg, 1997) and is thus likely a more robust measure of infants' perceptual discrimination than the VPC procedure. The CHT procedure also enables researchers to collect data from multiple trials (e.g., 24 trials) as opposed to the 1 or 2 trials in a VPC procedure. Furthermore, although Pascalis et al.'s

(2002) study found no evidence that adults could discriminate monkey faces, Mondlock, Maurer & Ahola (2006) showed that, if tested with more sensitive methodologies, adults can discriminate between monkey faces, albeit more poorly than between human faces. Specifically, they used a task in which participants were required to indicate whether two faces appearing one at a time belonged to the same individual or two different individuals. Therefore, it is possible that older infants and adults might have some ability to discriminate monkey faces, but that this ability was not revealed using the VPC procedure. Furthermore, amount of familiarization time has also been shown to influence the discrimination abilities of older infants. Fair, Flom, Jones & Martin (2012) found that although 12-month-olds failed to demonstrate discrimination of unfamiliar monkey faces following a 20s familiarization period, when familiarization was increased to 40s, 12-month-olds were successfully able to discriminate the monkey faces.

In the present study, we examined changes across age in infants' ability to discriminate human and rhesus monkey voices, but we did not directly compare performance on human versus monkey voices. This comparison was not appropriate as it is very difficult to equate the discriminability of two voice types. Equating discriminability of voices is difficult for two main reasons. First, it is difficult to determine whether human and monkey voice pairs are equally discriminable in general in the absence of experience with voices from either species. In fact, it is possible that human voices are intrinsically easier to process than rhesus monkey voices. Although there are fundamental similarities between rhesus monkey coos and human vowel sounds (Owren & Rendall, 2003, also see Stimuli), considerable differences exist between the vocal anatomies of these species and, therefore, the acoustic properties of their vocalizations, suggesting that voice quality differences could be easier to detect in human than rhesus voices. For example, human vocal folds are longer and thicker than those in rhesus monkeys, permitting a greater degree of variation between individuals' voices (Schon Ybarra, 1995; Titze, 1994). Human vowel sounds also typically contain two to four times the number of harmonics compared to rhesus coo calls (see Figure 1). These extra harmonics enable human vowels to be better defined than rhesus coos, potentially increasing the saliency of the variation between individuals in detailed aspects of their particular resonant patterns (Owren & Rendall, 2003).

The second difficulty in equating human and rhesus voices is more specific to our stimuli. Human vocalizations are typically more complex and contain consonant and vowel sounds whereas rhesus vocalizations contain only vowel-like sounds. We chose to use realistic vocalizations for each species. Our human stimuli consisted of the word "balloon" that contains both consonant and vowel sounds, whereas the monkey stimuli consisted of the very common rhesus coo call that only contains vowel-like sounds. Indeed, the possibility that some acoustic differences may be intrinsically more difficult to discriminate than others is consistent with the finding that acoustic salience influences perceptual narrowing, for example, during the processing of nasal-place phonetic distinctions (Narayan, Werker & Beddor, 2010). Our human vocalizations were also longer, on average, than the monkey vocalizations, so they might have contained more identifying information (about 600 ms versus 300 ms, respectively). In order to determine if length of utterance and the presence of consonants influences comparisons across species, and if our results would generalize to other human vocalizations, future studies could compare narrowing for human voices uttering sounds with vowels only to those uttering sounds with both consonants and vowels.

Nevertheless, the present study revealed a significant decrease in d' scores between 6 and 12 months for monkey voices, but not for human voices, so acoustic differences between the human and monkey voices cannot explain the present finding of perceptual narrowing for monkey voice discrimination between 6 and 12 months of age. This question could be investigated further in a cross-over design in which perceptual narrowing was investigated in both human and rhesus monkey species. If experience is the predominant variable rather than intrinsic properties of human and monkey voices, it would be expected that rhesus monkeys would improve at discriminating monkey voices, but get worse at discriminating the human voices, with increasing age.

We cannot ascertain from the present study whether the timeline of perceptual narrowing for voice discrimination might be underway earlier than that for processing consonant contrasts, faces from foreign species, musical scales and musical meters, for which there is no evidence of narrowing until after 6 months of age (e.g., Curtin & Werker, 2007; Hannon & Trehub, 2005a, 2005b; Hannon & Trainor, 2007; Pascalis et al., 2002). Indeed, examination of Figure 1 appears to indicate that our data more closely resemble the timeline found for the perceptual narrowing of faces from other races, vowel contrasts and the lexical tones within tonal languages such as Mandarin, for which narrowing appears to be underway prior to 6 months after birth (e.g., Kelly et al., 2007, 2009; Kuhl, 2004; Kuhl et al., 1992; Polka & Werker, 1994; Yeung, Chen & Werker, 2013). It is therefore important to test discrimination of these stimuli by infants younger than 6 months of age. Unfortunately, the head turn procedure used in the present study is not suitable for testing auditory discrimination in younger infants (Werker et al., 1998), but future studies could be conducted utilizing appropriate methods to determine whether narrowing might be underway for voice discrimination prior to 6 months of age. A further question of interest for future research is when the sensitive period ends for specialization for human voice discrimination and the influence of experience on this period of sensitivity. It is possible that some plasticity remains throughout the lifespan, such that even in adults, intensive experience with voices from another species would lead to better discrimination of individual voices within that species, although pre-narrowing levels might not be attainable. There is evidence that plasticity remains in infancy beyond the point at which perceptual narrowing appears to be accomplished. For example, Pascalis et al. (2005) demonstrated that the loss of ability to discriminate between primate faces observed between 6 to 9 months of age could be prevented by three months of exposure during this period to primate faces. Also, Anzures et al. (2012) showed that it was possible to reinstate the ability of 8- to 10-month-olds to recognize foreign-race faces with 3 weeks of exposure to these faces. In the language domain, one month of interaction with Mandarin Chinese speakers reversed the decline in ability to distinguish between Mandarin phonemes in English-learning infants previously observed at 12 months (Kuhl, Tsao, & Liu, 2003). Similarly, two weeks of exposure to foreign musical rhythms at 12 months reinstated sensitivity to these rhythms whereas a similar amount of exposure had no effect in adulthood (Hannon & Trehub, 2005b).

One of the first studies to directly investigate the effect of environmental exposure on sensitive periods for faces in a design that strictly controlled experience was conducted by Sugita (2008). In this study, infant monkeys were reared with no exposure to live faces for periods ranging from 6 to 24 months of age. Prior to exposure, the monkeys discriminated photographs of human faces as easily as photographs of monkey faces. After the end of the period of total deprivation for face stimuli, one month of subsequent exposure to either human or monkey faces resulted in the monkeys only being able to discriminate the face type to which they were exposed for that month. This demonstrates that the period of plasticity for face specialization can be extended for at least 2 years in the complete absence of experience. These results, in addition to the aforementioned studies by Pascalis et al. (2005), Anzures et al. (2012) and Kuhl et al. (2003), suggest that the timing of exposure as well as the type and amount of exposure affect perceptual narrowing. These variables should also be explored with respect to voices to determine the generality of the effects of timing, type and amount of exposure on sensitive periods.

Finally, questions remain as to whether the results of the present study generalize to foreign human languages, familiar voices, and to male in addition to female voices. It would also be interesting to know whether young infants can discriminate non-primate voices that are less similar to human voices than rhesus monkey voices, and whether the ability to discriminate voices from different species follows the same developmental trajectory as for primate voices.

Conclusions

The results indicate that young infants have the important social ability to discriminate individuals by voice, and that it is shaped by the auditory environment. Specifically, we found a significant decrease in the ability to discriminate foreign-species (rhesus monkey) voices between 6 and 12 months of age, accompanied by a non-significant trend for the ability to discriminate native-species (human) voices to increase.

NOTES

This research was supported by a grant to LJT from the Natural Sciences and Engineering Research Council of Canada. We thank Lauren Kutcher and Andrea Unrau for assistance in data collection.

REFERENCES

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. Journal of Experimental Child Psychology, 112, 484- 495.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. Psychological Science, 17, 159-163.

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(3), 339-350. doi:10.1016/0163-6383(95)90022-5.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer Version 5.1.04. Computer program. Available from http://www.praat.org/.

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. In M. G. Gaskell (Ed.), The oxford handbook of psycholinguistics. Oxford, UK: Oxford University Press.

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. Journal of Cognitive Neuroscience, 14(2), 199-209.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. Science, 208(4448), 1174-1176.

DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: Preference, discrimination, and reinforcing value. Developmental Psychobiology, 17(5), 481-491.

Fair, J., Flom, R., Jones, J., & Martin, J. (2012). Perceptual Learning: 12-Month-Olds' discrimination of monkey faces. Child Development, 83(6), 1996-2006.

Fitch, W. T. (2000). The evolution of speech: A comparative review. Trends in Cognitive Sciences, 4(7), 258-267.

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D. & Logothetis, N. K. (2007). Vocal-Tract Resonances as Indexical Cues in Rhesus Monkeys. Current Biology, 17, 425-430.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. Trends in Cognitive Sciences, 11(11), 466-472.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. Psychological Science, 16(1), 48-55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. Proceedings of the National Academy of Sciences, USA, 102(35), 12639-12643.

Hayden, A., Bhatt, R. S., Joseph, J. E., & Tanaka, J. W. (2007). The other-race effect in infancy: Evidence using a morphing technique. Infancy, 12, 95-104

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. Developmental Science, 14(5), 1002-1011.

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? Journal of Experimental Child Psychology, 104(1), 105-114.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy. Psychological Science, 18(12), 1084-1089.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., et al. (2005). Three- month-olds, but not newborns, prefer own-race faces. Developmental Science, 8, F31-F36.

Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., & Wang, Z. (2003). Effects of experience on fetal voice recognition. Psychological Science, 14(3), 220-224.

Kuhl, P. K. (1998). Effects of language experience on speech perception. The Journal of the Acoustical Society of America, 103(5), 2931.

Kuhl, P. K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo & L. Freund (Eds.), Infant pathways to language: Methods, models, and research directions (pp. 213-243). New York: Erlbaum.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. Nature Reviews Neuroscience, 5, 831-843.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. Developmental Science, 9(2), 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(15), 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(5044), 606-608.

Lecanuet, J. P., Granier-Deferre, C., Jacquet, A. Y., Capponi, I., & Ledru, L. (1993). Prenatal discrimination of a male and a female voice uttering the same sentences. Early Development and Parenting, 2(4), 217-228.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. Proceedings of the National Academy of Sciences, USA, 103(17), 6771-6774.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. Trends in Cognitive Sciences, 13(11), 470-478.

Loucks, J., & Sommerville, J. A. (2012). Developmental changes in the discrimination of dynamic human actions in infancy. Developmental Science, 15(1), 123-130.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. Attention, Perception and Psychophysics, 52(6), 599-608.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. Psychological Science, 1(4), 272-276.

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. Perception, 7(5), 491-497.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 144-178.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. Psychology, Public Policy and Law 7(1), 3-5.

Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., &

Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. Cerebral Cortex, 21(2), 254-261.

Mondlock, C. J., Maurer, D. & Ahola, S. (2006). Becoming a Face Expert. Psychological Science, 17(11), 930-934.

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

Narayan, C., Werker, J.F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. Developmental Science, 13(3), 407-420.

Owren, M. J. (1990). Acoustic classification of alarm calls by vervet monkeys (cercopithecus aethiops) and humans (homo sapiens): II. synthetic calls. Journal of Comparative Psychology, 104(1), 29-40.

Owren, M. J., & Rendall, D. (2003). Salience of caller identity in rhesus monkey (macaca mulatta) coos and screams: Perceptual experiments with human (homo sapiens) listeners. Journal of Comparative Psychology, 117(4), 380-390.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. Child Development, 83(2), 543-53.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. Behavioural Processes, 43, 87-96.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? Science, 296(5571), 1321-1323.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. Proceedings of the National Academy of Sciences, USA, 102(14), 5297-5300.

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

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. Proceedings of the National Academy of Sciences of the United States of America, 106(26), 10598-10602.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. Animal Behavior, 51(5), 1007-1015.

Sangrigoli, S., & De Schonen, S. (2004). Recognition of own-race and other-race faces by three-month-old infants. J Child Psychol Psychiatry, 45(7), 1219-27. doi:10.1111/j.14697610.2004.00319.

Schon Ybarra, M. (1995). A comparative approach to the non-human primate vocal tract: Implications for sound production. In E. Zimmermann, J. D. Newman & U. Jürgens (Eds.), Current topics in primate vocal communication (pp. 185-198). New York: Plenum.

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. Neuropsychologia, 48, 1857-1861.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. Current Directions in Psychological Science, 16(4), 197-201.

Scott, L.S., Shannon, R.W., & Nelson, C.A. (2006). Neural correlates of human and monkey face processing by 9-month-old infants. Infancy, 10(2), 171-186.

Simpson, E. A., Varga, K., Frick, J. E., & Fragaszy, D. (2010). Infants experience perceptual narrowing for nonprimate faces. Infancy, 16(3), 318-328.

Sommers, M. S., Moody, D. B., Prosen, C. A., & Stebbins, W. C. (1992). Formant frequency discrimination by Japanese macaques (macaca fuscata). Journal of the Acoustic Society of America, 91(6), 3499-3510.

Spence, M. J., & Freeman, M. S. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. Infant Behavior and Development, 19(2), 199-212.

Sugita, Y. (2008) Face perception in monkeys reared with no exposure to faces.

Proceedings of the National Academy of Sciences, 105(1), 394-398.

Titze, I. R. (1994). Principles of voice production. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones, & R. R. Fay (Eds.), Springer handbook of auditory research: Music perception (pp. 89-128). Heidelberg: Springer.

Trainor, L. J., & Unrau, A. J. (2012). Development of pitch and music perception. In L. Werner, R. R. Fay & A. N. Popper (Eds.), Springer handbook of auditory research: Human auditory development (pp. 223-254). New York: Springer.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. Child Development, 81(2), 517-527.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. Developmental Psychology, 24(5), 672-683.

Werker, J. F., Polka, L. & Pegg, J. E. (1997). The conditioned head turn procedure as a method for testing infant speech perception. Early Development and Parenting, 6 (34), 171-178.

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(1), 147-162.

Werker, J. F., Shi, R., Desjardins. R., Pegg, J.E., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. In A.M. Slater (Ed.), Perceptual development: Visual, auditory, and speech perception in infancy, (pp. 389-420) London: UCL Press.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development, 79(1), 49-63.

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(3), 233-251.

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

FIGURES

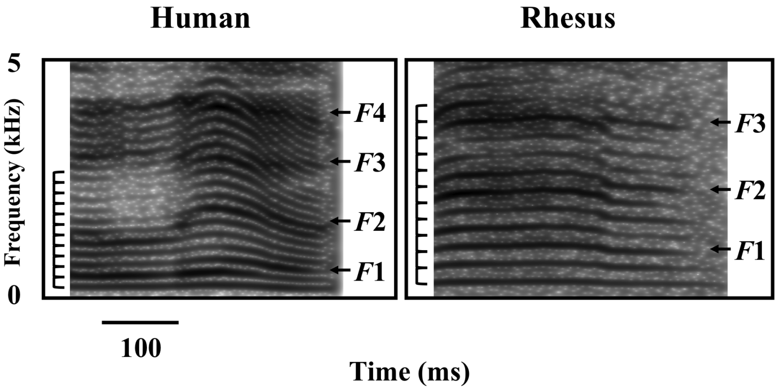

Figure 1. Spectrograms of the word "balloon" spoken by an adult human female and a "coo" vocalization produced by an adult female rhesus monkey. The sounds are generally similar in showing a stable fundamental frequency (F0), the lowest band on the spectrogram, with rich harmonic overtones that are filtered by the resonances, or formants, of the vocal tract (F1-F4). The sounds are different in that the F0 is higher in the rhesus monkey sounds (375 Hz versus 230 Hz in the human female sounds), which yields more widely spaced harmonic overtones in the rhesus monkey sounds (marked by the vertical grids at the left of each spectrogram). The formant frequencies are also higher in the rhesus monkey sounds. Note that the human vocalizations used were about 600 ms whereas the rhesus monkey vocalizations were about 300 ms.

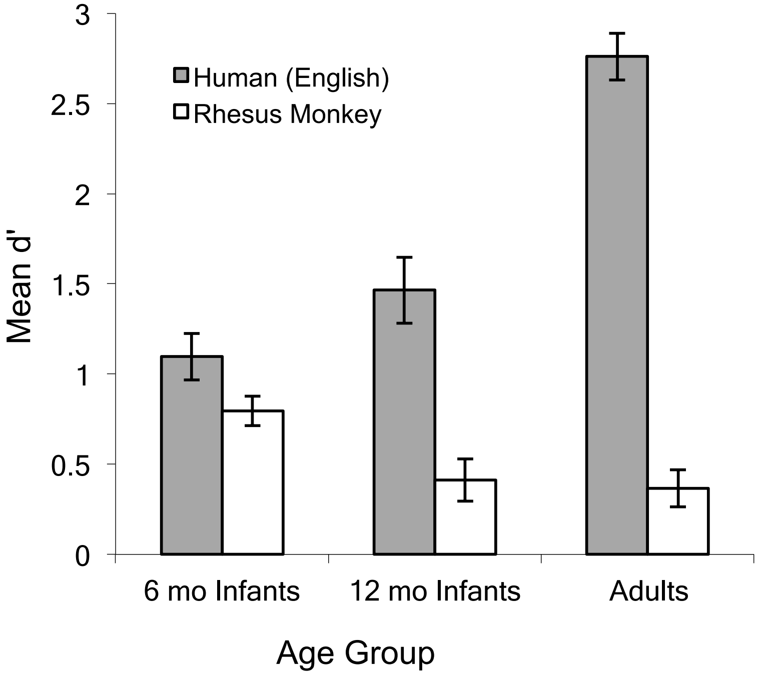

Figure 2. Adult (n = 24) and infant (6- and 12-month-olds, n = 48 each) discrimination of rhesus monkey and human voices, as measured by a same-different task and a conditioned head turn task, respectively. Although it is not appropriate to compare discrimination across voice types (see text), discrimination of human voices improved with age whereas discrimination of monkey voices declined with age. Error bars represent within-subjects SEM for adults and between-subjects SEM for infants.

CHAPTER 3: Own-species and own-language effects in English and Chinese adults' cross-species voice discrimination

Friendly, R. F., Hongting, L., Rendall, D. & Trainor, L. J. (2013). Own-species and own-language effects in English and Chinese adults' cross-species voice discrimination. *Journal of the Acoustical Society of America (Express Letters).* Manuscript prepared for submission.

Preface

In the previous chapter, perceptual narrowing was found to occur during the development of human (English-learning) participants' ability to discriminate voices. An advantage developed between 6- to 12-months of age for the processing of individual distinctions of native (human, English-speaking) voices, compared to foreign (rhesus macaque [*Macaca mulatta*]) voices, and this advantage was also present in English-speaking adults. In Chapter 3, I examine the ability of adults to discriminate the same human and monkey vocalizations cross-culturally. In particular, I investigated whether Mandarin-speaking adults from China, who had little or no exposure to the English language, would also demonstrate an advantage for discriminating human, English-speaking voices compared to the monkey voices, similar to the English adults in Chapter 2. The results reveal that the Chinese adults do indeed demonstrate a human-voice processing advantage. This suggests that the advantage for discriminating human voices documented for adults tested in Chapter 2 reflects familiarity with voices from the human species. However, the data also revealed evidence suggestive of an own-language superiority effect, in that English adults performed better than Chinese adults on the human (English) voice discrimination. This difference might be explained by other factors such as education level differences between the English-speaking and Mandarin-speaking groups, but it is likely that perceptual narrowing for voice discrimination involves specialization for both own-species and own-language voices.

ABSTRACT

Friendly, Rendall and Trainor (submitted) demonstrated that English-speaking adults and 12-month-old infants discriminate native-species (English-speaking human) individuals by their vocalizations much more easily than foreign-species (rhesus monkey) individuals, while 6-month-olds show little difference across species. Here we test whether this native-species superiority effect reflects familiarity with human vocalizations in general or the English language in particular, at least in adults. Mandarin Chinese-speaking adults with limited English exposure showed much better discrimination of English-speaking human individuals than monkey individuals. We also found suggestive evidence for an effect of language familiarity when comparing voice discrimination abilities of Mandarin- and English-speaking adults. The results indicate a robust own-species superiority effect for discriminating individuals by their voices.

1. Introduction

In order for social interaction to be successful, humans must be able to identify and distinguish between individuals in their environment. Variations among peoples' voices provide a cue to listeners regarding the identity of speakers. Adults are better at discriminating among voices speaking their familiar native language compared to voices speaking an unfamiliar foreign language (e.g., Goggin et al., 1991; Köster & Schiller, 1997). Recently, we found a similar familiarity effect for native-species over foreign-species voices. English-speaking adults were considerably better at discriminating between voices from their own (human) species than from a foreign (rhesus monkey) species (Friendly et al., submitted). This is consistent with studies showing stronger activations in an area of the brain associated with voice-processing (the superior temporal sulcus) when listening to human vocalizations compared to vocalizations from monkeys and other non-human animals (e.g., Fecteau et al., 2004), and parallels visual-perception studies showing that adults are better at individuating human than monkey faces (e.g., Pascalis & Bachevalier, 1998; Pascalis et al., 2002). Additionally, our study demonstrated that the ability to discriminate voices goes through a process of perceptual narrowing, with superiority for native- over foreign-species voice discrimination clearly developing between 6 months and 12 months of age, similar to the narrowing timelines for the discrimination of a number of other socially-relevant stimuli, such as phonemes, musical meters, and faces (see Lewkowicz & Ghazanfar, 2009; Scott et al., 2007, for reviews).

Nevertheless, it is unclear from our previous study if the bias for better performance with human compared to rhesus monkey voices is driven primarily by species familiarity (i.e., familiarity with the acoustic structure of human versus rhesus monkey voices in general) or rather by linguistic familiarity, as the human voice stimuli used were produced by speakers of the native language (English) of our participants, and used tokens of an English word. Here we examine this question in adults by testing a population of Mandarin-Chinese-speakers from China on their ability to discriminate the human English voices and rhesus monkey voices used in our previous study. We also compared their voice discrimination abilities to those of the English-speaking adults from Friendly et al. (submitted). The Chinese population was chosen because Mandarin is a tonal language and, as such, differs substantially in structure from English. Furthermore, the population we recruited in China did not speak, and had limited exposure to, English. Finding that Mandarin-speaking adults are similar to English-speaking adults in that they more easily discriminate human, English-speaking voices than rhesus monkey voices would support the hypothesis that species familiarity is a driving factor in the native (human) voice processing advantage observed in English-speaking adults.

2. Methods

2.1 Participants

Twenty-four English-speaking adults from Ontario, Canada (mean age = 21 years, SD = 6.26 years; range = 18- 40 years; 12 females) and 24 Mandarin-speaking adults from Zhejiang, China (mean age = 26 years, SD = 7.11 years; range = 18- 40 years; 12 females) participated. Informed consent was obtained from all participants. All reported normal hearing. The English-speaking adults were tested in a previous study (see Friendly et al., submitted) and received course credit for their participation. All reported (Canadian) English as their only spoken and understood language. Mandarin-speaking adults had as little experience with the English language as was possible. As English is generally studied in higher-level education in China, all participants were manual factory workers recruited by advertisements posted in a factory in Hangzhou City. One fifth of participants had only a primary school education, while the remainder had incomplete junior high educations. All participants reported speaking Mandarin 95% of the time or more. Two thirds of the participants reported never studying English, whereas one third had received a limited English education, learning a few basic English words and phrases but not having used them for many years. Chinese participants received 10 Renminbi (about 1.6 CAN/US dollars) as compensation for their participation in the study.

2.2 Stimuli and Apparatus

The stimuli outlined below are identical to those used by Friendly et al. (submitted).

2.2.1 Human Voice Stimuli

The vocalizations of eight English-speaking Canadian female graduate students were recorded using Felt Tip Sound Studio 2.1 software and an AKG Emotion D 770 (cardioid-directional dynamic) microphone via a USB Audio interface (US - 122) (sampling rate = 44.1 kHz; [intensity] resolution = 16-bit) on a Macintosh computer (Power Mac G5, OS X version 10.3.9). Each female speaker was asked to recite six sentences containing the word "balloon" in order to produce six different /balloon/ tokens, each with a different intonation contour. The sentences were: 1. Do you want a /balloon/? 2. What a great /balloon/! 3. This is a /balloon/. 4. Have you seen a /balloon/ today? 5. /Balloon/ stands sell lots of balloons! 6. This /balloon/ is very light! The program Cool Edit Pro (Syntrillium Software) was used to extract the /balloon/ tokens from the six carrier sentences and to normalize the tokens for peak intensity across the samples (Pair 1: voice 1 mean = 49 dB, range = 47-51 dB, voice 2 mean = 47 dB, range = 46-48 dB; Pair 2: voice 1 mean = 47 dB, range = 45-49 dB, voice 2 mean = 50 dB, range = 48-53 dB).

Acoustic analyses were performed using Praat software (Boersma & Weenink, 2009), where fundamental frequencies (F0) were extracted using the autocorrelation algorithm restricted to search between 100 and 600 Hz. Two pairs of female voices (pairs 1 and 2) were then chosen for use as test stimuli such that the six /balloon/ tokens from the two voices in each pair did not differ significantly in duration or in minimum, maximum or mean F0.

2.2.2 Rhesus Monkey Voice Stimuli

Rhesus macaque (*Macaca mulatta*) vocalizations were chosen for a number of reasons. First, there is evidence that rhesus monkeys can recognize both kin and other individuals based on some vocalizations in their repertoire (Rendall et al., 1996). Second, the ability to perceive formants, an acoustic feature of both human and nonhuman primate voices, is similar in both groups (Ghazanfar, Turesson, Maier, van Danther, Patterson & Logothetis, 2007; Owren, 1990; Sommers, Moody, Prosen, & Stebbins, 1992). Third, author DR had, for each of six female rhesus monkeys, several vocal samples of the same call category, the so-called “coo” call, which is a close-range contact call and is individually distinctive and discriminated by monkeys (for methodology on obtaining these recordings see Owren & Rendall, 2003; Rendall et al., 1996, 1998). Finally, human adults’ discrimination of rhesus coos has been previously tested and compared to the discrimination of human vowel sounds, as well as other types of rhesus vocalizations. These comparisons revealed that rhesus coo calls are more difficult for adults to discriminate than human vowels, but easier to discriminate than an alternative rhesus vocalization, namely the screams given in agonistic encounters

(Owren & Rendall, 2003). The rhesus coo calls were edited using Cool Edit Pro (sampling rate = 44.1 kHz, [intensity] resolution = 16-bit) and normalized for peak amplitude across the sample (Pair 1: voice 1 mean = 49 dB, range = 46-54 dB, voice 2 mean = 49 dB, range = 47-51 dB; Pair 2: voice 1 mean = 55 dB, range = 53-57 dB, voice 2 mean = 56 dB, range = 55-57 dB). Two pairs of primate voices (6 different tokens of the coo call for each individual monkey) were chosen based on acoustic analyses using Praat. As with the human voices, the six tokens from each pair were matched for duration and for minimum, maximum and mean F0.

For both the English-speaking and Mandarin-speaking participants, the voice samples were presented using Presentation (Neurobehavioral Systems). For English participants, Presentation was run on an LG (Antec, Windows XP) computer through headphones (Sennheiser, HDA 200). For Chinese participants, Presentation was run on a Toshiba (M100, Windows XP) computer through headphones (Philips, SHM 1900).

2.3 Procedure

Each adult completed two blocks of 40 same-different trials. One block contained rhesus monkey voices and the other human voices. The order of the blocks was counterbalanced across participants. Each block contained 20 “different” trials (one voice token was spoken by each of two different individuals) and 20 “same” trials (two different voice tokens were spoken by the same individual). On different trials, the two voices presented were always from one of the two matched pairs (see Stimuli). Across all trials, tokens were chosen randomly with the constraints that an equal number of each token per voice be presented during testing, no pair from a "same" trial contain the same token repeated twice, and no token pair be presented more than once. Participants were instructed that they would hear two vocalizations, and that they were to press one button if the vocalizations were spoken by two different individuals, or a second button if the vocalizations were spoken by the same individual. Percent correct scores were converted to d' separately for the monkey and human voice conditions for each adult.

3. Results

A repeated measures analysis of variance (ANOVA) with d' score as the dependent variable, voice type as a within-subject variable (human, rhesus) and language spoken as a between-subject variable (English, Mandarin) revealed a significant interaction between voice type and language spoken, $F(1,46) = 37.59$, $p < .001$ (see Figure 1). Paired samples t-tests conducted separately for the two language groups revealed that both English-speaking adults [$t(23) = 17.90$, $p < .001$] and Mandarin-speaking adults [$t(23) = 6.57$, $p < .001$] discriminated human voices more easily than rhesus monkey voices, demonstrating an own-species superiority effect. One-sample t-tests conducted separately for each adult language type also revealed that both English-speaking (human: $t(23) = 21.21$, $p <$

.001; monkey: t(23) = 3.58, p = .002) and Mandarin-speaking types (human: t(23) = 8.94, p < .001; monkey: t(23) = 4.30, p < .001) adults' discrimination was significantly above chance levels (expected d' value of 0) for both voice types. A one-way ANOVA revealed that although English-speaking and Mandarin-speaking adults showed no differences in their discrimination abilities for rhesus monkey voices, English-speaking adults discriminated the human, English-speaking voices more easily than did the Mandarin-speaking adults, *F*(1,46) = 32.69, *p* < .001, suggesting that there might be a language familiarity effect in addition to the species familiarity effect.

4. Discussion

Mandarin-speaking adults were better at discriminating English-speaking human voices than rhesus monkey voices even though the human voices uttered words from an unfamiliar language. This provides strong evidence for an own-species superiority effect for voice discrimination in adults, which parallels the own-species superiority effects found for faces (Pascalis et al., 2002). These data complement recent results showing better performance at discriminating voices from one's native than from a foreign (monkey) species at 12-months of age and in adulthood, but little difference between human and monkey voices at 6 month of age (Friendly et al., submitted). The present study suggests that the superior performance on own- than foreign-species voices observed in English-speaking adults in Friendly et al. (submitted), results in large part from experience with own-species voices. It is important to note that although both adult populations demonstrated a human voice processing advantage, they also both had d' scores for monkey voices that were above chance levels. This indicates that adults are able to discriminate monkey voices, albeit poorly in comparison to human voices. Thus, although adults specialize for processing human voices, some sensitivity still remains to distinctions in monkey voices.

At the same time, English-speaking adults performed better than Mandarin-speaking adults at discriminating voices uttering the English words. Mandarin adults might have performed less well for a number of reasons. First, it is possible that this result reflects differences in task comprehension and/or performance due to the education levels of the two groups. The English-speaking participants were largely university students whereas the Mandarin-speaking participants were factory workers who had not completed high school, which was a deliberate precaution to ensure limited exposure to English. Second, it is possible that the necessary use of two different sets of testing equipment (one in China, and one in Canada) could have led to the observed differences. Third, the motivation of the two samples to complete the task might have differed. However, the most likely explanation is that the difference reflects the additional operation of an own-language superiority effect, consistent with previous studies (e.g., Goggin et al., 1991; Köster, & Schiller, 1997). The finding that native voices are processed more easily than foreign voices is also in line with numerous studies that find similar biases in adults for the processing of various social stimuli including linguistic phonemes, musical meter, and intersensory face/voice stimuli (see Lewkowicz & Ghazanfar, 2009 for a review). Nevertheless, as the present study did not use a crossover design (comparing the discrimination of English- and Mandarin-voices by English- and Mandarin-speaking adults), further investigation is required before firm conclusions can be made with respect to the language effect.

Understanding the basis of the specialization for own-species voices requires a consideration of differences in vocal tract shape between human and rhesus monkeys, and the resulting differences in acoustic features of the sounds they produce. For instance, humans lack the air sac that non-human primates have and tend to have a longer oral cavity and lower larynx (e.g., Fitch, 2000). Compared to rhesus monkeys in particular, humans have longer supralaryngeal vocal tracts (Fitch, 1997; Fitch & Giedd, 1999), longer and thicker vocal chords (Schon Ybarra, 1995), and produce two to four times the number of harmonics in their vowel sounds (Owren & Rendall, 2003) (see Figure 2). This is true for human vocalizations that include vowels regardless of the language being spoken. Thus, it appears that, through experience with human voices, human perceptual systems become tuned to discriminate fine differences in stimuli with relatively rich harmonic spectra and strong resonance (formant) features characteristic of human voices, compared to stimuli, like rhesus monkey coos, that are broadly similar to human vowels in formant structure but have much less densely populated harmonic spectra.

Of course, it is possible that other differences beyond vocal anatomy, such as the complexity the vocalizations, might have influenced the pattern of results. For instance, the average duration of the rhesus monkey stimuli was approximately 300ms shorter than for the human voices. As well, the Human English stimuli contained consonant and vowel sounds, while the monkey stimuli only contained vowel sounds. Including consonants was necessary in order to determine if there was a language effect, but future studies could compare the discrimination of rhesus monkey coos to that of human voices producing only vowels in order to determine if the complexity of the human voices contributed to the human voice processing advantage found in both English- and Mandarin-speaking adults. It needs to be kept in mind, however, that such differences are very unlikely to explain the entire bias for better discrimination of native- compared to foreign-species voices because this bias is not found in young infants who discriminate native- and foreign-species voices similarly well (Friendly et al., submitted).

In sum, adults show an own-species superiority effect for discriminating individuals by voice that extends beyond familiarity with a particular language.

5. Acknowledgments

This research was supported by a grant to LJT from the Natural Sciences and Engineering Research Council of Canada and a grant to HT Li from the Zhejiang Qianjiang Talent Foundation (1013371-M). We thank Lauren Kutcher and Andrea Unrau for assistance in data collection.

References and Links

Boersma, P., & Weenink, D. (2009). *Praat: Doing Phonetics by Computer Version 5.1.04*. Computer program. Available from http://www.praat.org/.

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage, 23*(3), 840-848.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*(3), 1511-1522.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America, 102*(2), 1213-1222.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences, 4*(7), 258-267.

Friendly, R. H., Rendall, D., & Trainor, L. J. (submitted). Learning to differentiate individuals by their voices: Infants' individuation of native- and foreign- species voices. *Developmental Psychobiology: Special Issue on Perpetual Narrowing.*

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D. & Logothetis, N. K. (2007). Vocal-Tract Resonances as Indexical Cues in Rhesus Monkeys. *Current Biology, 17*, 425-430.

Goggin, J. P., Thompson, C. P., Strube, G., & Simental, L. R. (1991). The role of language familiarity in voice identification. *Memory & Cognition, 19*(5), 448-458.

Köster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics, 4*(1), 18-28.

Lewkowicz, D. J. & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13* (11), 470-478.

Owren, M. J., & Rendall, D. (2003). Salience of caller identity in rhesus monkey (macaca mulatta) coos and screams: Perceptual experiments with human (homo sapiens) listeners. *Journal of Comparative Psychology, 117*(4), 380-390.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioral Processes, 43*(1), 87-96.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science, 296*(5571), 1321-1323.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behavior, 51*(5), 1007-1015.

Rendall, D., Owren, M.J. & Rodman, P.S. (1998) The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations.

Journal of the Acoustical Society of America 103: 602-614.

Schon Ybarra, M. (1995). A comparative approach to the non-human primate vocal tract: Implications for sound production. In E. Zimmermann, J. D. Newman & U. Jürgens (Eds.), *Current topics in primate vocal communication* (pp. 185-198). New York: Plenum.

Scott, L.S., Pascalis, O. & Nelson, C. A. (2007). A domain general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16*(4), 197-201

Sommers, M. S., Moody, D. B., Prosen, C. A., & Stebbins, W. C. (1992). Formant frequency discrimination by Japanese macaques (*macaca fuscata*). *Journal of the Acoustic Society of America, 91*(6), 3499-3510.

Figures

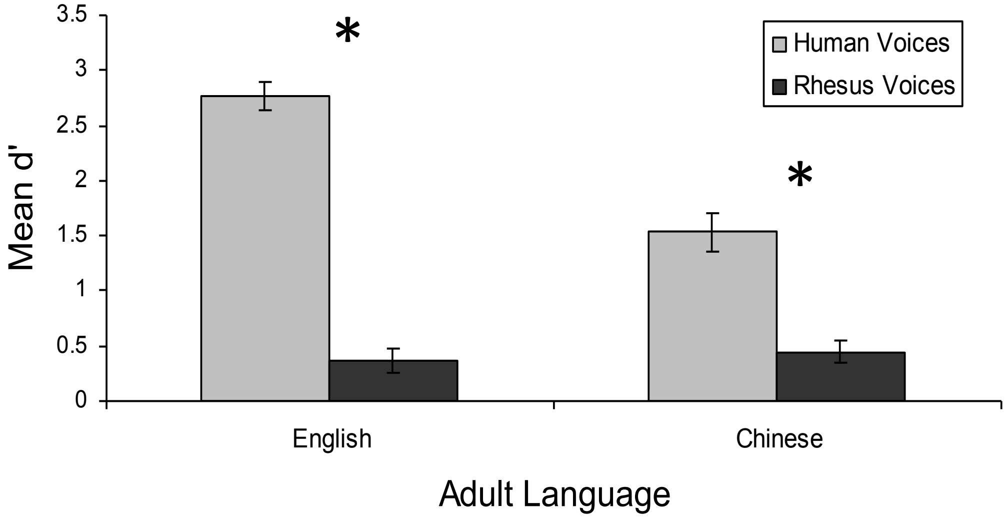

Figure 1. English- (n = 24) and Mandarin- (n=24) speaking adults' discrimination of rhesus monkey individuals and human individuals by their voices. Both English-speaking and Mandarin-speaking adults demonstrated a species familiarity effect, discriminating between human voices much more easily than between rhesus monkey voices (both p's < .001). In addition, English-speaking adults better discriminated between the English-speaking human voices than did the Mandarin-speaking adults, suggesting a language-familiarity effect (p < .001). Error bars represent within-subjects SEM.

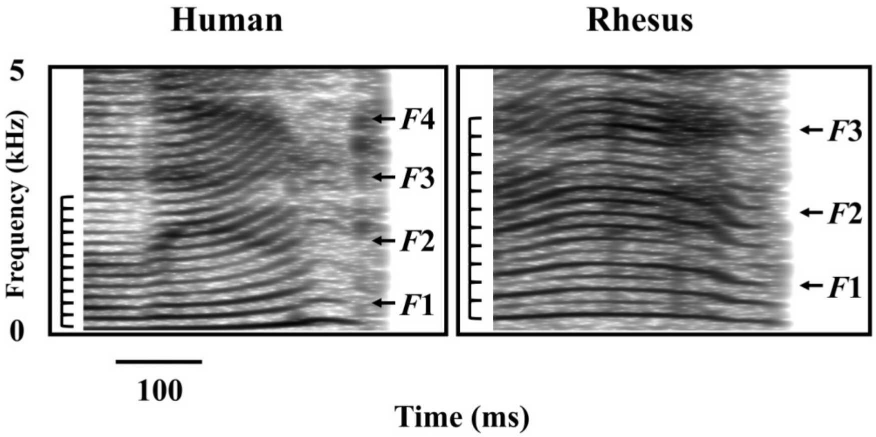

Figure 2. A "coo" call vocalization produced by a female adult rhesus monkey and the English word "balloon" produced by a female adult human are represented here in the form of spectrograms. In both vocalizations, the lowest band on the spectrogram (i.e., the fundamental frequency [F0]) is stable, and both vocalizations contain harmonic overtones, which are filtered by formants (F1-F4), or the resonances of the vocal tract. However, the F0 is higher in the rhesus "coo" vocalization (418 Hz versus 209 Hz) and thus the formant frequencies are more widely spaced than in the human "balloon".

CHAPTER 4: Plasticity after perceptual narrowing for voice perception: Reinstating the ability to discriminate monkeys by their voices at 12 months of age

Friendly, R. F., Rendall, D. & Trainor, L. J. (2013). Plasticity after perceptual narrowing for voice perception: Reinstating the ability to discriminate monkeys by their voices at 12 months of age. *Frontiers in Auditory Cognitive Neurosciences.* Manuscript prepared for submission.

Preface

In the previous two chapters, I provide evidence that perceptual narrowing occurs during the development of the ability to discriminate voices. These studies demonstrate that human infants become specialized for processing distinctions relevant to individual identity in voices from their native- (human) species during their first year after birth, while decreasing in sensitivity to such qualities in foreign- (rhesus macaque [*Macaca mulatta*]) species voices. Previous research indicates that exposure during a sensitive period is necessary for perceptual narrowing to occur (although one animal study suggests that the end of the sensitive period can be extended in the absence of all relevant stimulation, Sugita, 2008). The critical role of exposure is also revealed by studies that show that sensitivity to distinctions in foreign stimuli can be reinstated (once lost) by providing a fixed period of exposure to the foreign stimulus at the end of the first year (e.g., Kuhl, Tsao & Liu, 2003). In Chapter 4, I investigate whether two weeks of twice-daily exposure to foreign-species (rhesus monkey) voices at home, in the form of a CD-narrated storybook, can reinstate 12-month-olds' sensitivity to distinctions between individual monkey voices. The results support the hypothesis that exposure plays a critical role in the narrowing of voice discrimination abilities during the first year and suggest that the window of plasticity for the processing of foreign voice-types remains open until at least 12 months of age. Discussion focuses on remaining questions and future directions for this line of research, including consideration of the degree to which this window of plasticity might remain open past infancy into childhood and adulthood.

ABSTRACT

Learning to differentiate individuals by their voice is an important social skill for infants to acquire. In a previous study, we demonstrated that the ability to discriminate individuals by their voices follows a pattern of perceptual narrowing (Friendly, et al., submitted). Six-month-olds are able to discriminate between two foreign-species (rhesus monkey) voices nearly as easily as between two native-species (human) voices, but 12-month-olds are much better at human-voice than monkey-voice discrimination, and demonstrate worse ability to discriminate monkey voices than 6-month-old infants. Here we investigate the extent to which the system remains plastic at 12 months after perceptual narrowing has occurred. We found that 12-month-olds who received two weeks of monkey-voice training were significantly better at discriminating between rhesus monkey voices than untrained 12-month-olds. Furthermore, their discrimination was reinstated to a level comparable to 6-month-olds who have not yet fully undergone perceptual narrowing, suggesting that voice-processing abilities remain considerably plastic at the end of the first year.

1. Introduction

Human perception becomes specialized for socially-relevant information in faces, voices, music, and language through a process of *perceptual narrowing,* whereby perception improves for native stimuli experienced in the environment, and becomes worse for foreign stimuli not experienced in the environment (for reviews, see Lewkowicz & Ghazanfar, 2009; Scott et al., 2007). Perceptual narrowing likely contributes to facilitating identification with a group and becoming a fully functioning member of that group. This specialization enables people to discriminate between individuals, identify group and species members, and discern who is from one's own group and who is from outside one's group in order to help inform decisions such as whether to approach or withdraw from a situation. A number of studies indicate that, although perceptual narrowing appears to be accomplished largely by the end of the first year after birth, a certain amount of plasticity remains beyond this age. Here we examine whether a relatively small amount of experience around 12 months of age with voices from a foreign species with which infants have little experience can reinstate the ability to discriminate pairs of voices from that foreign species.

An advantage for processing differences in native compared to foreign stimuli by 12 months of age has been documented across a number of domains, including faces (e.g.,, Kelly et al., 2007; Lewkowicz & Ghazanfar, 2006; Pascalis et al., 2002; Pons et al.,,2009; Simpson et al., 2010), voices (Friendly et al., submitted), music (e.g., Hannon & Trainor, 2007; Hannon & Trehub, 2005a, 2005b; Lynch et al., 1990; Trainor 2005; Trehub & Hannon, 2006), language (e.g., Palmer et al., 2012; Kuhl, 2004; Kuhl et al., 2006; Kuhl et al., 1992; Polka & Werker, 1994; Tsao et al., 2000; Werker & Tees, 1984; for reviews see Curtin & Werker, 2007; Kuhl, 2008, Werker & Tees, 2005), and even action (Loucks & Sommerville, 2012). For example, 6-month-olds are equally good at discriminating two monkey faces as they are at discriminating two human faces, but 9-month-olds and adults are much better with human faces (Pascalis et al., 2002). Similarly, in the language domain, 6-month-old infants are equally good at discriminating consonant speech sounds from two native or two foreign phonemic categories whereas 10- to 12-month-olds are much better with native categories (e.g., Kuhl at al., 2006; Werker & Tees, 1984).

Although less studied than speech and faces, voices are important social stimuli for infants. People can be identified by the unique characteristics of their voices, which is especially useful when visual cues are poor, and voices provide cues to the listener about the size, gender, and age of a talker (e.g., Smith & Patterson, 2005). Recently, Johnson et al. (2011) found that by 7 months of age, infants exhibit an own-language effect when discriminating voices, just as they exhibit an own-race effect when discriminating faces, with better discrimination of voices speaking their native language compared to a foreign language. Subsequently, Friendly et al. (submitted) found that infants demonstrate a similar native-voice processing advantage for voices from their native species, analogous to the specialization infants show for discriminating own-species faces toward the end of their first year after birth. Friendly et al. found that 6-months-olds were able to discriminate among two native-species (human) voices as well as among two foreign-species (rhesus monkey) voices, showing little difference in their discrimination ability for the two voice types. However, by 12 months of age, infants were worse than 6-month-olds at discriminating rhesus monkey voices, and demonstrated a clear advantage for differentiating the human voices.

The important role of experience in perceptual narrowing is indicated by research showing that sensitivity to foreign stimuli can in some cases be *maintained* with exposure to a foreign stimulus during the period when loss typically occurs (e.g., Burns et al., 2003; Pascalis et al., 2005; Scott & Monesson, 2009, 2010) or *reinstated* with exposure to a foreign stimulus after the period of loss (e.g., Anzures, et al., 2012; Hannon & Trehub, 2005b; Kuhl et al., 2003). For example, Pascalis et al. (2005) found that after 2 weeks of daily exposure to monkey faces at 6 months of age, followed by 2.5 months of less frequent exposure, 9-month-olds maintained the ability to discriminate a novel set of monkey faces at a level comparable to that shown at 6 months of age. With respect to reinstatement, 8- to 12-month-olds who received training with other-race faces (Anzures et al., 2012), musical (Hannon & Trehub, 2005b), or linguistic (Kuhl et al., 2003) stimuli foreign to their environment, demonstrated successful processing of those stimuli, whereas their untrained counterparts did not. In the present paper, we investigate the generality of reinstatement by testing whether experience with rhesus monkey voices around 12 months of age can reverse the decrement in ability to discriminate such voices that has been documented in previous research (Friendly et al., submitted).

In designing our training protocol, we considered two features that appear to be important for exposure or training in infancy to result in learning. The first is individual-level encoding (Archambault et al., 1999; Scott & Monesson 2009, 2010; Scott et al., 2006, 2008; Waxman & Braun, 2005). In the visual domain, Scott and Monesson (2009) repeated the Pascalis et al. (2005) monkey-face maintenance study, but used three different types of training regimen. One group of infants was trained at the individual level, similar to the infants in Pascalis et al.'s study, with parents labelling each monkey face with its own unique name (e.g., Dario, Flora, Boris, etc.). A second group of infants was trained at the category level, with parents labelling all monkey faces with the identifier "monkey". The third group of infants was trained through passive viewing alone, with no label being read by parents during exposure. Interestingly, only the infants who received individual-level training with monkey faces showed maintenance of the ability to discriminate novel monkey faces at 9 months of age. Scott and Monesson concluded that individual-level exposure was critical for obtaining the effects found in Pascalis et al.'s study because it focused infants' attention on the features that were unique in each face, rather than on what the monkey faces had in common. Similarly, Anzures et al. (2012) successfully reinstated sensitivity to foreign-race faces in 8- to 10-month-olds after exposing them to daily videos of foreign-race women who introduced themselves by name. However, this was not compared to a no-name condition making it unclear if labelling the faces by name influenced reinstatement in this study. Nevertheless, considering the findings from Scott and Monesson (2009), we designed our training procedure such that particular monkey voices would be associated with unique monkey names and characters.

The second feature that promotes improved learning is social interaction. Kuhl et al. (2003) demonstrated that English-learning infants who were exposed to 12 sessions of Mandarin Chinese training between 9 and 10 months of age showed only a reinstatement in the ability to discriminate between Mandarin-specific phonemes at 10 to 11 months of age if they interacted with the Mandarin speakers in person. Infants who received the same training in the form of an audio-visual video or who received audio-alone training did not show a reverse in the decline of the ability to discriminate Mandarin phonemes. Similar benefits of social interaction have been found for native-language training using Baby Einstein© videos, where infants required parental interaction in order to learn the words that were featured in the videos (Richert et al., 2010). Likewise, active musical interaction between infants and parents has been found to lead to earlier musical pitch enculturation compared to passive music listening (Gerry et al., 2012; Trainor et al., 2012). On the other hand, Anzures et al. (2012) found evidence of reinstatement, at 8 to 10 months of age, of sensitivity to foreign-race faces after 3 weeks of daily exposure to audio-visual videos of other-race females and Hannon and Trehub (2005b) found reinstatement of sensitivity to foreign musical rhythms in 12-month-old infants who listened passively to these rhythms at home on a CD. However, while parents were instructed to avoid drawing the infants' attention to the music and instead to go about their regular routines, the CD may have been played during some type of social interaction between the parent and infant.

In the present study, we gave 11.5-month-old infants specific exposure to rhesus monkey voices under conditions that promoted individual-level encoding in the context of social interaction. In particular, we designed a storybook and accompanying audio CD narration that parents listened to with their infants twice a day for a 2-week period. The storybook contained a number of exemplars of each of four monkey characters' voices. Following this exposure, we tested infants' ability to discriminate a new set of monkey voices. We compared their discrimination of monkey voices to that of the 12- and 6-month-olds in Friendly et al., submitted) who did not receive any training. We aimed to determine whether this exposure would reinstate the ability to discriminate the monkey voices to the original level found at 6 months of age, before perceptual narrowing was fully underway

2. Materials and methods

2.1 Participants

Twenty-four infants (mean age = 12.0 months, SD = 0.19 months at the time of testing; 10 females) received two weeks of monkey-voice training prior to testing (Trained-12 month Group). They were compared to two groups of infants from our previous study that were tested in the identical procedure, but who did not receive any training (Friendly et al., submitted). One group was also 12 months of age (n = 24, mean age = 12.0 months, SD = 0.19 months; 9 females; Untrained-12 month Group) and the other group was 6 months of age (n = 24; mean age = 6.1 months, SD = 0.22 months, 11 females; Untrained-6 month Group). Parents gave informed consent and reported normal hearing for all infants. Parents also reported all infants as hearing English 98-100% of the time in their home environment. An additional 20 infants across the three groups were excluded from the final sample due to fussiness (n = 4), failure to pass training (see Procedure below; n = 7), receiving less than 23 training sessions (n = 5), being too old at time of testing (n = 1) and hearing non-English languages more than 2% of time in their home environment, as reported by parents (n = 3).

2.2 Stimuli and Apparatus

2.2.1 Training Stimuli. A CD-narrated picture storybook was created for the monkey voice training. Entitled “Beach Day for the Monkey Family”, it contained colourful illustrations of four members of the Monkey Family going out for a day at the beach. Each monkey (labelled as Daddy, Mommy, Sister and Brother Monkey) was shown individually on a separate page, in consecutive order, six times throughout the storybook (for an example, see Figure 1). The CD-narration for the storybook was read by a monolingual English-speaking adult female and spoken in an infant-directed manner. Parents were instructed to listen to the accompanying CD and turn the page only at the sound of the chime, which occurred 4 seconds after the last vocalization on each page. The CD was designed so that every time one monkey was being viewed in the storybook, infants heard two vocalizations produced by a real rhesus macaque (*Macaca mulatta*). Thus, on the CD, each monkey character in the storybook was associated with the vocalizations of only one particular rhesus monkey. Twelve rhesus voice recordings were heard for each of the four rhesus monkey characters on the CD (6 tokens of the “coo” call category, heard two times each). The rhesus monkey recordings were obtained from author DR (for methodology on obtaining these recordings see Owren & Rendall, 2003; Rendall et al., 1996), edited using Cool Edit Pro (Syntrillium Software; sampling rate = 44.1 kHz, [intensity] resolution = 16-bit), and normalized for peak intensity across the sample. On the CD, the recordings of each monkey were ordered randomly, with the stipulation that the same "coo" token was never presented twice in a row and that the two tokens heard on each page formed a unique pair. The four rhesus monkeys on the CD were different from the monkeys used for testing. Those on the CD formed two matched pairs (Daddy/Mommy monkey; Sister/Brother monkey) such that the set of tokens for each voice in the pair were matched for mean duration (mean = .367 s, SD = .065 s; mean = .477 s, SD = .080 s, for the two pairs, respectively) and minimum (mean = 218.78 Hz, SD = 116.45 Hz; mean = 469.45 Hz, SD = 24.78 Hz), maximum (mean = 458.48 Hz, SD = 70.31 Hz; mean = 535.57 Hz, SD = 20.50 Hz) and mean (mean = 340.12 Hz, SD = 86.48 Hz; mean = 514.46 Hz, SD = 15.55 Hz) F0 (analysed using Pratt software`s autocorrelation algorithm, F0 searched for between 100-600 Hz; Boersma & Weenink, 2009). An excerpt from the training storybook and CD (full version: 6 min 40 s in duration) can be found at: http://psycserv.mcmaster.ca/ljt/RSM/RSM.html.

2.2.2 Testing Stimuli. The rhesus monkey vocalizations used for testing were identical to those used in our previous study (see Friendly et al., submitted), but different from those used during training. Six vocal samples of the "coo" call from each of four female rhesus monkeys were obtained from author DR (see Owren & Rendall, 2003; Rendall et al., 1996), edited using Cool Edit Pro (sampling rate = 44.1 kHz, [intensity] resolution = 16-bit), and normalized for peak amplitude across the sample (Pair 1: voice 1 mean = 49 dB, range = 46-54 dB, voice 2 mean = 49 dB, range = 47-51 dB; Pair 2: voice 1 mean = 55 dB, range = 53-57 dB, voice 2 mean = 56 dB, range = 55-57 dB).

Two pairs of primate voices (six different "coo" call tokens for each individual monkey) were paired based on acoustic analyses using Praat software, such that their sets of tokens were matched for mean duration (mean = .302 s, SD = .047 s; mean = .274 s, SD = .061 s, for the two pairs, respectively) and minimum (mean = 282.36 Hz, SD = 66.28 Hz; mean = 502.48 Hz, SD = 18.93 Hz), maximum (mean = 351.84 Hz, SD = 49.87 Hz; mean = 562.59 Hz, SD = 32.49 Hz), and mean (mean = 320.50 Hz, SD = 48.72 Hz; mean = 542.04 Hz, SD = 26.30 Hz) F0. Four conditions (1A, 1B, 2A and 2B) were created for testing infants so that, for each voice pair (1 and 2), one voice in the pair served as the "change" voice and the other as the "background" voice for condition A (see Procedure). The change and background voices were switched for condition B.

2.3 Procedure

2.3.1 Training procedure. Two weeks prior to testing, infants in the Trained Group were mailed a package containing the illustrated storybook and accompanying CD. The package also contained a music and language questionnaire, daily reading log, and instructions for the infant's training schedule. Parents were instructed to play the CD twice a day at home for two weeks, for a total of 28 training sessions, following along in the storybook with their infant. In order to ensure that the voice of each rhesus monkey on the CD was associated with a particular character in the book (labeled either Daddy, Mommy, Sister, or Brother Monkey), parents were instructed to turn each page of the storybook only when they heard a musical chime sound. To make sure that infants were engaged actively during the monkey-voice training, parents were instructed to listen to the storybook together with their infant, interacting to engage their infant's attention as much as possible. Infants were reported to have received between 23 and 29 sessions of training (mean = 28 sessions). One day after the completion of the two-week training period, infants were brought into the lab for testing.

2.3.2 Testing Procedure. Infants in the Trained Group were tested in the identical Conditioned Head Turn (CHT) procedure as infants in the Untrained Groups from Friendly et al. (submitted) (also see Werker et al., 1998). Infants were assigned randomly to one of four stimulus conditions (1A, 1B, 2A or 2B), where the A and B conditions reversed which voice of the pair was the background and which the change voice.

During the testing phase of the CHT procedure, a loudspeaker located 90^{o} to the infant's left played the six *"coo*" tokens from the background voice repetitively in a quasi-random order such that the same token was never repeated consecutively (stimulus onset asynchrony = 2 seconds). The parent sat across from the experimenter with the infant seated on his/her lap and listened to masking music through headphones in order to eliminate potential parental influence on the infant's behaviour. The experimenter likewise listened to masking music during testing. Throughout the experiment, tokens from the background voice were played continuously. The experimenter pressed one button when the infant was paying attention and facing forward (toward the experimenter), indicating to the computer that the infant was ready for a trial. There were 24 trials. Half (12) were control (no-change) trials that were indistinguishable from the repeating background. The other half (12) were change trials, on which the background voice was replaced by one of the six tokens of the changed voice for one repetition. Across the 12 trials, each of the six change-voice tokens was presented twice in a random order. The order of change and control trials was quasi-random, with the constraint that no more than two control trials were presented in a row. The experimenter pressed a second button when the infant made a head-turn response of 45 degrees or more to the left towards the speaker from which the sounds were played. Head turn responses occurring on control trials (i.e., false alarms) were not rewarded by the computer. In contrast, head turns on change trials (i.e., hits) that occurred within 1.5s of the onset of the changed voice were rewarded by the computer with 2s of an animated light and toy display. The proportion of hits and false alarms was converted into d' scores for data analysis.

Before infants began the testing phase of the CHT procedure, they first had to pass an initial familiarization phase, which involved a training session designed to familiarize them with the rule that when they made a head-turn response to a change from one monkey's voice to another, they would be rewarded with an animated toy display. In this phase, only two of the six change-voice tokens were used and there were no control trials. Furthermore, during familiarization the change voice was played, on average, 8 dB louder than the repeating background voice (see Stimuli for testing phase levels) in order to make it a noticeable difference that would attract the infant's attention to look toward the loudspeaker. In order to pass the familiarization phase and proceed to the testing phase, infants were required to make four correct head-turn responses in a row within 20 training trials. Once in the testing phase, all six 'coo' tokens of the change and background voices were presented at equal intensities. Infants who did not pass this familiarization criterion were excluded from the final data set (see Participants).

3. Results

As can be seen in Figure 2, infants in the Trained-12 month group performed quite well at discriminating the monkey voices. Although all three groups performed significantly above chance levels (Untrained-12 month: t(23) =

3.53, p = .002**;** Untrained-6 month: t(23) = 9.70, p < .001; Trained-12 month: t(23) = 8.69, p < .001), a one-way ANOVA with group indicated a significant difference across groups in d' scores, F(2, 69) = 8.09, p = .001. A follow-up independent samples t-test indicated that infants in the Trained-12 month Group performed significantly better than age-matched infants in the Untrained-12 month Group, *t*(46) = 3.62, *p* = .001, Cohen's *d* = 1.07. As well, a comparison of infants in the Trained-12 month Group to the younger infants who had not yet achieved perceptual narrowing (Untrained-6 month Group) indicated no significant difference between these groups, *t*(46) = 1.49, *p* = .14, Cohen's *d* = 0.44. These findings suggest that two weeks of exposure to monkey voices at 11.5- 12 months of age (after perceptual narrowing has occurred) can reinstate sensitivity to voices from a foreign species to a 6-month-old level.

4. Discussion

In previous work, we showed that the development of specialization for own-species voice discrimination follows a pattern of perceptual narrowing, with a decrease in infants' ability to individuate rhesus monkey voices between 6 and 12 months of age (Friendly et al., submitted). In the present paper, we found that after 2 weeks of twice-daily exposure to rhesus monkey voices in the form of a CD-narrated storybook, 12-month-old infants demonstrated significantly better discrimination of novel rhesus monkey voices not heard during training compared to 12-month-old infants who received no such training. Furthermore, the performance of the 12-month-olds who received exposure to the monkey voices was not significantly different from that of 6-month-old infants. The fact that monkey voice exposure enhanced discrimination of monkey voices in 12-month-olds to a level seen in 6-month-old infants indicates that the processes underlying perceptual narrowing for voice identification retain considerable plasticity at least until 12 months of age. This conclusion is consistent with studies in other domains that indicate that exposure to a socially-relevant foreign stimulus can either maintain (Burns et al., 2003; Pascalis et al., 2005; Scott & Monesson, 2009, 2010) or reinstate (Anzures et al., 2012; Hannon & Trehub, 2005b; Kuhl et al., 2003) sensitivity past the period during which perceptual narrowing normally occurs.

Previous studies on face processing suggest that maintenance of sensitivity for individuating faces from a foreign species requires that the exposure to those faces be at an individual level, with different labels, such as names, being applied to the different faces in the exposure set (e.g., Scott & Monesson, 2009; Scott et al., 2006). Those studies found that having infants simply observe foreign-species faces, or experience them in the context of a common label applied to all faces (e.g., “monkey”), does not lead to maintenance of the ability to discriminate these faces. As well, Anzures et al. (2012) found that exposure to labelled foreign-race faces at 8- to 10-months reinstates infants' ability to recognize foreign-race faces, although they did not test under conditions with no labelling. Scott and Monesson (2009) suggest that labelling faces by name might draw infants' attention to the differences between individuals, rather than to what the individuals have in common. On the other hand, phoneme categories in speech and metrical structures in music do not apply to individual people or individuals from other species (although changes in phonemes can signal changes in word meaning), and explicit individuation though the use of labels in these cases does not seem to be necessary for exposure to foreign categories or structures to disrupt perceptual narrowing (Hannon & Trehub, 2005b; Kuhl et al., 2003). The case of distinguishing individuals by their voice would seem to be similar to the case of distinguishing individuals by their face, suggesting that attention to differences between individuals during exposure might be critical for reinstating sensitivity to voices from a foreign species. This could be tested in future research on distinguishing voices by using a training protocol that either applies no label at all or that applies the label "monkey" to each rhesus monkey voice sample during training.

Social interaction also has been identified as important for plasticity during the period of perceptual narrowing (Kuhl et al., 2003; Gerry et al., 2012). On the other hand, Hannon and Trehub (2005b) found that 2 weeks of twice-daily passive exposure to foreign musical rhythms was sufficient to reinstate 12-month-olds' ability to detect violations in foreign rhythmic structure. In the present study, infants receive training in the social context of parental interaction. It is possible that passive exposure to foreign stimuli can affect perceptual narrowing, but that exposure in a social context is more powerful. It remains for future research to determine whether the social interaction during the training phase of the present study was a necessary condition for the reinstatement of sensitivity to monkey voices.

In interpreting the effects of different kinds of experience on reinstatement of abilities at 12 month of age, it is important to consider the results from a study by Fair et al. (2012), in which sensitivity to distinctions in foreign-species (monkey) faces at 12 months was observed without a prescribed at-home training period, by simply extending the length of the familiarization period at the time of testing. Specifically, Fair et al. demonstrated that 12-month-olds showed no evidence of discriminating unfamiliar monkey-face after 20s of familiarization, but did discriminate them after 40s of familiarization. This extended familiarization period could be considered a relatively brief form of training (although with only one face), but it is surprising that reinstatement could be achieved after such a brief training.

A final question concerns the age range over which the window of plasticity remains open with respect to learning to discriminate foreign voices. In other domains, there is evidence that some plasticity remains throughout the lifespan in that exposure to foreign stimuli in childhood (e.g., Feinman & Entwisle, 1976; Macchi Cassia et al., 2009; Sangrigoli et al., 2005; Wang & Kuhl, 2003) or adulthood (e.g., Bradlow et al., 1999; de Heering & Rossion, 2008; Iverson et al., 2005; Kuefner et al., 2008; McCandliss et al., 2002; Pruitt et al., 2006; Scott et al., 2006, 2008; Tees & Werker, 1984; Zhang et al., 2009) results in improved processing of those stimuli, particularly if the training contains highly variable and numerous stimuli (Zhang et al., 2009), if differences between the stimuli are exaggerated (McCandliss et al., 2002), or if the person had some exposure to the stimuli earlier in life (e.g., Lenneberg, 1967; Macchi Cassia et al., 2009; Newport et al., 2001; Oh et al., 2010; Sangrigoli et al., 2005; Tees & Werker, 1984). However, completely native-like processing of foreign stimuli appears to be very difficult, if not impossible, to achieve in adulthood (e.g., Flege et al., 1999; Hannon & Trehub, 2005b; Iverson et al., 2005; McCandliss et al., 2002; Takagi, 2002; Takagi & Mann, 1995, for reviews see Birdsong, 2006; Hernandez & Li, 2007). Nevertheless, Sangrigoli, et al. (2005) found native-like discrimination of Caucasian faces by Korean adults who were adopted by French families between 3 to 9 years of age, suggesting that it is possible to demonstrate native levels of foreign-race face processing under certain conditions. In the present study, informal feedback from parents who participated with their infants suggests that parents had a difficult time distinguishing between the four monkey voices used in the storybook, even after listening to the story with their infant for the 2-week, twice-daily period. However, it is possible that with sufficient training, adults could become proficient at discriminating foreign voices, as their ability to discriminate the monkey voices is poor but above chance levels without training. Again, this could be investigated in a future study.

In summary, perceptual narrowing achieved by the end of the first year after birth for discriminating voices can be modified by two weeks of exposure to voices from a foreign species, indicating a period of flexibility and plasticity following narrowing. It remains for future research to determine the time course of this plasticity across the lifespan, the effects of the social context, and whether individual-level training is important in perceptual narrowing for voice discrimination.

Acknowledgements

This research was supported by a grant to LJT from the Natural Sciences and Engineering Research Council of Canada. We thank Ren Ee Choo, Aliyah Mohamed and Andrea Unrau for assistance in data collection.

References

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology*, *112,* 484-495.

Archambault, A., O'Donnell, C. & Schyns, P. G. (1999). Blind to object changes: When learning the same object at different levels of categorization modifies its perception. *Psychological Science 10*(3), 249-255.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*, 9-49.

Boersma, P., & Weenink, D. (2009). *Praat: Doing Phonetics by Computer Version 5.1.04*. Computer program. Available from http://www.praat.org/

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and / l/: Long-term retention of learning in speech perception and production. *Perception & Psychophysics*, *61*(5), 977-985.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. In B. Beachley, A. Brown & F. Conlin (Eds.), *Proceedings of the 27th annual Boston university conference of language development* (pp. 173-

184). Cascadilla, Sommerville, MA.

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. In M. G. Gaskell (Ed.), *The oxford handbook of psycholinguistics*. Oxford, UK: Oxford University Press.

de Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLoS ONE, 3*(5), e2317.

Fair, J., Flom, R., Jones, J., & Martin, J. (2012). Perceptual Learning: 12-Month-Olds' Discrimination of Monkey Faces. *Child Development* *83*(6), 1996-2006.

Feinman, S. & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development, 47,* 506-510.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.

Friendly, R. H., Rendall, D., & Trainor, L. J. (submitted). Learning to differentiate individuals by their voices: Infants' individuation of native- and foreign- species voices. *Developmental Psychobiology: Special Issue on Perpetual Narrowing.*

Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental Science, 15*(3), 398-407.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*(11), 466-472.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48-55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, USA, 102*(35), 12639-12643.

Hernandez, A., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638-650.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*(5), 1002-1011.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy. *Psychological Science, 18*(12), 1084-1089.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of*

Experimental Psychology: Human Perception and Performance, 34(4), 811-817.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

Kuhl, P. K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research directions* (pp. 213-243). New York: Erlbaum.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*(2), 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, USA, 100*(15), 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*(5044), 606-608. Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley & Sons, Inc.,

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences, USA, 103*(17), 6771-6774.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13*(11), 470-478.

Loucks, J., & Sommerville, J. A. (2012). Developmental changes in the discrimination of dynamic human actions in infancy. *Developmental Science, 15*(1), 123-130.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1*(4), 272-276.

Macchi Cassia, V., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing: Evidence for the reactivation of dormant effects. *Psychological Science, 20*(7), 853-859.

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: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective & Behavioral Neuroscience, 2*(2), 89-108.

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Doupoux (Ed.), *Language, brain and cognitive development: Essays* *in honor of Jacques Mehler* (pp. 481- 502). Cambridge, MA: MIT Press.

Oh, J. S., Au, T. K. F., & Jun, S. A. (2010). Early childhood language memory in the speech of international adoptees. *Journal of Child Language, 37*(5), 1123-1132.

Owren, M. J., & Rendall, D. (2003). Salience of caller identity in rhesus monkey (macaca mulatta) coos and screams: Perceptual experiments with human (homo sapiens) listeners. *Journal of Comparative Psychology, 117*(4), 380-390.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development, 83*(2), 543-53.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science, 296*(5571), 1321-1323.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences, USA, 102*(14), 5297-5300.

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

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009).

Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences, USA, 106*(26), 10598-10602.

Pruitt, J. S., Jenkins, J. J., & Strange, W. (2006). Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *The Journal of the Acoustical Society of America, 119*(3), 1684-1696.

Rendall, D., Rodman, P. S., & Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behavior, 51*(5), 1007-1015.

Richert, R. A., Robb, M. B., Fender, J. G., & Wartella, E. (2010). Word learning from baby videos. *Archives of Pediatrics and Adolescent Medicine, 164*, E1-E5.

Sangrigoli, S., Pallier, C., Argenti, A., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*(6), 440-444.

Scott, L. S., & Monesson, A. (2009). The origins of biases in face perception. *Psychological Science, 20*(6), 676-680.

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. *Neuropsychologia, 48*, 1857-1861.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in*

Psychological Science, 16(4), 197-201.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience, 18*(9), 1453-1465.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L. & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: behavioral and neurophysiological study. *Brain Research, 1210*, 204-215.

Simpson, E. A., Varga, K., Frick, J. E., & Fragaszy, D. (2010). Infants experience perceptual narrowing for nonprimate faces. *Infancy, 16*(3), 318-328.

Smith, D. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgments of speaker size, sex, and age. *Journal of the Acoustical Society of America, 118*(5), 3177-3186.

Sugita, Y. (2008) Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences, USA, 105*(1), 394-398.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*(6), 2887-2896.

Takagi, N., & Mann, V. A. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics, 16*(4), 379-405.

Tees, R. C., & Werker, J. F. (1984). Perpetual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 34*(4), 579-590.

Trainor, L. J. (2005). Are there critical periods for music development? *Developmental Psychobiology, 46,* 262-278.

Trainor, L. J., Marie, C., Gerry, D., Whiskin, E., & Unrau, A. (2012). Becoming musically enculturated: Effects of music classes for infants on brain and behavior. *Annals of the New York Academy of Sciences, 1252,* 129-138.

Trehub, S. E. & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73-99.

Tsao, F., Liu, H., Kuhl, P. K., & Tseng, C. (2000). Perceptual discrimination of a Mandarin fricative-affricate contrast by English-learning and Mandarin-learning infants. *Poster presented at the meeting of the International Society on Infant Studies*, Brighton, England.

Wang, Y., & Kuhl, P. K. (2003). Evaluating the "critical period" hypothesis: Perceptual learning of Mandarin tones in American adults and American children at 6, 10 and 14 years of age. *15th ICPhS,* Barcelona. 1537-1540.

Waxman, S. R. & Braun, I. E. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*, B59-B68.

Werker, J. F., Shi, R., Desjardins, R., Pegg, J. E., Polka, L., & Patterson, M. (1998). Three methods for testing infant speech perception. In A. M. Slater

(Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy* (pp. 389-420). London: UCL Press.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 79*(1), 49-63.

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*(3), 233-251.

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., Kawakatsu, M., Tohkura, Y., & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage, 46*, 226-240.

FIGURES

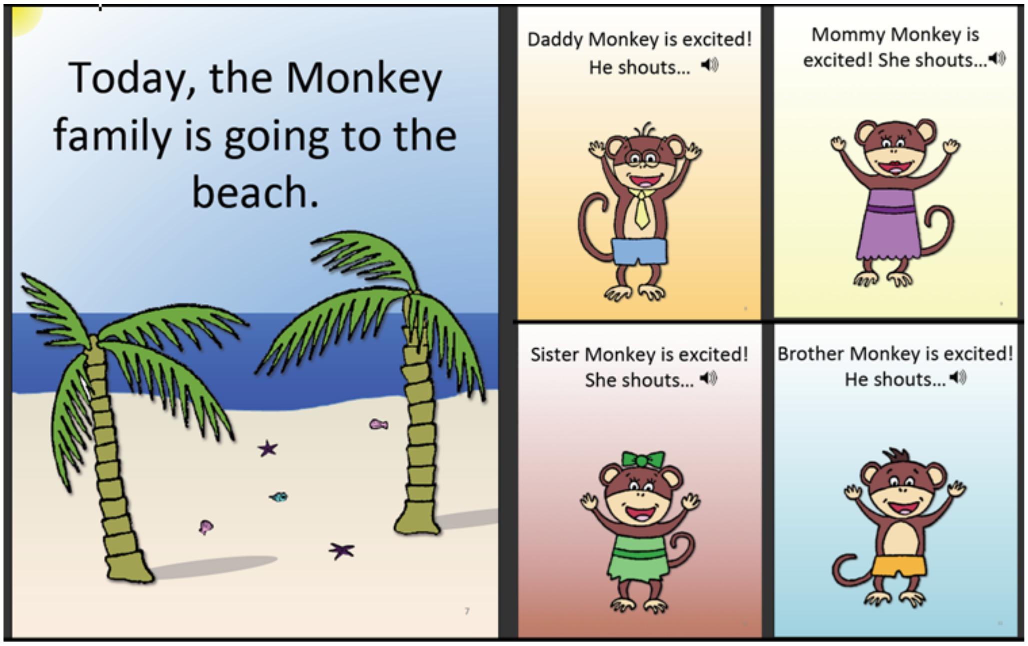

Figure 1. An excerpt from the training stimulus storybook “Beach Day for the Monkey Family”. A CD that contained the story narration and vocalizations belonging to four individual rhesus monkeys (six different coo tokens per monkey) accompanied the storybook. Each rhesus voice was always associated with the same monkey character in the story (either Daddy, Mommy, Sister, or Brother Monkey), and the monkey characters were presented in the same order six times throughout the book. The full storybook and CD can be found at: http://psycserv.mcmaster.ca/ljt/RSM/RSM.html

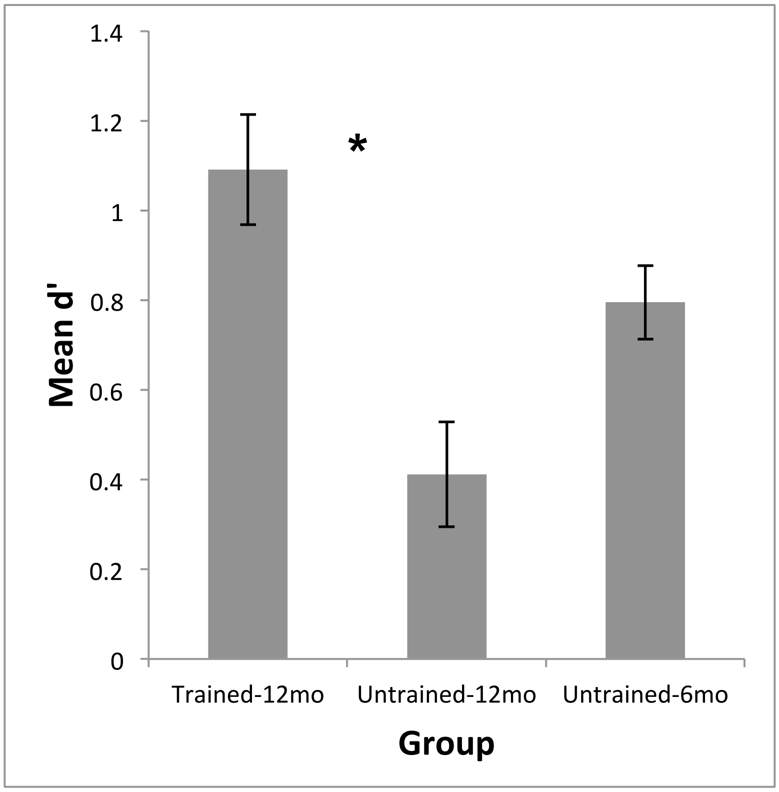

Figure 2. Rhesus monkey voice discrimination abilities of 12-month-old infants (n = 24) who received monkey-voice training (Trained-12 month) in the present study compared to 6- (Untrained-6 month) and 12-month-old (Untrained-12 month) infants (n = 24 each) who did not receive training and were tested in a previous study (Friendly et al., submitted). Discrimination was significantly better for trained than untrained 12-month-old infants (p = .001). There was no significant difference between trained 12-month-old and untrained 6-month-old infants, suggesting that two weeks of exposure to monkey voices at 12 months reinstates processing of foreign-species voices to a level observed at 6 months, before perceptual narrowing is fully underway. Error bars represent SEM.

CHAPTER 5: General Discussion

Thesis Findings and Unique Contributions

In the present thesis, I examined the development of voice discrimination abilities during the first year after birth. In Chapter 2, I investigated whether the ability to discriminate voices might develop through a process of perceptual narrowing, with infants becoming specialized for processing differences between voices in their native species by the end of their first year, similar to the perceptual narrowing previously observed for other socially-relevant stimuli from auditory (e.g., language, music) and visual (e.g., faces) domains (for reviews see Lewkowicz & Ghazanfar, 2009; Scott, Pascalis & Nelson, 2007). Although a previous study found evidence that 7-month-old infants demonstrate language-specificity when individuating voices speaking a foreign versus their native language (Johnson, Westrek, Nazzi, & Cutler, 2011), results of the present thesis are the first to reveal evidence of perceptual narrowing for the unique characteristics of human voices, with specialization developing between 6 and 12 months of age for human compared to rhesus macaque (*Macaca mulatta*) voices. Specifically, at 6 months infants are similar in their abilities to discriminate between two human voices and between two monkey voices whereas at 12 months infants are much better at discriminating human than monkey voices.

In Chapter 3 of the present thesis, I examined whether the processing advantage for human voices observed for adults in Chapter 2 was driven primarily by participants' familiarity with the language spoken (English) or with the species type (human) of the vocalizations. Since the human vocalizations used in Chapter 2 consisted of females uttering the English word "balloon", and given that specialization for voices speaking the native language has been demonstrated in 7-month-olds (Johnson et al., 2011), it was unclear to what extent the results of the adults in Chapter 2 reflected specialization for the unique characteristics of human voices in general, or for the English language in particular. To examine this, I investigated whether an advantage for processing these vocalizations would be evident in an adult Chinese population for which English was a foreign, rarely heard language. The findings revealed that Mandarin-speaking adults, who had little or no exposure to the English language, did exhibit an own-species effect, specifically demonstrating much better discrimination of the human (English) voices than the rhesus monkey voices. At the same time, however, the English-speaking adults from Chapter 2 performed significantly better at discriminating the human (English) vocalizations than did the Mandarin-speaking adults tested in Chapter 3, which is in line with a few previous studies that have found own-language effects operating during the individuation of voices in adulthood (e.g., Goggin, Thompson, Sturbe & Simental, 1991; Köster, & Schiller, 1997). However, there were also differences between the populations in terms of education level, so it is not clear whether these results reflect the operation of an other-language effect. In any case, the important result is that the data clearly show the operation of an other-species effect in adults and support that idea of specialization for discrimination of voices with characteristics of the human vocal tract.

In Chapter 4 of the present thesis, I described the first study to explore the influence of exposure to foreign-voices after perceptual narrowing has occurred at the end of the first year after birth. Previous studies provide evidence that experience plays a key role in the process of perceptual narrowing. For example, exposure to foreign consonant contrasts (Kuhl, Tsao, & Liu, 2003), other-race faces (Anzures et al., 2012) and musical rhythms (Hannon & Trehub, 2005b) at the end of the first year can *reinstate* perceptual sensitivity that was previous dampened during the narrowing process. In Chapter 4, I investigated whether a similar reinstatement of sensitivity to distinctions in foreign-species voices would be observed with exposure to these voices at the end of the first year after birth. To test this, I compared the discrimination of foreign-species (rhesus monkey) voices by 12-month-olds who received two weeks of twice-daily training (with a different set of rhesus monkey voices than those used during testing) with the 12-month-olds from Chapter 2, who received no such training. The findings revealed superior discrimination of monkey voices by the trained 12-month-olds compared to the untrained 12-month-old infants. Furthermore, there is evidence that trained

12-month-olds were able to discriminate the monkey voices at least as well as 6-month-old infants. These findings suggest that exposure to foreign-species' voices can reinstate sensitivity to these voices once perceptual narrowing has occurred. The findings are in line with the previous research demonstrating reinstatement of sensitivity to foreign auditory stimuli (phonemes, musical rhythms) (Hannon & Trehub, 2005b; Kuhl et al., 2003), and foreign-races faces (Anzures et al, 2012). They are also in line with research demonstrating that sensitivity to distinctions in individual monkey faces can be maintained with exposure to monkey faces labelled at an individual-level using distinct names (Pascalis et al., 2005, Scott & Monesson, 2009). An important implication of this study is that the window of plasticity for sensitivity to foreign-species voices appears to remain open until at least 12 months of age.

Further Questions Related to Perceptual Narrowing

One question for further study is whether the window of plasticity for specialization of voice discrimination remains open after one year of age. The ability to distinguish foreign stimuli has been demonstrated to depend on the timing and type of exposure received. In infancy, experience with foreign-species faces during the perceptual narrowing period can *maintain* sensitivity to foreign-species faces (Pascalis et al., 2005), and experience just after the narrowing period can *reinstate* sensitivity to foreign-race faces, foreign consonant contrasts and foreign musical rhythms (Anzures et al., 2012; Hannon & Trehub, 2005b; Kuhl et al., 2003) to levels seen prior to narrowing. After the first year, the effects of exposure appear to decrease with age (e.g., Flege, Yeni-Komshian, & Liu, 1999; Hannon & Trehub, 2005b). Although exposure to foreign stimuli in childhood (e.g., Macchi Cassia, Kuefner, Picozzi & Vescovo, 2009; Sangrigoli, 2005) and adulthood (e.g., Iverson, Hazan & Bannister, 2005; Kuefner, Macchi Cassia, Picozzi & Bricolo, 2008) increases an individual's ability to discriminate these foreign stimuli, adults generally never learn to process them at a native-like level (e.g., Hannon & Trehub, 2005b; for reviews see Birdsong, 2006; Hernandez & Li, 2007). Nevertheless, evidence suggests that training is more effective for adults who were exposed to foreign faces (Macchi Cassia et al., 2009; Sangrigoli et al., 2005) and languages (e.g., Oh, Au & Jun, 2010; Tees & Werker, 1984) during infancy and/or childhood, but not after, compared to adults who were never exposed previously to the foreign stimuli.

Despite this evidence for diminishing plasticity with age, in some cases, considerable plasticity appears to remain throughout the lifespan. For example, preschool teachers, who receive extensive exposure to children's faces, do not demonstrate the *other age effect* (i.e., poorer discrimination of children's than adults' faces), which is observed for most adults (de Heering & Rossion, 2008). Furthermore, if exposure to foreign faces is continuous from childhood to adulthood, and there is no exposure to native faces during this period, then processing of foreign faces can reach native-like levels. Sangrigoli, Pallier, Argenti, Ventureyra, and de Schonen (2005) found that Korean adults, who were adopted by Caucasian (French) families between 3 to 9 years of age, could discriminate successfully among Caucasian (French) faces at a level akin to that observed in French adults, and better than the level of Korean adults who moved to France as adults. With respect to non-human primates, rhesus monkeys were able to match the faces and voices of the human caregivers and researches they were familiar with (Sliwa, Duhamel, Pascalis & Wirth, 2011). However, the ability to process foreign stimulus distinctions at a native-like level can come at the expense of the ability to process distinctions in native stimuli, at least for linguistic processing. After their adoption by French-speaking families, Korean adults who have no exposure to their first language (Korean), lose the processing advantages for their first language, and do not regain much advantage following exposure (Pallier at al., 2003; Ventureyra, Pallier & Yoo, 2004).

With respect to voice discrimination, it remains for future research to investigate the extent to which exposure to foreign-species voices can improve the ability of human children and adults to discriminate foreign-species voices, and whether the effectiveness of training decreases with age. Such investigations will shed light on whether there are sensitive periods during which exposure to foreign voice types has particular impact, and when during the lifespan these sensitive periods may occur. For instance, if it is found that exposure does not lead to improvements in children's and adults' ability to discriminate foreign-species voices, or if the improvements found are significantly less than those shown at 12 months of age in Chapter 4 of the present thesis, then this would suggest that the sensitive period is restricted to the end of the first year after birth, at least for sensitivity to rhesus monkey voices. On the other hand, if children or adults show levels of improvement following exposure similar to those found in 12-month-olds, then this would suggest that the sensitive period might extend into childhood or adulthood, respectively.

Another remaining question in the literature on perceptual narrowing concerns the timing of sensitive periods in the absence of all stimulation. In other words, does the timing of perceptual narrowing remain constant, regardless of experience with native stimuli, or does exposure to native stimuli affect the timing of perceptual narrowing? To date, there has been only one highly controlled investigation of this question. Sugita (2008) examined the development of face processing in monkeys that were reared without exposure to any faces from birth for periods ranging from 6 to 24 months. During the deprivation period, the monkeys showed evidence of equal ability for discriminating photographs of human and monkey faces. Following facial deprivation, monkeys were exposed to either human or monkey faces (in the form of interaction with individuals) for a month, after which the monkeys were successfully able to discriminate photographs only of the face type to which they had been exposed. That study demonstrates that, in the absence of all experience with faces, the sensitive period for face specialization can be extended, and perceptual narrowing for face discrimination can be postponed, to at least 2 years of age in monkeys. The answer to this question remains unknown for voices.

A further question concerns the neural mechanism(s) by which perceptual narrowing takes place. Perceptual narrowing has been proposed to be a domain-general process through which the brain is prepared to process distinctions that are critical in the individual's environment, since this phenomenon has been documented across multiple socially-relevant stimuli in the auditory and visual domains (e.g., Black, Jones, Nelson & Greenough, 1998; see Scott et al., 2007 for a review). Currently, however, perceptual narrowing has not been demonstrated in the olfactory, gustatory, or tactile domains, which leaves question as to whether narrowing is in fact domain-general, or if it is specific to the auditory and visual domains. If narrowing is indeed domain-general, and if it follows a similar time frame across domains, then it is expected that young infants initially would be sensitive to important distinctions in the smells, tastes and tactile feel of stimuli from both their native and from foreign environments, but that this sensitivity would become specific to native stimuli by the end of the first year. Thus, future research is needed to investigate whether perceptual narrowing occurs, and if it occurs during the first year after birth, for domains other than auditory and visual.

A number of neural mechanisms have been proposed to underlie the narrowing process. There is currently little definitive evidence that perceptual narrowing results from any of these proposed mechanisms and it should be noted that they are not necessarily mutually exclusive. One popular model posits that perceptual narrowing occurs as a result of *pruning* (or elimination) of an initial over-abundance of neuronal connections in the brain (Changeux & Danchin, 1976; Cowan, Fawcett, O'Leary & Stanfield, 1984; Goldman-Rakic, 1987; Huttenlocher, 1979; Low & Chen, 2006; Scott et al., 2007; Webb, Monk, & Nelson, 2001). In other words, this model suggests that young infants have a large number of neural connections, but that the connections between neurons that receive relatively little or no concurrent stimulation are selectively eliminated as the brain develops, while those that do receive concurrent stimulation are strengthened. A second model of neural plasticity suggests that perceptual narrowing results from elaboration (or the creation of new neural connections), rather than from the pruning of relatively weakly stimulated connections (see Lewkowicz & Ghazanfar, 2009, Box 2). Evidence supporting the second model comes from neural studies of non-human animals showing that, for example, regions of the rodent and ferret brains (e.g., the somatosensory and visual cortices, respectively) lack profuse neural connections early in development (Agmon, Yang, O'Dowd & Jones, 1993; Quartz & Sejnowski, 1997), and that the rhesus monkey brain continues to increase in terms of the number of neural connections for a relatively long period after birth (Bourgeois & Rakic, 1993; Purves, White & Riddle, 1996). However, the models are not mutually exclusive, as pruning of unused neural connections has been shown to follow elaboration or strengthening of used connection during development across a wide range of mammalian species, including humans (e.g., Bourgeois, 1997; Sowell et al., 2003).

A third mechanism suggested to facilitate perceptual narrowing is a shift in brain regions utilized for perceptual discrimination during infancy. Evidence for this shift comes from a study investigating activations in the human infant brain during the processing of emotion in intersensory face/voice stimuli (Vogel, Monesson & Scott, 2012). Vogel et al. found that between 5 and 9 months of age, intersensory processing narrows, becoming specialized for native-race emotional stimuli. In addition, between 5 and 9 months, there is a shift, for the most part, in the neural networks activated in response to incongruous emotional face-voice congruency in native- and foreign-race stimuli, from an anterior event-related potential (ERP) component associated with attention to a posterior ERP component associated with perception. Additionally, there is evidence from brain imaging studies with adults of stronger brain activation in a region associated with voice processing (the superior temporal sulcus) in response to native- compared to foreign-species voices (Fecteau, Armony, Joanette & Belin, 2004, see Belin, 2006 for review). However, it remains unknown as to what regions of the brain are associated with the processing of these voice types during infancy, and whether these brain regions change over the course of development. Investigations of changes over infancy in brain regions activated during voice processing, and comparisons with changes in brain regions activated for face and intersensory processing, would help to clarify whether perceptual narrowing is associated with shifts in the brain regions processing these stimuli

Considering the aforementioned models, which attempt to explain the neural mechanisms underlying perceptual narrowing, it is important to note that ERP research on perceptual narrowing provides evidence that some neural sensitivity remains to foreign distinctions in infancy, even after infants are no longer behaviourally showing evidence of sensitivity to these distinctions (Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Scott, Shannon, & Nelson, 2006). For example, although behavioural studies reveal no evidence of sensitivity to foreign phoneme contrasts by about 10 months of age, a longitudinal ERP study demonstrated that from 7 to 11 months of age, infants retain neural sensitivity to foreign phoneme contrasts (Rivera-Gaxiola et al., 2005). Similarly in the visual domain, 9-month-olds generally show behavioural specialization for own-species (human) over foreign-species (monkey) faces (Pascalis, de Haan & Nelson, 2002). Neural evidence also demonstrates that there is some neural specialization for human compared to monkey faces at this age, but 11-month-old infants still show a differential ERP response to familiar versus unfamiliar monkey faces, suggesting that some sensitivity is maintained to individual distinctions in foreign-species faces (Scott et al., 2006). As this ERP response is preconscious, this research suggests that the brain retains sensitivity to foreign distinctions, even when participants are not consciously aware of these distinctions. Scott et al. (2007) suggest that although the brain may retain some sensitivity to distinctions in foreign faces and speech sounds, experience with these distinctions might be required in order for this sensitivity to manifest behaviourally. Thus, remaining neural sensitivity might contribute to the reinstatement of infants' ability to manifest behavioural sensitivity to distinctions in foreign social stimuli after experience. In Chapter 2 of the present thesis, I demonstrate that *behavioural* sensitivity to individual distinctions in foreign-species (monkey) voices decreases considerably between 6 to 12 months of age. Although neural evidence of species specificity has been found during the processing of voices in human adults (Fecteau et al., 2004, see Belin, 2006 for review), *neural* sensitivity to foreign voices has not been measured in infants. Furthermore, it is unknown whether this sensitivity remains later during the first year after birth. If neural sensitivity to foreign-species voices is found during this period, this would lend support to the idea that remaining preconscious neural sensitivity to foreign distinctions contributes to the effects of experience on maintaining or reinstating sensitivity to these distinctions during infancy.

Beyond understanding the mechanisms underlying the narrowing process, it is important to consider the benefits of perceptual narrowing. With respect to the present thesis, one might consider why infants' perceptual abilities become specialized for native-species voices. As early as the third prenatal trimester, there is evidence suggesting that the fetus' responsiveness to sounds reaches sophisticated levels, as documented in particular in relation to the perception of speech (see DeCasper & Fifer, 1980 for review). In addition, many of the auditory abilities required to differentiate individuals by voice are already present in newborns, such as sensitivity to intonation, variation in frequency, rhythmicity and the phonetic components of speech (see DeCasper & Fifer, 1980; Floccia, Nazzi & Bertoncini, 2000 for reviews). These early abilities also facilitate language acquisition. In adulthood, familiarity with voices also has been shown to facilitate linguistic performance. For example, identifying novel words masked by varying levels of noise is easier when listening to a familiar compared to unfamiliar speaker (e.g., Nygaard, Sommers & Pisoni, 1994). However, language comprehension is unlikely the only benefit for becoming specialized for discriminating human voices in one's native language group.

Another proposed benefit of infants' early auditory competency is that it might facilitate mother-infant bonding (see DeCasper & Fifer, 1980 for review). Indeed, the ability to discriminate the voice of one's mother from that of a female stranger might encourage the bonding process. However, perceptual narrowing for species-specific voice discrimination appears to occur largely between 6 and 12 months of age, long after a bond has been established between the infant and its mother. During this period of time, the infant is becoming more aware of other people in his or her environment. Perhaps, then, at this point in development it is important to be able to discriminate and identify the individuals in the immediate environment, and perceptual narrowing helps in this endeavor. It is reasonable to postulate that the perceptual narrowing of voice discrimination (in addition to face, phonetic, and musical discrimination) abilities facilitates infants' learning of the social group to which they belong (the 'in-group') and which they do not (the 'out-group'). Identifying one's in-group, such as one's native race or species, has potential evolutionary benefits. In the environments of our early ancestors, members of one's in-group generally were encountered more frequently than out-group members, making them more relevant in terms of decision-making related to survival and reproductive fitness (Ackerman et al., 2006). As infants progress through their first year, they become increasingly mobile and can begin to make decisions about whom they approach and from whom they withdraw. In-group members tend to be more likely than out-group members to provide these infants with resources important to their survival, such as food and shelter. As development progresses through to adulthood, social interactions that are most important to an individual's outcomes (e.g., mate selection, negotiation, and exchange of status hierarchies) occur mainly within the in-group, which was the case historically, and remains so today (Ackerman et al., 2006; Fiske, 1992). In contrast, interaction with out-group members tends to be less frequent and generally present at the group, rather than individual, level (Ackerman et al., 2006). Thus, there are significant advantages to specializing for identifying between members of one`s in- versus out-group by the end of the first year after birth.

Limitations and Future Directions

The present thesis is the first to examine the time line of perceptual narrowing for voice-discrimination abilities during the first year after birth. However, some key questions remain unanswered. First, why do the 6-month-olds tested in Chapter 2 demonstrate a slight processing advantage for native-species (human) voices? Second, until what age does the window of plasticity (i.e., the sensitive period) remain open for learning to discriminate distinctions in foreign-species voices after the first year? Third, what is/are the neural mechanism(s) driving the narrowing process during the early development of voice discrimination? And fourth, do the perceptual narrowing results found in the present study generalize to other voice types, or other types of human vocalizations?

In Chapter 2, the results suggest that the beginnings of specialization for processing individual distinctions in human voices already might be evident as early as 6 months of age, which is in line with the narrowing time frame for some previously-examined stimuli (other-race faces, vowels), but out of line with the time frame for other stimuli (monkey faces, consonants, musical pitch and rhythm) that do not show evidence of narrowing until after 6 months of age (see Lewkowicz & Ghazanfar, 2009; Scott et al., 2007 for reviews). One possible explanation for this relatively early narrowing time line, proposed in Chapter 2, is that perceptual narrowing occurs earlier in development for stimuli that enable individual identification, than for stimuli that enable communication between individuals. However, it is possible that the small human-voice processing advantage found at 6 months does not reflect exclusively early experience-driven perceptual narrowing.

An alternative explanation for the pattern of results observed at 6 months of age involves inherent properties of the vocal stimuli used for testing rather than experience. This alternative explanation is that the human voice stimuli used for testing in Chapters 2 through 4 might have been intrinsically easier to distinguish than the monkey voice stimuli. Although similarities exist between human and rhesus monkey vocalizations, there are also a number of differences related to the anatomy of each species' vocal tracts and the spectral qualities of their vocal output as described in Chapters 2 to 4 (Fant, 1960; Fitch, 1997; Fitch & Giedd, 1999; Owren & Rendall, 2003; Schon Ybarra, 1995; Titze, 1994). These differences could enable participants to detect distinctions in the quality of individual human voices more easily than in rhesus monkey voices. This idea is supported by research demonstrating that certain acoustic differences may be intrinsically more difficult to discriminate than others, such as phonemic sounds that have relatively low acoustic salience (Narayan, Werker & Beddor, 2010). It remains for future studies to investigate the intrinsic similarity of human and monkey vocalizations in order to determine if internal qualities of the testing stimuli (rather than, or in addition to, experience-driven perceptual narrowing) influences the developmental perception of individual voices across species. Three ways the inherent discriminability of the human and monkey vocal stimuli could be tested are (1) to add a control condition where participants are required to discriminate backward vocalizations (removing effects of prosody and language), (2) to test Chinese- and English- speaking adults on their discrimination of Chinese and English voices, as well as monkey vocalizations, to determine if the results in Chapter 3 truly reflect an own-language effect or rather stimuli differences in discriminability, or (3) to create synthetic vocal stimuli that differ only on one/some of the acoustic qualities inherent in the human and primate vocal stimuli used in the present thesis. The latter method would enable researchers to determine whether one particular quality (or combination of qualities) is related directly to the discriminability of voices by infants and adults. For example, in Chapters 2 to 4, the human vocalizations used for testing were approximately 300 ms longer on average than the monkey vocalizations. Also, the fundamental and formant frequencies of the monkey vocalizations were substantially higher than those of the human vocalizations (see Chapter 2 methods for values). It is unclear whether the monkey vocalizations were more difficult to discriminate because they were shorter (possibly containing less information) or higher in fundamental frequency than the human voices. In terms of the fundamental frequency of the voices, it has been shown that 6-month-olds are worse at discriminating synthetic English vowel sounds that are higher in pitch (presented in a female, infant-directed pitch range) compared to those that are lower in pitch (presented in a female, adult-directed range) (Trainor & Desjardins, 2002). One possibility is that this is because voices with lower pitch have more narrowly-spaced harmonics than higher-pitched voices, with harmonic energy closer to the frequencies of the formant resonances, leading to it being easier to detect the location of the formants (Lieberman & Blumstein, 1988). This supports the idea that the human voices might be easier to discriminate because their pitch is lower than that of the monkey voices. Therefore, a study examining the influence of such acoustic qualities would help identify whether acoustic traits of monkey voice stimuli might be intrinsically more difficult to discriminate than those of human voice stimuli.

One might argue that the pattern of results found in the present thesis for infants could reflect 6-month-olds simply being worse at the head-turn task than 12-month olds. However, the findings argue against this as these groups did not differ significantly in their discrimination of human voices, suggesting that the head-turn task was not more difficult for infants at 6 compared to 12 months of age. This, in addition to the fact that infants' ability to discriminate monkey voices decreases between 6 and 12 months of age, suggests that acoustic differences between the human and monkey voices do not explain the perceptual narrowing findings. In the future, it would be useful to conduct a crossover design, comparing rhesus monkeys' ability to discriminate among the human and monkey vocalizations used in the present study to those of humans. This comparison would shed light on whether the monkey vocalizations are inherently easier to discriminate than the human vocalizations, because in this case the monkeys should demonstrate better discrimination of the human voices, rather than the expected own-species voice discrimination advantage. Furthermore, a language crossover design, testing the discrimination of Mandarin vocalizations by both English-and Mandarin-speaking adults, would be useful in supporting the proposed language effects reported in Chapter 3 of this thesis.

Another explanation for the pattern of results observed at 6 months of age that does not involve experience relates to the different developmental trajectories for the maturation of spectral and temporal auditory processing. Previous research has found earlier evidence of perceptual narrowing for vowel sounds than for consonant sounds (see review by Kuhl, 2008). While temporal (duration) features tend to define consonant sounds, spectral (sound frequency) features primarily define vowel sounds (e.g., Rosen, 1992). A number of studies suggest that fine temporal processing is relatively immature during early infancy and that the mechanisms that enable temporal processing develop more slowly than those for spectral processing (Clarkson & Clifton, 1995; Smith, Trainor & Shore, 2006; Werner, Marean, Halpin, Spetner & Gillenwater, 1992). For this reason, perceptual narrowing may occur earlier for sounds involving primarily spectral information, such as vowels, as well as individual vocal quality (or timbre), than for sounds involving primarily temporal information, such as consonants and musical rhythm. On the other hand, this explanation does not account for the finding that perceptual narrowing occurs relatively early for face processing (e.g., Kelly et al., 2007); nor does it account for findings that musical pitch processing narrows relatively late during development (e.g., Lynch, Eilers, Oller, & Urbano, 1990; Trainor & Trehub, 1992). In the present thesis, both the human "balloon" and monkey "coo" vocalizations contained primarily vowel sounds. However, the human, but not the monkey, vocalizations also included consonants (e.g., /b/). This makes it hard to rule out the role temporal processing might play in participants' discrimination of the human voices. Thus, the idea that voice discrimination narrows relatively early because it depends largely on spectral auditory processing requires further examination, such as comparing participants' discrimination of rhesus "coos" to that of human vowel sounds that do not contain consonants.

A related question is whether specialization for human voices is present at an age younger than 6-months-old. For example, would the trend for a human-voice discrimination advantage disappear at 4 months of age or is the small advantage for processing human voices found at six months of age also found at 4 months, or even at birth, representing an initial developmental condition in humans? There is some evidence to suggest that general changes between birth and 6 months occur for the perception of primate and human voices. For example, differential brain activation to human and primate vocalizations is found in Japanese 4-month-olds (Minagawa-Kawai et al., 2011), and general preferences for human speech over rhesus monkey vocalizations are found at 3 months of age, but not in neonates (Vouloumanos, Hauser, Werker, & Martin, 2010). Still, these differences could be due to the aforementioned disparity in the acoustic features between human and rhesus monkey voices. On the other hand, infants can match dynamic monkey faces with monkey vocalizations as early as 4 months of age (Lewkowicz & Ghazanfar, 2006), although this ability has been attributed more to sensitivity to temporal synchrony than to the monkey stimuli specifically (Lewkowicz, Sowinski & Place, 2008). Thus, if specialization for human voice discrimination follows the same early developmental time line found for the processing of faces and vowels, infants younger than 6 months should not demonstrate a clear advantage for processing human over monkey voices.

A second key question remaining in the present line of research concerns when the sensitive period ends for learning to discriminate voices from a foreign species. In particular, although Chapter 4 demonstrated that a period of exposure to individually-named rhesus monkey voices could reinstate sensitivity to individual distinctions in novel examples of this category of voices at 12 months of age, it remains unclear as to how long this window of plasticity remains open. Research on non-human primates highlights the important role of exposure on sensitive periods, demonstrating that early facial deprivation, followed by a period of exposure to own- or foreign-species faces, can extend the period of specialization for experienced facial stimuli for at least 2 years (Sugita, 2008). Human studies suggest that the timing of exposure (or training) influences the ability to reinstate sensitivity to foreign distinctions, as this ability decreases with age. Although adults can learn to discriminate foreign distinctions (e.g., foreign phonemes), the reinstatement is limited and sensitivity to foreign distinctions rarely reaches the level observed in young infants (e.g., Flege et al., 1999; Hannon & Trehub, 2005b; Iverson et al., 2005; McCandliss, Fiez, Protopapas, Conway & McClelland, 2002; Takagi, 2002; Takagi & Mann, 1995, for reviews see Birdsong, 2006; Hernandez & Li, 2007). It remains for future research to test the effects of exposure to foreign-species voices in children, adolescents, and adults in order to determine if the effects of training decreases with age. If there is significantly more improvement in ability to discriminate foreign-species voices at 12 months of age than at older ages, or if there is no improvement at these older ages, this would support the idea that a sensitive period for foreign voice exposure ends during infancy. However, if the levels of improvement following exposure are comparable across these age groups, then this suggests that the period of sensitivity extends past the first year after birth.

Sensitivity to foreign distinctions is influenced not only by the timing of exposure but also by the type of exposure received. For example, adult discrimination of foreign phoneme contrasts generally improves when the training stimuli are numerous and variable in nature (Zhang et al., 2009), and differences between them are exaggerated (McCandliss et al., 2002). As well, in infancy, discrimination of foreign phonemes requires social interaction (e.g., Kuhl et al., 2003). Furthermore, training at an individual-level (e.g., labeling monkey faces using unique names), rather than a categorical-level (e.g., labeling all monkey faces with the word "monkey"), is crucial in order for infants to learn to individuate foreign-species faces (e.g., Scott & Monesson, 2009). In Chapter 4, efforts were made to produce training that aligned with the type of exposure shown to be important for reinstating sensitivity to foreign distinctions during infancy. Specifically, parents were instructed to interact with their infants during training, and each monkey voice was associated with a particular name in order focus attention at an individual level. This type of training was effective in reinstating sensitivity to distinctions in foreign-species voices. However, it remains unclear whether social interaction and individual-level exposure are necessary for effective reinstatement of foreign-voice sensitivity. In order to investigate this, future studies could instruct parents to avoid interaction with their infant during training, or train infants at a categorical, rather than individual, level by associating each monkey voice with the word “monkey”, rather than an individual name.

A third key question that remains in the present line of research pertains to the mechanisms supporting perceptual narrowing. As discussed above, it was postulated originally that perceptual narrowing results from a process of neural pruning (Changeux & Danchin, 1976; Cowan et al., 1984; Goldman-Rakic, 1987; Huttenlocher, 1979; Low & Chen, 2006; Scott et al., 2007; Webb et al., 2001). However, there is also evidence suggesting that perceptual narrowing may be supported by a process of neural elaboration (see Lewkowicz & Ghazanfar, 2009, Box 2). Evidence from a variety of mammalian species, including humans, suggests that these two models are not mutually exclusive, as pruning tends to follow elaborations across the lifespan during neural development (e.g., Bourgeois, 1997; Sowell et al., 2003). However, one way it may be possible to start examining the relative contributions of these two proposed mechanisms is by examining the changes in grey matter (GM) and white matter (WM) volumes in the developing human brain. Such changes suggest that elaboration might play a larger role than pruning in brain development during the first year after birth, the period when perceptual narrowing generally takes place (see Lewkowicz & Ghazanfar, 2009; Scott, Pascalis & Nelson, 2007 for reviews). Evidence from MRI studies suggests that the neonatal brain is about half the volume of the adult brain, growing to about 90% of the adult brain volume by the second year after birth (Pfefferbaum et al., 1994). Particularly, during the first postnatal year, there is a substantial growth in GM volume (Gilmore et al. 2007, 2011; Hüppi et al. 1998; Knickmeyer et al., 2008; Matsuzawa et al., 2001), and relatively small increases in WM (Knickmeyer et al., 2008). Increases in GM and WM volume have been proposed to relate to neurodevelopmental events that are progressive in nature, such as neural elaboration (e.g., Giedd et al., 1999; Pfefferbaum et al., 1994; Purves, 1988). In particular, from 5 to 10 years of age, increases in GM are likely associated with cell proliferation, arboration, and myelination (Cowan, Fawcett, O'Leary & Stanfield, 1984; Jacobson, 1991; Purves 1988), Similarly, increases in WM are likely associated with myelination and axonal growth (e.g., Holland, Haas, Norman, Brant-Zawadski, & Newton, 1986; Purves, 1988). On the other hand, reduction of GM, which has been documented in some studies after 5 years of age (Jernigan & Talla, 1990; Lange, Giedd, Castellanos, Vaituzis, & Rapaport, 1997; Hüppi et al., 1998), has been associated with neural pruning (Huttenlocher, 1979). Nevertheless, this method does not allow for examination of the perceptual narrowing process specifically, and may rather reflect processes associated with general brain development. Additionally, a third mechanism that has been suggested to support perceptual narrowing is a switch, for the most part, from prominent activation of brain regions associated with attention, to those associated with perception during the first year after birth (Vogel et al., 2012). Further investigation is required to better understand the role of neural pruning, elaboration and regional shifts in facilitating perceptual narrowing in general, and the specialization for own-species voices in particular.

Finally, questions remain concerning the generalization of the present study to other voice types. It is important to examine whether infants demonstrate the same perceptual narrowing and exposure effects when processing voices from foreign non-primate species, foreign human languages, male in addition to female voices, and from familiar individuals. Such investigations will provide additional evidence with which to support or refute previous claims that perceptual narrowing is a general learning mechanism (for review see Lewkowicz & Ghazanfar, 2009).

Conclusions

In summary, the present thesis examined whether infants' ability to discriminate individual voices develops through a process of perceptual narrowing, such that it becomes specialized for native-species voices during the first year after birth, and whether experience can reinstate discrimination of foreign voices after narrowing has occurred. These experiments address an important aspect of development, as the ability of infants to discriminate members of their native group is likely to contribute highly to successful group membership. This thesis is the first to provide evidence that (1) perceptual narrowing occurs for the discrimination among native as opposed to foreign-species voices, between 6 to 12 months of age, resulting in a specialization for processing human (English) voices by 12 months of age; (2) the voice specialization found in adulthood is influenced not only by the language of the voices, but also by the whether the voices are from one's own species; and (3) two weeks of exposure at 12 months of age to voices from a foreign species can reinstate sensitivity to individual distinctions in foreign-species voices, suggesting that the voice-processing system remains plastic at least until the end of the first year after birth. Future research is required to examine cross-species voice discrimination prior to 6 months of age, to determine how long the window of plasticity remains open after 12 months of age, and to examine the effects of the timing, type and amount of exposure on the process of perceptual narrowing for voices.

REFERENCES

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., Maner, J. K., & Schaller, M. (2006). They All Look the Same to Me (Unless They're Angry): From Out-Group Homogeneity to Out-Group Heterogeneity. *Psychological Science*, *17*(10), 836-840.

Agmon, A., Yang, L. T., O'Dowd, D. K., & Jones, E. G., (1993). Organized growth of thalamocortical axons from the deep tier of terminations into layer IV of developing mouse barrel cortex. *The Journal of Neuroscience, 13*(12), 5365-5382.

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology, 112,* 484- 495.

Belin, P. (2006). Voice processing in human and non-human primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 361*(1476), 2091-2107.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*, 9-49.

Black, J., Jones, T., Nelson, C., & Greenough, W. (1998). Neuronal plasticity and the developing brain. In J. Noshpitz (Ed.), *Handbook of child and adolescent psychiatry: Vol. 6. Basic psychiatric science and treatment* (pp.

31-53). New York: Wiley.

Bourgeois, J. P. (1997). Synaptogenesis, heterochrony and epigenesis in the mammalian neocortex. *Acta Paediatrica*, *86*(S422), 27-33.

Bourgeois, J. P. & Rakic, P. (1993). Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. The *Journal of Neuroscience, 13*(7), 2801–2820.

Changeux, J. P., & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature, 264*, 705-712.

Clarkson, M. G., & Clifton, R. K. (1995). Infants' pitch perception: Inharmonic tonal complexes. *Journal of the Acoustical Society of America, 98*, 1372-1379.

Cowan, W. M., Fawcett, J. W., O`Leary, D. D. M., & Stanfield, B. B. (1984). Regressive events in neurogenesis. *Science, 225*(4668), 1258-1265.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176.

de Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLoS ONE, 3*(5), e2317.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2004). Is voice processing species specific in human auditory cortex? An fMRI study. *Neuroimage*

23(3), 840-848.

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*, 689-723.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America, 102*(2), 1213-1222.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*(3), 1511-1522.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language, 41*, 78-104.

Floccia, C., Nazzi, T. & Bertoncini, J. (2000). Unfamiliar voice discrimination for short stimuli in newborns. Developmental Science, 3(3), 333-343.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C. & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*(10), 861-863.

Gilmore, J. H., Lin, W., Prastawa, M. W., Looney, C. B., Vetsa, Y. S., Knickmeyer, R. C., Evans, D. D., Smith, J. K., Hamer, R. M., Lieberman, J. A., & Gerig, G. (2007). Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *Journal of*

Neuroscience, 27(6), 1255–1260.

Gilmore , J. H., Shi, F., Woolson, S. L., Knickmeyer, R. C., Short, S. J., Lin, W., Zhu, H., Hamer, R. M., Styner, M., Shen, D. (2011). Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cerebral Cortex*, doi: 10.1093/cercor/bhr327.

Goggin, J., Thompson, C., Sturbe, G. & Simental, L. (1991). The role of language familiarity in voice identification. *Memory & Cognition 19*(5), 448-458.

Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development, 58*(3), 601-622.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, USA, 102*(35), 12639-12643.

Hernandez, A., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638-650.

Holland, B. A., Haas, D. K., Norman, D., Brant-Zawadski, M., Newton, T. H. (1986). MRI of normal brain maturation. *American Journal of Neuroradiology, 7*, 201-208.

Hüppi, P. S., Warfield, S., Kikinis, R., Barnes, P. D., Zientara, G. P., Jolesz, F. A., Tsuji, M. K., Volpe, J. J. (1998). Quantitative magnetic resonance imaging of brain development in premature and mature newborns. *Annals of Neurology, 43*(2), 224–235.

Huttenlocher, P. (1979). Synaptic density in human frontal cortex: Developmental changes and effects of aging. *Brain Research, 163*(2), 195-205.

Huttenlocher, P. R. & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*(2), 167-178.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*(5), 1002-1011.

Jernigan, T. L. & Tallal, P. (1990). Late childhood changes in brain morphology observable with MRI. *Developmental Medicine & Child Neurology, 32*(5), 379-385.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy. *Psychological Science, 18*(12), 1084-1089.

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., Hamer, R. M., Lin, W., Gerig, G., & Gilmore, J. H. (2008). A structural MRI study of human brain development from birth to 2 years. *The Journal* *of Neuroscience, 28*(47), 12176 –12182.

Köster, O. & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics 4*(1), 18-28.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 811-817.

Kuhl, P. K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research directions* (pp. 213-243). New York: Erlbaum.

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 of the United States of America, 100*(15), 9096-9101.

Lange, N., Giedd, J. N., Castellanos, F. X., Vaituzis, A. C., & Rapaport, J. L. (1997). Variability of human brain structure size: Ages 4-20 years. *Psychiatry Research: Neuroimaging, 74*(1), 1-12.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences of the United States of America, 103*(17), 6771-6774.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences, 13*(11), 470-478.

Lewkowicz, D., Sowinski, R., & Place, S. (2008). The decline of cross-species intersensory perception in human infants: Underlying mechanisms and its developmental persistence. *Brain Research*, 1242, 291-302.

Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. New York: Cambridge University Press.

Loucks, J., & Sommerville, J. A. (2012). Developmental changes in the discrimination of dynamic human actions in infancy. *Developmental Science, 15*(1), 123-130.

Low, L. K., & Cheng, H. J. (2006) Axon pruning: An essential step underlying the developmental plasticity of neuronal connections. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 361*(1473)*,* 1531-1544.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1*(4), 272-276.

Macchi Cassia, V., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing: evidence for the reactivation of dormant effects. *Psychological Science, 20*(7), 853-859.

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: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective & Behavioral Neuroscience, 2*(2), 89-108.

Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral Cortex, 21*(2), 254-261.

Narayan, C., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science, 13*(3), 407-420.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*(1), 42-46.

Oh, J. S., Au, T. K. F., & Jun, S. A. (2010). Early childhood language memory in the speech of international adoptees. *Journal of Child Language, 37*(5), 1123-1132.

Owren, M. J., & Rendall, D. (2003). Salience of caller identity in rhesus macaque (macaca mulatta) coos and screams: Perceptual experiments with human (homo sapiens) listeners. *Journal of Comparative Psychology, 117*(4), 380-390.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex, 13*,155-161.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development, 83*(2), 543-553.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296(5571), 1321-1323.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences of the United States of America 102*(14), 5297-5300.

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A Quantitative Magnetic Resonance Imaging Study of Changes in Brain Morphology From Infancy to Late Adulthood. *Archives of Neurology, 51*(9), 874-887.

Purves, D. (1988). Body and Brain. Cambridge, Mass: Harvard University Press.

Purves, D., White, L. E., & Riddle, D. R. (1996). Is neural development Darwinian? *Trends in Neurosciences, 19*(11), 460-464.

Quartz, S. R., & Sejnowsky, T. J. (1997) A neural basis of cognitive development:

A constructivist manifesto. *Behavioral and Brain Sciences, 20*, 537-556.

Rivera-Gaxiola, M., Silva-Pereyra, J. & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*(2), 162-172.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society B, 336*, 367-373.

Sangrigoli, S., Pallier, C., Argenti, A., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*(6), 440-444.

Schon Ybarra, M. (1995). A comparative approach to the non-human primate vocal tract: Implications for sound production. In E. Zimmermann, J. D. Newman & U. Jürgens (Eds.), *Current topics in primate vocal communication* (pp. 185-198). New York: Plenum.

Scott, L. S., & Monesson, A. (2009). The origins of biases in face perception. *Psychological Science, 20*(6), 676-680.

Scott, L. S., & Monesson, A. (2010). Experience-dependent neural specialization during infancy. *Neuropsychologia, 48*, 1857-1861.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science, 16*(4), 197-201

Scott, L. S., Shannon, R. W., & Nelson, C. A. (2006). Neural correlates of human

and monkey face processing by 9-month-old infants. *Infancy*, *10*(2), 171-186.

Sliwa, J., Duhamel, J., Pascalis, O., & Wirth, S. (2011). Spontaneous voice–face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences of the USA, 108*(4), 1735-1740.

Smith, N. A., Trainor, L. J., & Shore, D. I. (2006). The development of temporal resolution: Between-channel gap detection in infants and adults. *Journal of Speech, Language, and Hearing Research, 49(5), 1104-1123.*

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L. & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6*(3), 309-315.

Sowell, E. R., Thompson, P. M. & Toga, A. W. (2004). Mapping changes in the human cortex throughout the span of life. *The Neuroscientist, 10*, 372-392.

Sugita, Y. (2008) Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences, USA, 105*(1), 394-398.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*(6), 2887-2896.

Takagi, N., & Mann, V. A. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics, 16*(4), 379-405.

Tees, R. C., & Werker, J. F. (1984). Perpetual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 34*(4), 579-590.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice-Hall, Inc. Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 394-402.

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review, 9,* 335-340.

Ventureyra, V. A. G., Pallier, C. & Yoo, H. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79-91.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental Science, 15*(3), 359-372.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development, 81*(2), 517-527.

Webb, S. J., Monk, C. S., & Nelson, C. A. (2001). Mechanisms of postnatal neurobiological development: Implications for human development. *Developmental Neuropsychology, 19*(2), 147-171.

Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development, 63*, 260-272.

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., Kawakatsu, M., Tohkura, Y., & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *Neuroimage, 46*, 226-240.