Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants

Feng-Ming Tsao[a)]
Department of Psychology, National Taiwan University, Taipei, Taiwan 106 Taiwan, Republic of China

Huei-Mei Liu
Department of Special Education, National Taiwan Normal University, Taiwan, Republic of China

Patricia K. Kuhl
Institute for Learning and Brain Sciences, University of Washington, USA

(Received 19 September 2005; revised 25 July 2006; accepted 26 July 2006)

Previous studies have shown improved sensitivity to native-language contrasts and reduced sensitivity to non-native phonetic contrasts when comparing 6–8 and 10–12-month-old infants. This developmental pattern is interpreted as reflecting the onset of language-specific processing around the first birthday. However, generalization of this finding is limited by the fact that studies have yielded inconsistent results and that insufficient numbers of phonetic contrasts have been tested developmentally; this is especially true for native-language phonetic contrasts. Three experiments assessed the effects of language experience on affricate-fricative contrasts in a cross-language study of English and Mandarin adults and infants. Experiment 1 showed that English-speaking adults score lower than Mandarin-speaking adults on Mandarin alveolo-palatal affricate-fricative discrimination. Experiment 2 examined developmental change in the discrimination of this contrast in English- and Mandarin-leaning infants between 6 and 12 months of age. The results demonstrated that native-language performance significantly improved with age while performance on the non-native contrast decreased. Experiment 3 replicated the perceptual improvement for a native contrast: 6–8 and 10–12-month-old English-learning infants showed a performance increase at the older age. The results add to our knowledge of the developmental patterns of native and non-native phonetic perception. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2338290]

PACS number(s): 43.71.Hw, 43.71.Ft, 43.71.Es, 43.70.Fq [ALF] Pages: 2285–2294

I. INTRODUCTION

The initial state of the infant's mind and the mechanisms for developmental change have stood at the heart of the nature-nurture debate regarding phonetic perception (Kuhl, 2000; 2004; Werker and Curtin, 2005; Best, 1995; Nittrouer, 2001; Aslin *et al.*, 2002). Young infants are able to discriminate phonetic contrasts from both their native (e.g., Eimas, Siqueland, Jusczyk, and Vigorito, 1971) and from foreign languages (e.g., Lasky *et al.*, 1975; Streeter, 1976; Trehub, 1973), while adult listeners generally find non-native discrimination difficult (e.g., Miyawaki *et al.*, 1975; Strange and Jenkins, 1978). Studies suggest that listening to native language speech alters infants' phonetic discrimination to produce a language-specific bias during the first year of life (Best *et al.*, 1995; Kuhl *et al.*, 1992; Werker and Tees, 1984).

Many studies have demonstrated a decline in infants' discrimination of non-native phonetic contrasts between 6 and 12 months of life (Best and McRoberts, 2003; Best, McRoberts, LaFleur, and Silver-Isentadt, 1995; Kuhl *et al.*, 2006; Werker and Tees, 1984). For example, Werker and Tees (1984) tested speech discrimination for non-native contrasts in 6- to 12-month-old infants from English-speaking families with two non-native place contrasts: the Hindi retroflex/dental stops [ʈa] vs [ta] and the velar/uvular glottalized voiceless stops [k̀i] vs [q̀i] in the Thompson language of the Salish Indians. The results demonstrated that infants aged 6–8 months were sensitive to differences between these non-native contrasts but that this perceptual sensitivity to foreign contrasts was significantly reduced in 10–12-month-old infants. However, not all studies on non-native contrasts show a decline between 6 and 12 months of age (Best *et al.*, 1988; Polka *et al.*, 2001).

Fewer tests have examined developmental change for native-language contrasts between 6 and 12 months (Eilers, Wilson, and Moore, 1977; Kuhl *et al.*, 2006; Polka *et al.*, 2001). Kuhl *et al.* (2006), in a recent cross-language study using American English /r-l/ in tests on American and Japanese 6–8 and 10–12-month-old infants, showed that native-language phonetic perception improves between 6 and 12 months, while non-native phonetic perception showed the typically observed decline. Previous studies on native language phonetic perception only suggested a developmental change *during* the first year of life (Eilers *et al.*, 1977) or have shown improvement *after* the first year of life (Polka *et al.*, 2001; Sundara *et al.*, 2006). Studies using a neural measure of discrimination have shown the pattern of facilitation in native language perception and the decline in non-native

[a)]Author to whom correspondence should be addressed; Department of Psychology, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, TAIWAN 106, Phone: +886–2–3366–4467, Fax: +886–2–2363–1463. Electronic mail: tsaosph@mail2000.com.tw

perception in the first year of life—event-related potential studies of phonetic perception show larger amplitude differences for native phonetic discrimination in 10–12 month olds when compared to 6–8-month-old infants (Cheour, 1998; Rivera-Gaxiola *et al.*, 2005), and a decline for non-native discrimination.

It is not clear why there is an age variation in the patterns of developmental change for native and non-native speech. It has been suggested that timing differences of the developmental change in perception could be attributable either to the acoustic fragility of certain phonetic contrasts, such as fricatives, or the frequency of occurrences of the phonetic unit in unstressed contexts, which would also diminish the acoustic cues (Polka *et al.*, 2001; Sundara *et al.*, 2006). Without tests on additional phonetic contrasts, these questions cannot be answered. This study examined infants' perceptual sensitivity to affricate-fricative contrasts, which have not been previously studied, using a cross-language design. Infants in two age ranges, 6–8 months of age and 10–12 months of age, were tested in Taiwan and the United States to investigate the patterns of developmental change for these consonants during the second-half of the first year of life.

Because cross-cultural studies on adults have not been done using these Mandarin contrasts, our studies began with an examination of the acoustic events that signal the affricate-fricative distinction contained in Mandarin, and perceptual studies on Mandarin and English adult speakers. Many studies on adult speakers show that the phonetic contrasts of foreign languages can be difficult to discriminate (Goto, 1971, Miyawaki *et al.*, 1975; Sheldon and Strange, 1982; Trehub, 1976; Werker, Gilbert, Humphery, and Tees, 1981). The classic case is the difficulty that Japanese speakers have in discriminating English /r-l/, a phonemic contrast that is not utilized in Japanese (Miyawaki *et al.*, 1975). In the present study, it is important to note that speakers of both languages (Mandarin and English) utilize an affricate-fricative manner distinction phonemically in their language; however, the distinction in Mandarin utilizes a different place of articulation, and there are subtle acoustic differences in the way the affricate-fricative distinction is realized in the two languages. We therefore hypothesized that adult speakers of the two languages would emphasize different acoustic cues in perception, and that adult speakers of English would therefore find it more difficult to discriminate the Mandarin contrast.

To summarize, this study's goals were twofold: (a) Experiment 1 examined adult perception of affricate-fricative consonants in a cross-language test on Mandarin and English speakers, and (2) Experiments 2 and 3 examined cross-language patterns of developmental change in native (Exp 2 and 3) and non-native (Exp 2) phonetic perception using affricate-fricative consonants during the first year of life. We hypothesized that (a) adult Mandarin speakers would outperform adult English speakers on Mandarin contrasts in Experiment 1, (b) that infants tested in Experiments 2 and 3 on native-language affricate-fricative contrasts would show significant improvement between 6 and 12 months of age, and (c) that infants for whom the affricate-fricative contrast was non-native in Experiment 2 would show a decline between 6 and 12 months of age.

II. EXPERIMENT 1: ENGLISH- AND MANDARIN-SPEAKING ADULTS ON MANDARIN FRICATIVE VS AFFRICATE DISCRIMINATION

A reasonable starting point to explore phonetic discrimination difficulty for non-native contrasts is to identify and test contrasts that are phonemic in one language but not in another language. The alveolo-palatal affricate vs fricative contrasts of Mandarin Chinese, e.g., /tɕ/ vs /ɕ/ and /tɕʰ/ vs /ɕ/, are appropriate for this purpose. First, the manner difference between affricate and fricative is phonemic in English (e.g., /tʃ/ vs /ʃ/ and /dʒ/ vs /ʒ/), but the place of articulation of the Mandarin contrasts, alveolo-palatal, does not occur in English. The English palato-alveolar consonants have a constriction in the vocal tract that is forward of alveolo-palatal sounds (Ladefoged and Maddieson, 1995). This articulation difference results in different acoustic features between Mandarin and English. For example, the spectral peaks of Mandarin alveolo-palatal consonants (around 4900 Hz, Liu, 1996) are located between English palato-alveolar (3800 Hz) and alveolar consonants (6839 Hz) (Jongman *et al.*, 2000).

Second, the exact phonetic features that distinguish affricate and fricative consonants differ in the two languages. Mandarin has three voiceless alveolo-palatal sounds, including two affricates /tɕ/, /tɕʰ/ and one fricative /ɕ/. The phonetic feature of aspiration distinguishes the affricate /tɕ/ [−aspirated] from its counterpart /tɕʰ/ [+aspirated]. In contrast, the English palato-alveolar consonants (e.g., /tʃ/ vs /ʃ/ and /dʒ/ vs /ʒ/) are distinguished with the phonetic features of articulation manner (affricate vs fricative) and voicing (voiced vs voiceless) (Chomsky and Halle, 1968). The phonetic differences between English and Mandarin result in different acoustic correlates in the two languages. Amplitude rise time and frication duration are the two relevant acoustic cues distinguishing the affricate vs fricative contrasts in English (Cutting and Rosner, 1974; Howell and Rosen, 1983; Kluender and Walsh, 1992; Hedrick, 1997). However, frication duration is perceptually more prominent than amplitude rise time for identification of the affricate and fricative (Kluender and Walsh, 1992). Studies on the acoustic correlates of Mandarin affricate and fricative contrasts show that affricates have shorter frication duration and higher occurrence of an initial burst (Liu *et al.*, 2000). However, no study has measured the amplitude rise time of this Mandarin phonetic contrast. In addition, it is unclear whether frication duration and amplitude rise time are both perceptually relevant for differentiating affricates and fricatives in Mandarin Chinese. In brief, both the phonetic and acoustic differences between Mandarin and English provide a reasonable basis for hypothesized performance differences between English and Mandarin adult speakers on the perception of Mandarin alveolo-palatal contrasts.

Experiment 1 tested the hypothesis that English adult speakers will show significantly poorer performance when compared to adult Mandarin speakers on the discrimination of Mandarin alveolo-palatal affricate and fricative contrasts.

TABLE I. Acoustic features of Mandarin affricate versus pricative contrasts in experiment 1.

Phonetic contrasts	Phonetic feature differences	Stimulus pairs	Amplitude rise time (+3 dB SPL from start in ms)	Frication noise duration (in ms)
Set 1:	Aspirated	1	30 vs 100	Same (130 ms)
/tɕʰ/ vs /ɕ/	Affricate vs	2	30 vs 60	Same (130 ms)
	Fricative	3	30 vs 60	Same (160 ms)
Set 2:	Aspirated	4	Same (30 ms)	130 vs 80
/tɕʰ/ vs /ɕ/	Affricate vs	5	Same (30 ms)	100 vs 50
	Unaspirated	6	(30 ms)	100 vs 80
Set 3:	Unaspirated	7	Same (50 ms)	55 vs 95
/tɕ/ vs /ɕ/	Affricate vs	8	Same (50 ms)	55 vs 85
	Fricative	9	(50 ms)	65 vs 95

Performance differences between adult speakers of the two languages were also expected to guide the selection of the specific speech stimuli that were used in Experiment 2.

A. Method

1. Participants

Thirty-six undergraduate and graduate students of the University of Washington without history of severe language or hearing impairments participated in this study. One group consisted of 18 native English speakers (ten females, eight males). Another group included 18 native Mandarin Chinese speakers (nine females, nine males) from either Taiwan or China. All participants were tested in Seattle. English speakers were recruited from the Psychology Subject Pool and received class credits for their participation. Mandarin speakers were recruited from a solicitation on web pages of international student organizations and received $20 for their participation.

2. Stimuli

Nine synthesized speech pairs were created for /tɕi/, /tɕʰi/, and /ɕi/ tokens using the HLsyn speech synthesizer 2.2 (1996); they were sampled at 11025 Hz. These pairs tested three sets of phonetic contrasts: /tɕi/ vs /tɕʰi/; /tɕi/ vs /ɕi/; and /tɕʰi/ vs /ɕi/. The synthetic stimuli varied in either amplitude rise time or frication noise duration. The acoustic differences for each stimulus pair are shown in Table I. The values of frication duration in each pair were chosen based on an acoustic analysis of Mandarin alveolo-palatal affricate and fricative consonants (Liu, 1996). However, no acoustic data are available for amplitude rise time in Mandarin affricate and fricative consonants. For amplitude rise time in the aspirated affricate /tɕʰ/ vs fricative /ɕ/ pair, perceptual studies in English (Cutting and Rosner, 1974; Howell and Rosen, 1983) suggested that phonetic boundaries between palato-alveolar affricate and fricative pairs were located around 50 ms. Therefore, the values of amplitude rise time and frication duration were varied across phonetic boundaries to generate separate phonetic categories for Mandarin-speaking adults. Amplitude rise time is defined as the time to reach the maximum amplitude (+3 dB SPL) of frication noise from the onset of the syllable.

The spectral peak of frication noise is around 4700 Hz, and the frication noise energy mainly occurs above 2500 Hz, typical values for Mandarin speakers (Liu, 1996). During the 245 ms vowel portion of the syllables, the formant frequencies were: 293, 2274, 3186, and 3755 Hz, respectively, for F1 through F4. The bandwidths of F1–F4 were 80, 90, 150, and 350 Hz, respectively. The fundamental frequency (pitch) of the syllable was 120 Hz, a typical value for male Mandarin speakers (Huang, 1996). Finally, tokens were equalized in rms amplitude. The appropriateness of these speech stimuli were judged by native Mandarin speakers from Taiwan and China in a pilot study.

3. Procedures

A computer presented pairs of these stimuli in an AX discrimination task. On each trial, the participant heard two tokens separated by 350 ms through earphones at a comfortable listening level of approximately 65 dBA in a sound-attenuated booth. Participants were asked to decide whether the pair of stimuli were the same or different. The probability of stimulus presentation order, i.e., A-A, A-B, B-A, and B-B, was equal and stimulus pairs were randomly presented across subjects. Prior to the test, participants practiced with synthetic stop-vowel (/ba/ vs /pa/) stimuli that were easy for both language groups and received feedback on their responses. During the test stage, 180 test pairs (=9 pairs $\times$4 presentation orders$\times$5 repetitions) were presented with no feedback. Every subject completed this experiment in 30–40 min.

B. Results and discussion

A bias-free measure of sensitivity (d') was calculated for the two language groups for each stimulus pair and the values are illustrated in Fig. 1. As shown, the Mandarin group shows better discrimination across all pairs (average $d'=1.74$, percent correct=77.62%) when compared with the

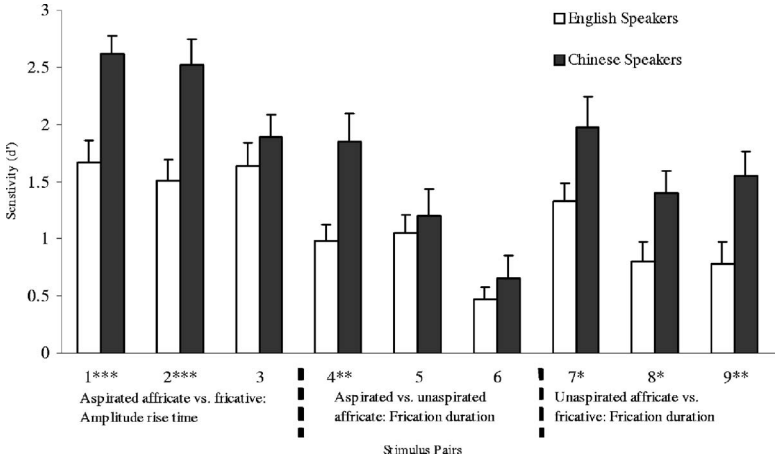


FIG. 1. Sensitivity (d') in discriminating Mandarin contrasts (*SE* in error bars, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$).

American English group (average $d'=1.13$, percent correct $=67.72\%$). A mixed-design two-way ANOVA (between subject factor: Language group $\times$ within subject factor: Stimulus pair) of d' reveals that both language $[F(1,34)=17.132]$ and stimulus factors $[F(8,272)=17.377]$ are significant at $p<0.001$. The nonsignificant interaction $[F(8,272)=1.842, p>0.1]$ of these two factors reveals that Mandarin speakers consistently performed better than American English speakers in discriminating these Mandarin contrasts.

The results of Experiment 1 indicate that amplitude rise time and frication duration contribute significantly to Mandarin speakers' discrimination of the three sets of phonetic contrasts. Amplitude rise time is effective for discriminating the aspirated affricate /tɕʰ/ vs fricative /ɕ/ contrast, $F(2,34)=5.445, p<0.01$. Frication duration is utilized by Mandarin speakers to distinguish the aspirated /tɕʰ/ vs unaspirated /tɕ/ affricate, $F(2,34)=14.563, p<0.001$, and unaspirated affricate /tɕ/ vs fricative /ɕ/ contrasts, $F(2,34)=3.478, p<0.05$.

An interesting question is the relative contribution of amplitude rise time and frication duration in the discrimination of Mandarin contrasts. Studies have shown that both rise time and frication duration covary in the English affricate and fricative distinction although frication duration is perceptually more effective (Kluender and Walsh, 1992). A different weighting of acoustic cues to the affricate and fricative distinction between English and Mandarin Chinese might be one of the key factors in the performance difference between the two language groups. This experiment examined the performance differences between the two language groups by utilizing only a small set of acoustic values. Further studies that covary rise time and frication duration will be needed to investigate the relative contribution of both acoustic cues in the discrimination of Mandarin affricative vs fricative consonants.

To summarize, the results of Experiment 1 show that English speakers experience more difficulty overall when compared to Mandarin speakers in the discrimination of Mandarin alveolo-palatal fricative vs affricate contrasts, indicating that language experience affects perceptual discrimination for non-native contrasts. In addition, the results suggest that amplitude rise time and frication duration are important for discriminating Mandarin alveolo-palatal contrasts. These data were used to guide selection of a contrast to use in tests on Mandarin- and English-learning infants in Experiment 2.

III. EXPERIMENT 2: TESTING MANDARIN- AND ENGLISH-LEARNING INFANTS ON THE PERCEPTION OF A MANDARIN AFFRICATE-FRICATIVE CONTRAST

The goal of this experiment was to explore the time course of developing native and non-native phonetic perception using a cross-language design. In Experiment 1, English-speaking adults performed significantly poorer than Mandarin-speaking adults in the discrimination of Mandarin Chinese alveolo-palatal fricative vs affricate tokens. The largest difference between English- and Mandarin-speaking adults was shown in the aspirated-affricate vs fricative contrast when the amplitude rise time was varied, and so this contrast was utilized in tests on infants.

A. Method

1. Participants

The participants were 69 infants, 37 American and 32 Taiwanese. Of the 37 American infants, 19 were in the age range of 6–8 months (mean age at test=7.2 months; range =6.8–7.7 months; boys=nine, girls=ten), and 18 were in the age range of 10–12 months (mean age at test=10.9 months; range=10.8–11.1 months; boys=nine, girls=nine). Of 32 Taiwanese babies, half of them were 6–8 months (mean age at test=7.4 months; range=6.9–8.4 months; boys=eight, girls=eight), and half were in the age range of 10–12 months (mean age at test=11.3 months; range =10.7–12.2 months; boys=11, girls=five). An additional 19 infants failed to complete testing due to an inability to pass the training (17), an equipment failure (1), or a failure to return for all of the required sessions (1). Infants who failed to pass the training did not differ by age or language: five 6–8-month-old American infants (drop-out rate=20.8%), four 6–8-month-old Taiwanese infants (drop-out rate =20.0%), two 10–12-month-old American infants (drop-out rate=10.0%), and six 10–12-month-old Taiwanese infants (drop-out rate=27.3%) failed to meet the criterion. Results of Fisher's exact probability test on the drop-out rate indi-

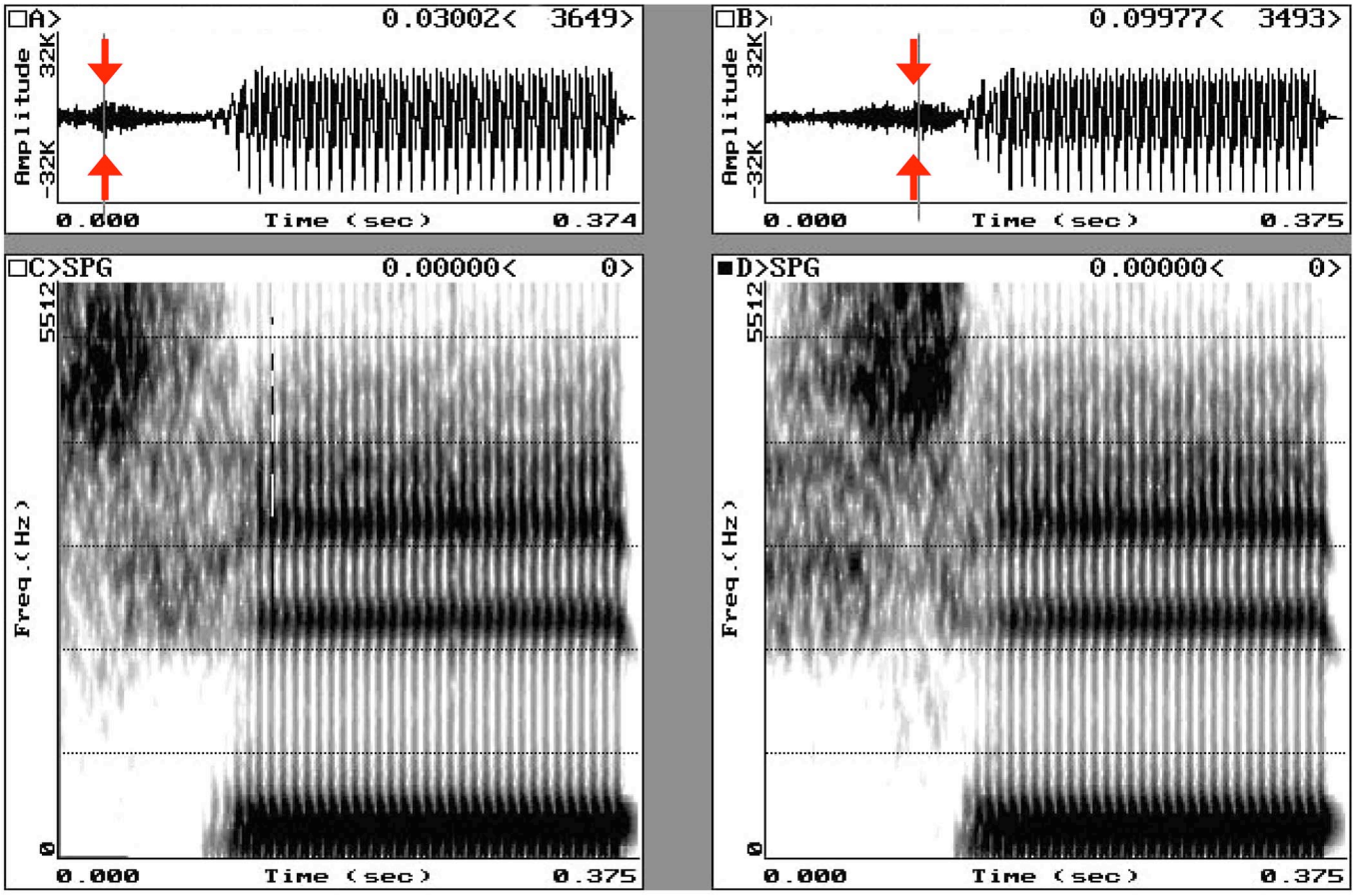


FIG. 2. (Color online) Wave forms (top panel) and spectrograms (bottom panel) of affricate (left) and fricative (right) stimuli in Exp. 2. Arrows indicate the location of the maximum amplitude in frication noise.

cated neither the age nor language effect reached significance. Pre-established criteria for inclusion in the study were that infants had no known visual or auditory deficits, were full term (born ± 14 days from due date), had uncomplicated deliveries, were normal birth weight (6–10 lbs), were developing normally, and that members of their immediate families had no history of hearing loss. Parents were paid $30 for completing the experiment.

American infants were recruited through the database of names contained in the Infant Studies Subject Pool (ISSP) at the University of Washington. Taiwanese infants were recruited either through listings of names on the House Registry of the Lin-Ya Area, Kaohsuing City, Taiwan, or from notices soliciting families' participation which were displayed at the Kaohsiung Chung-Kung Children's Hospital. Although Taiwan is a multi-lingual society, Mandarin Chinese is the most dominant language in homes. A Mandarin-dominant (or only) language environment was verified for the Taiwanese infants through a language background questionnaire in Chinese that was administered to the caregiver before the study began. The criteria of judging Mandarin the dominant language were: (1) parents and infant's caretakers used Mandarin when addressing their infants, and (2) infants heard Mandarin since birth.

2. Stimuli

The stimuli were pair 1 from Experiment 1, which consisted of computer synthesized male tokens of Mandarin Chinese alveolo-palatal /tɕʰi/ and /ɕi/ syllables. Figure 2 illustrates wave forms and spectrograms for these stimuli (see Table I for acoustic differences). Tokens were played to infants at a comfortable listening level of approximately 65 dBA. Mandarin-speaking adults (n=14) tested in the same conditioned head-turn procedure as infants discriminated this contrast at 95.21% (SD=5.84).

3. Apparatus

Stimuli were presented using a digital signal processor (CAC Bullet III) controlled by a portable computer (Dell Inspiron 3500). The computer was also used to record infant head-turn responses. The sounds were reproduced with 11.025 k 16 bit samples per second, and were low-pass filtered at a 5.5 kHz cutoff frequency. Stimuli were amplified (Shure FP42) and delivered to subjects in an adjoining sound-treated test room via a loudspeaker (Boston Acoustics CR7). Parents and experimenters wore headphones (David Clark H3050) and listened to masking music during the tests so they could not distinguish between the stimuli presented to infants. Infants' responses were monitored in the control room via use of a closed-circuit camera (RCA TC7011) and a video monitor. Both American and Taiwanese infants were tested with the same apparatus, experimenters, and test protocols to carefully control the experimental task between the two different countries.

4. Test suite

The test suite consisted of two rooms. In the test room, an infant was held on its parent's lap, facing forward while the assistant was seated at a 60° angle to the infant's right side. An Assistant maintained the infant's attention by manipulating a series of engaging, silent toys to bring the child's gaze to midline (straight ahead of the infant). A bank of two visual reinforcers, located at a 60° angle to the infant's left side, each consisted of a dark Plexiglas box (13 in. × 13 in. × 13 in.) containing a commercially available mechanical toy (e.g., a bear pounding a drum). The toys were not visible until they were activated and lights mounted inside the box were illuminated. The visual reinforcers were placed on either side of the loudspeaker, and were at eye level for the infant. A camera, located in front of the infant, but hidden from view by a curtain with a hole cut for the lens, fed an image of the test room to the adjoining control room, where the Experimenter observed the infant's behavior. In all phases of training and testing, trials are initiated by the Assistant. The Experimenter, who cannot hear the stimuli presented during trials, and who is unaware of the type of trial selected automatically by the compute, indicates infants' head turns by pressing a computer key.

5. Procedure

The conditioned head-turn technique was used to assess infants' discrimination abilities (Kuhl, 1985; Werker *et al.*, 1997). The "background" speech sound, /ɕi/, was repeated once every 2 s (Interstimulus interval, ISI=1625 ms). Infants first were trained to produce a head turn for visual reinforcement whenever the background speech sound was changed to the "target" speech sound, /tɕʰi/. Only one direction of stimulus change was tested because potential directional effects in infant testing (e.g., Kuhl *et al.*, 2006; Polka and Bohn, 1996, 2003; Polka *et al.*, 2001) might complicate the developmental pattern. The experimental protocol required a two-step Conditioning Phase: (1) an intensity cue was initially added to assist infants in detecting the sound change and (2) the intensity cue was eliminated to establish infants' abilities to discriminate the contrast in the absence of a loudness difference. During Conditioning, all trials involve a change in the stimulus from the background to the target stimulus (Change Trials). After this two-step Conditioning Phase, a Test Phase was initiated; during the Test Phase, an equal number of Change and Control trials are run in random order. All phases of the experiment were under computer control. The same basic procedure has been used in previous infant studies in this laboratory (Liu, Kuhl, and Tsao, 2003; Kuhl, Tsao, and Liu, 2003; Tsao, Liu, and Kuhl, 2004).

During the first step of Conditioning, infants were trained to associate presentation of the target speech sound with the activation of the visual reinforcers. The Assistant initiated a trial when infants appear ready (focused on the toys held by the assistant). Then, the target sound interrupted the repetitive presentation of the background speech sound, and was presented at a level 4 dBA higher than the background speech sound. The target stimulus was presented three times, and infants had to produce a head turn in response to the sound change within this 6 s period. The louder target speech sound facilitated infant learning to turn their heads and associate the sound change with the presentation of the reinforcer. When the infant produced a head turn on two consecutive trials, the infant proceeded to the second Conditioning phase, during which the target sound is presented at the same level as the background sound; infants can only use the phonetic difference between sounds as a cue. Infants must produce three consecutive head turns within 30 training trials to advance to the Test Phase. During the Conditioning phase, infants were cued to produce a head turn by the activation of the reinforcers near the end of the trial if they had not produced a head turn.

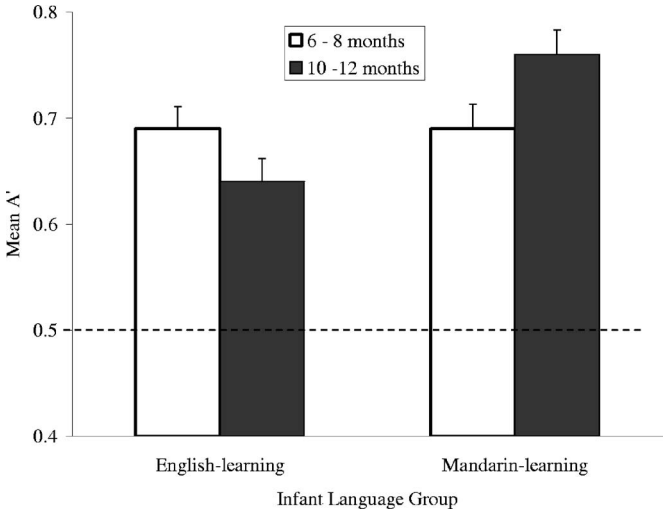


FIG. 3. Phonetic discrimination of 6–8 and 10–12-month-old English- and Mandarin-learning infants on Mandarin contrast (dotted line shows the chance level in Exp 2; *SE* in error bars).

The Test Phase consisted of 30 trials, an equal number of Change and Control (no-change) Trials, presented in random order. Infants were tested in 20 min sessions on consecutive days, when possible, but all completed testing within one week. Most infants were tested in two days (Conditioning on Day 1 and then test on Day 2), but up to three sessions were allowed to complete the test. When infants returned on the second day for testing, three Conditioning trials (+ intensity cue) were used to refresh the infants on the experimental procedure before the test trails began. Infants who failed to pass the two-phase Conditioning in two sessions were eliminated from the experiment.

B. Results and Discussion

The results provided information about the time course of development of native- and non-native phonetic perception for affricate-fricative consonants. Figure 3 provides the A′ scores as a function of age and language group for the Taiwanese and American infants. A′ is a distribution-free sensitivity measure that takes both "hits" and "false alarms" into consideration to provide an estimate of an infant's accuracy in detecting the sound change. This measure (range=0 −1, chance level=0.5) is similar to d' (Grier, 1971) and has been used in infant speech discrimination studies (e.g., Anderson, Morgan, and White, 2003; Polka *et al.*, 2001).

As shown in Fig. 3, increasing age affects discrimination capacity very differently in the two language groups. Mandarin-learning infants show an increase in performance over time while English-learning infants show a decline. The difference in the effect of age on discrimination performance in the two groups was tested by analysis of variance. Taiwanese infants demonstrated a substantial improvement in their performance on the discrimination task, 6–8 month olds ($M=0.69$, $SD=0.08$) and 10–12 month olds ($M=0.76$, $SD=0.08$), $F(1,30)=5.41$, $p=0.027$. In contrast, American infants showed a trend toward declining sensitivity, younger infants ($M=0.69$, $SD=0.10$) and older infants ($M=0.64$, $SD=0.10$), $F(1,35)=3.09$, $p=0.088$. Performance of the two language groups at the younger age was nonsignificant, $F(1,33)<1$, and both above chance level at $p<0.001$, one-sample t test, Taiwanese infants, $t(15)=9.81$; American infants, $t(18)=8.37$. In contrast, the performance of older infants was significantly different for the two language groups, $F(1,32)=14.33$, $p<0.001$; performance of both older infant groups is significantly above chance level at $p<0.001$, one-sample t test, Taiwanese infants, $t(15)=12.41$; American infants, $t(17)=6.21$. A two-way ANOVA (Language group X Age) on the A′ was conducted. The main effect of age was not significant, $F(1,65)<1$, though that of language did reach significance, $F(1,65)=6.28$, $p=0.015$. The age x language interaction was highly significant, $F(1,65)=7.88$, $p=0.007$, and indicated a divergent trend for developing native and non-native language perception during the first year of life.

The results show a divergent trend in the development of native and non-native phonetic perception. Infants listening to a native language contrast show a significant improvement between 6 and 12 months, as observed recently by Kuhl *et al.* (2006), while those listening to a non-native contrast show a pattern of decline in the ability to discern differences between speech sounds of the non-native language, though one that does not reach significance. The pattern of facilitation in the first year, seen for /r-l/ in American infants (Kuhl *et al.*, 2006) and in the present study for affricate-fricative contrasts in Taiwanese infants, suggests that experience with native language leads infants to develop increased sensitivity to native contrasts.

IV. EXPERIMENT 3: PHONETIC DISCRIMINATION OF ENGLISH-LEARNING INFANTS ON THE ENGLISH AFFRICATE VS FRICATIVE CONTRAST

Experiment 2 demonstrated that the Mandarin affricate-fricative contrast shows an increase between 6–8 and 10–12 months for Mandarin-learning infants. Experiment 3 was designed to extend these findings to American infants listening to their native-language affricate-fricative (palato-alveolar) contrast.

A. Method

1. Participants

The participants were 17 American infants aged 6–8 months (mean age at test=6.9 months; range =6.8–7.7 months; boys=ten, girls=seven) and 19 American infants aged 10–12 months (mean age at test=10.7 months; range=10.5–11.1 months; boys=ten, girls=nine). Infants were recruited through the database of the Infant Studies Subject Pool (ISSP) at the University of Washington. An additional 17 infants failed to complete testing due to inability to pass the Conditioning Phase (11), or failure to return for all of the required sessions (six). Of the 11 infants who failed to pass Conditioning, the drop-out rate was not significantly different between age groups, younger infants ($n=5$, drop-out rate=22.7%) and older infants ($n=6$, drop-out rate=24.0%), Fisher's exact test, $p>0.1$. The inclusion and exclusion criteria of subject selection were the same as in Experiment 2. Parents were paid $30 when their infants completed the experiment.

2. Stimuli

The stimuli were computer synthesized tokens of English palato-alveolar affricate /tʃi/ and fractive /ʃi/ syllables created using a male voice. They were matched in all acoustic details other than the temporal features during the initial portion of the consonants. Previous studies showed that frication duration is the primary acoustic cue for the affricate-fricative distinction in English (Kluender and Walsh, 1992).

Frication duration of the affricate and fricative tokens were 80 and 180 ms, respectively. The amplitude rise time was 30 ms shorter than the frication duration to generate more natural-sounding speech tokens for English speakers. Therefore, both amplitude rise time and frication duration differed in the two tokens. The spectral peak frequency of this pair of tokens was 2800 Hz. Acoustic parameters of the vowel /i/ were exactly the same as in the previous experiments. The duration of the vowel was 245 ms. The two stimuli were judged to be good instances of English native categories and were easily discriminated by English-speaking adults in a pilot study. Tokens were equalized in rms amplitude and were played to infants at a comfortable listening level of approximately 65 dBA.

3. Apparatus and Procedure

The procedure and apparatus were identical to that used to test perceptual development of infants' speech discrimination on native and non-native contrasts in Experiment 2. The background syllable was fricative /ʃi/ and the target syllable was affricate /tʃi/.

B. Results and Discussion

Experiment 3 tested two specific predictions. First, based on the results of Experiment 2, English-learning infants' performance on the native palato-alveolar affricate-fricative distinction was hypothesized to show an increase with age. Second, performance of 11-month-old English-learning infants was expected to exceed that shown for the nonnative affricate-fricative contrast tested in Experiment 2.

The results of this experiment support these two predictions. Figure 4 illustrates the A′ scores of the two English-learning groups on the discrimination of English /tʃ/ vs /ʃ/. Older English-learning infants were more sensitive ($M=0.78$, $SD=0.05$) than younger infants ($M=0.70$, $SD=0.11$)

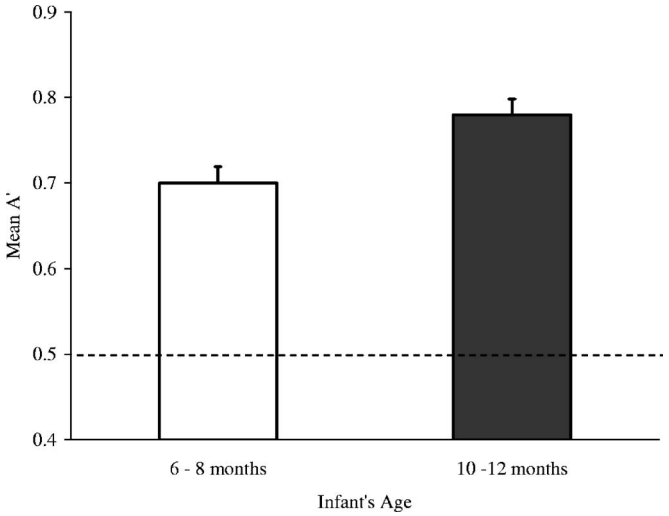


FIG. 4. Phonetic discrimination of 6–8 and 10–12-month-old English-learning infants on English contrast (dotted line shows the chance level in Exp 3; *SE* in error bars).

to the differences in the two phonemes, $F(1,34)=6.39$, $p=0.016$. In addition, English-learning 11-month-old infants were more accurate in detecting the acoustic differences between the native affricate-fricative contrast than they were in detecting the affricate-fricative non-native contrast in Experiment 2, $F(1,35)=30.35$, $p<0.001$. The younger English-learning infants performed similarly for both English and Mandarin contrasts, $F(1,34)<1$. Combining the results of young and old English-learning infants on the discrimination of both native (Experiment 3, English palato-alveolar affricate-fricative) and non-native (Experiment 2, Mandarin alveolo-palatal affricate-fricative) contrasts, a two-way ANOVA (Age × Phonetic contrast) reveals a significant contrast effect, $F(1,69)=11.64$, $p<0.001$, and nonsignificant age effect, $F(1,69)<1$. The interaction of Age and Phonetic contrast is significant, $F(1,69)=8.67$, $p=0.004$. Therefore, the results demonstrate the divergent trend of perceptual development for native and non-native contrasts. In addition, 11-month-old English-learning infants performed at the same level on the discrimination of native affricate vs fricative contrast as Mandarin-learning infants at the same age discriminating their native affricate-fricative contrast. No significant difference was evident in a one-way ANOVA on infant language group, $F(1,33)<1$. This suggests that the developmental pace of perceiving affricate and fricative contrasts is similar for infants raised in two very different language environments.

V. GENERAL DISCUSSION

The effects of language experience on developmental change in speech perception were examined using a cross-language design with both adults and infants on the perception of native and non-native speech sounds. The results of Experiment 1 demonstrate the impact of language experience on phonetic discrimination in adults. Mandarin-speaking adults are more accurate than English-speaking adults in distinguishing the acoustic differences between Mandarin alveolo palatal fricative and affricate consonants.

English and Mandarin-learning infants were compared in Experiment 2 to examine sensitivity change to native and non-native contrasts during the second half of the first year of life using a Mandarin affricate vs fricative contrast from Experiment 1. The results of Experiment 2 revealed a divergent trend in perceptual development for native and non-native contrasts and are consistent with the view that infants develop language-specific processing around their first birthday. The results of Experiment 2 provide support for the idea that native contrasts show a pattern of facilitation over time; perceptual sensitivity of 10–12-month-old infants improved over that seen in 6–8-month-old infants. This improvement for native contrasts is not mirrored in performance on non-native contrasts. Older infants perform less accurately than younger infants on the discrimination of a non-native contrast. Furthermore, Experiment 3 buttresses these findings on facilitation for native-language contrasts by showing that 10–12-month-old English-learning infants are significantly more sensitive than 6–8-month-old English-learning infants in discriminating their native English affricate-fricative distinction. This pattern of facilitation for native-language phonetic learning in the first year has been suggested by neural studies (Cheour, 1998; Rivera-Gaxiola *et al.*, 2005), and clearly shown in a behavioral test for the American English /r-l/ contrast (Kuhl *et al.*, 2006). Thus the facilitation pattern is seen for liquids (Kuhl *et al.*, 2006) and the two affricate-fricatives tested in the present experiments; a third contrast shows facilitation, though later in development (Polka *et al.*, 2001; Sundara *et al.*, 2006). One study observed a decline in a difficult fricative (/s-z/) native-language contrast (Best and McRoberts, 2003).

What accounts for the variance across phonetic contrasts in the pattern of facilitation or decline seen developmentally? The timing of developmental change could be due to the amount of experience with specific phonetic units in the native language. The fact that the frequency of occurrence of native consonants is not equally distributed in language input to infants suggests that different patterns of development may exist for different consonants. It is estimated that the coronal stops are more frequent than the dorsal stops in the English infant-directed speech (Anderson *et al.*, 2003). For non-native perception, one recent study demonstrated that 8.5-month-old English-learning infants performed less accurately distinguishing the non-native coronal stops than the dorsal stops (Anderson *et al.*, 2003). Infants utilize the distributional probabilities of phonetic features in the native language to perceive speech sounds. For example, Maye, Werker, and Gerken (2002) examined the impact of distributional properties of speech sounds in infants by varying these properties in 2 min exposures to a series of eight stimuli from a voice-onset time continuum; infants' discrimination abilities were improved by "bimodally" distributed experience (see Maye and Weiss, 2003, for discussion). Kuhl *et al.* (1992) examined the impact of distributional properties by testing perception in 6-month-old infants from two countries whose experience with natural language provided them with vowels whose distributional properties differed in language input; infants' vowel categorization abilities were enhanced for native-language vowel categories (see Kuhl, 2004 for dis-

cussion). These studies suggest a statistical basis of phonetic learning and predict that listening experience with native language enhances the perception of native-language phonetic categories while reducing sensitivity to non-native contrasts. If infants are sensitive to distributional properties of phonetic categories when listening to native language, as suggested by previous research, the possibility exists that the frequency of consonants and their distribution in ambient speech, as well as their acoustic properties, may be shown to affect perceptual development of different native contrasts.

The results of infant experiments and previous studies clearly demonstrate that infants are born with language-general processing abilities (e.g., Best and McRoberts, 2003; Kuhl, 2004; Werker and Tees, 1984), and the results of this study further demonstrate the divergence in development of the perception of native and non-native phonetic contrasts during the first year of life. What is needed is a hypothesis that explains all the patterns seen to date in developmental phonetic perception data. What determines when facilitation for native contrasts will occur; and what determines which non-native contrasts will decline and which do not? Various hypotheses have been developed: Best and McRoberts (2003) hypothesize knowledge of articulatory organs, Maye, Werker and Gerken (2002) argue that a distributional frequency hypothesis accounts for the data, and Kuhl and her colleagues argue that a combination of motherese and distributional frequency accounts for native-language learning with social and cognitive factors playing a critical role (Liu *et al.*, 2003; Kuhl *et al.*, 2003; Kuhl *et al.*, in press). Additional contrasts will need to be tested to examine why some contrasts show developmental change prior to 12 months, while others do not; such studies will also allow comparisons among theories.

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

Aslin, R. N., Werker, J. F., and Morgan, J. L. (**2002**). "Innate phonetic boundaries revisited (L)," J. Acoust. Soc. Am. **112**, 1257–1260.

Best, C. T. (**1995**). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 171–204.

Best, C. T., and McRoberts, G. W. (**2003**). "Infant perception of non-native consonant contrasts that adults assimilate in different ways," Lang Speech **46**, 183–216.

Best, C. T., McRoberts, G. W., and Sithole, N. N. (**1988**). "The phonological basis of perceptual loss for nonnative contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants," J. Exp. Psychol. Hum. Percept. Perform. **14**, 345–360.

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

Cheour, M. (**1998**). "Development of language-specific phoneme representations in the infant brain," Nat. Neurosci. **1**, 351–353.

Chomsky, N., and Halle, M. (**1968**). *Sound Pattern of English* (Harper and Row, New York).

Cutting, J. E., and Rosner, B. S. (**1974**). "Categories and boundaries in speech and music," Percept. Psychophys. **16**, 564–570.

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

Eilers, R. E., Wilson, W. R., and Moore, J. M. (**1977**). "Developmental changes in speech discrimination in infants," J. Speech Hear. Res. **20**, 766–780.

Goto, H. (**1971**). "Auditory perception by normal Japanese adults of the sounds "L" and "R"," Neuropsychologia **9**, 317–323.

Grier, J. B. (**1971**). "Nonparametric indexes for sensitivity and bias: Computing formulas," Psychol. Bull. **75**, 424–429.

Hedrick, M. (**1997**). "Effect of acoustic cues on labeling fricatives and affricates," J. Speech Lang. Hear. Res. **40**, 925–938.

Howell, P., and Rosen, S. (**1983**). "Perception of rise time and explanations of the affricate/fricative contrast," Speech Commun. **2**, 164–166.

Huang, K. I. (**1996**). "Articulation, acoustic features, and perception of vowels," in *Fundamentals of Speech Pathology Vol. 2*, edited by C. H. Tseng (Psychology, Taipei, Taiwan), pp. 1–31.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Kluender, K. R., and Walsh, M. A. (**1992**). "Amplitude rise time and the perception of the voiceless affricate/fricative distinction," Percept. Psychophys. **51**, 328–333.

Kuhl, P. K. (**1985**). "Methods in the study of infant speech perception," in *Measurement of Audition and Vision in the First Year of Life: A Methodological Overview*, edited by G. Gottlieb and N. A. Krasnegor (Ablex, Norwood, NJ), pp. 223–251.

Kuhl, P. K. (**2004**). "Early language acquisition: Cracking the speech code," Nat. Rev. Neurosci. **5**, 831–843.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (**in press**). "Developmental phonetic perception: Native language magnet theory expanded (NLM-e)," Philos. Trans. R. Soc. London, Ser. B

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (**2006**). "Infants show a facilitation effect for native language phonetic perception between 6 and 12 months," Dev. Sci. **9**, F13–F21.

Kuhl, P. K., Tsao, F. M., and Liu, H. M. (**2003**). "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning," Proc. Natl. Acad. Sci. U.S.A. **100**, 9096–9101.

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

Ladefoged, P., and Maddieson, I. (**1995**). *The Sounds of the World's Languages* (Blackwell, Malden, MA).

Lasky, R. E., Syrdal-Lasky, A., and Klein, R. E. (**1975**). "VOT discrimination by four to six and a half month old infants from Spanish environments," J. Exp. Child Psychol. **20**, 215–225.

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

Liu, H. M., Tseng, C. H., and Tsao, F. M. (**2000**). "Perceptual and acoustic analysis of speech intelligibility in Mandarin-speaking young adults with cerebral palsy," Clin. Linguist. Phon. **14**, 447–464.

Liu, H. M. (**1996**). "Perceptual and acoustic analysis of speech intelligibility in young adults with cerebral palsy," Unpublished Master's Thesis, National Kaoshuing Normal University, Taiwan.

Maye, J., and Weiss, D. J. (**2003**). "Statistical cues facilitate infants' discrimination of difficult phonetic contrasts," *Proceedings of the 27th Annual Boston University Conference on Language Development*, pp. 508–518.

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

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., and Fujimura, O. (**1975**). "An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English," Percept. Psychophys. **18**, 331–340.

Nittrouer, S. (**2001**). "Challenging the notion of innate phonetic boundaries," J. Acoust. Soc. Am. **110**, 1598–1605.

Polka, L., and Bohn, O. S. (**1996**). "Cross-language comparison of vowel perception in English-learning and German-learning infants," J. Acoust. Soc. Am. **100**, 577–592.

Polka, L., and Bohn, O. S. (**2003**). "Asymmetries in vowel perception," Speech Commun. **41**, 221–231.

Polka, L., Colantonio, C., and Sundara, M. (**2001**). "A cross-language comparison of /d/-/ð/ perception: Evidence for a new developmental pattern," J. Acoust. Soc. Am. **109**, 2190–2201.

Rivera-Gaxiola, M., Silva-Pereyra, J., and Kuhl, P. K. (**2005**). "Brain potentials to native- and non-native speech contrasts in seven and eleven-month-old American infants," Dev. Sci. **8**, 167–172.

Sheldon, A., and Strange, W. (**1982**). "The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception," Appl. Psycholinguist. **3**, 234–261.

Strange, W., and Jenkins, J. (**1978**). "Role of linguistic experience in the perception of speech," in *Perception and Experience*, edited by R. D. Walk and H. L. Pick (Plenum, New York), 125–169.
Streeter, L. A. (**1976**). "Language perception of 2-month-old infants shows effects of both innate mechanisms and experience," Nature (London) **259**, 39–41.
Sundara, M., Polka, L., and Genesee, F. (**2006**). "Language-experience facilitates discrimination of /d-ð/; in monolingual and bilingual acquisition of English," Cognition **100**, 369–388.
Trehub, S. E. (**1973**). "Infants' sensitivity to vowel and tonal contrasts," Dev. Psychol. **9**, 91–96.
Trehub, S. E. (**1976**). "The discrimination of foreign speech contrasts by infants and adults," Child Dev. **47**, 466–472.
Tsao, F. M., Liu, H. M., and Kuhl, P. K. (**2004**). "Speech perception in infancy predicts language development in the second year of life: a longitudinal study," Child Dev. **75**, 1067–1084.
Werker, J. F., and Tees, R. C. (**1984**). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," Infant Behav. Dev. **7**, 49–63.
Werker, J. F., Gilbert, J. H. V., Humphery, K., and Tees, R. C. (**1981**). "Developmental aspects of cross-language speech perception," Child Dev. **52**, 349–355.
Werker, J. F., Polka, L., and Pegg, J. E. (**1997**). "The conditioned head turn procedure as a method for testing infant speech perception," Early Dev. Parenting **6**, 171–178.
Werker, J. F., and Curtin, S. (**2005**). "PRIMIR: A developmental framework of infant speech processing," Lang. Learn. Dev. **1**, 197–234.