Perception of Similar and Dissimilar Lexical Tones by Non-Tone-Learning Infants

Rushen Shi

Université du Québec à Montréal

Elsa Santos

Université du Québec à Montréal

Jun Gao

Institute of Linguistics, Chinese Academy of Social Sciences

Aijun Li

Institute of Linguistics, Chinese Academy of Social Sciences

Keywords: lexical tone, perceptual development, phonetic categories, infancy, psychoacoustic processing

Corresponding author:

Rushen Shi

Département de Psychologie

Université du Québec à Montréal

C.P. 8888, Succursale Centre-Ville

Montréal, Québec, Canada H3C 3P8

Email: shi.rushen@uqam.ca

Tel: 1(514) 987 3000 ext. 7604

Abstract

This study examined the perception of Mandarin lexical tones by non-tone-learning infants. We tested French-learning 4-, 8- and 11-month-olds on Tone 2 (rising) versus Tone 3 (dipping), the most acoustically similar contrast in Mandarin, and on Tone 1 (high) versus Tone 4 (falling), an acoustically more dissimilar contrast. We hypothesized that sensitivity should decline with age for the similar contrast, and the dissimilar contrast should remain more discriminable. Infants were habituated to one tone, and tested with the same tone versus the contrastive tone. Results showed that the dissimilar T1-T4 contrast was consistently discriminated by all three ages, whereas the discrimination of the similar T2-T3 contrast revealed a tendency to decline over the ages. These results suggest that relative to consonants and vowels, perceptual sensitivity to lexical tones remains stronger in non-tone-learning infants. Discrimination of lexical tones is affected by both acoustic salience and perceptual attunement to the native-language phonology.

Keywords: lexical tone, perceptual development, phonetic categories, infancy, psychoacoustic processing

Infants begin life with remarkable sensitivity to speech sounds, discriminating both native and non-native phonetic contrasts. During the second half of the first year of life, infants become attuned to their native-language sound system, maintaining or improving their discrimination of contrasts in their language while declining in their discrimination of non-native contrasts. This development was first shown for consonants (e.g., Werker & Tees, 1984), and subsequently for vowels (e.g., Kuhl, et al., 2006; Polka & Werker, 1994).

Besides consonants and vowels, many languages use lexical tones for distinguishing word meaning. Lexical tones are typically carried by the syllable or the vowel. Fundamental frequency (F0, the fundamental rate of vocal fold vibration, related to pitch perception) is the primary acoustic correlate to tones, although duration and amplitude also serve as cues. A limited number of studies have investigated infants' perception of lexical tones, showing native-language influences during the first year of life: Tone-learning infants maintain sensitivity to lexical tones, whereas non-tone-learning infants decline in their tone sensitivity. In Mattock and Burnham (2006) English-learning infants declined in their discrimination of the Thai low versus rising tones from 6 to 9 months of age, unlike Chinese infants, who discriminated the Thai contrast at both ages. French-learning infants showed the same pattern of discrimination decline for the Thai tones, as did English-learning infants (Mattock, Molnar, Polka, & Burnham, 2008). Yeung, Chen and Werker (2013) tested Cantonese tones (mid versus high-rising), and found that Cantonese- and Mandarin-learning infants showed tonal preferences reflecting the specific patterns of their native languages, while English-learning infants declined in their sensitivity to these Cantonese tones from 4 to 9 months of age. These findings suggest a development from language-universal to language-specific perception during the first year of life.

However, different findings of tone perception by non-tone-learning children have also been reported, indicating continuing sensitivity to lexical tones long after infancy. Liu and Kager (2014) tested Dutch-learning 5-18-month-olds' discrimination of high (T1) versus falling (T4) tones in Mandarin, and found that infants across ages discriminated the contrast. English-learning 19-month-olds distinguished the Mandarin rising (T2) and falling (T4) tones in an auditory discrimination task, although they failed to encode the contrast lexically in a word-object association task (Hay, Graf Estes, Wang, & Saffran, 2015). In Singh, Hui, Chan, and Golinkoff, (2014), on the other hand, English-learning toddlers aged 18 months encoded this contrast during word learning. Success in discriminating Mandarin tonal contrasts involving T1 (high), T2 (rising) and T3 (dipping) was even obtained in French-speaking 9-17-year-olds, whose performance matched that of their Mandarin-speaking peers (Pierce, Kline, Chen, Delcencerie, & Genesee, 2014).

There is evidence that acoustical salience affects perceptual development. For example, Filipino infants' discrimination of syllable-initial alveolar and velar nasals (n-ng), an acoustically similar contrast that occurs syllable-initially in their native language, does not emerge until approximately 10-12 months of age, whereas English-learning infants aged 4-12 months consistently failed to discriminate this contrast (Narayan, Werker, & Beddor, 2010). Regarding lexical tones, Tsao (2008) found that Mandarin-learning 10-12-month-olds discriminated acoustically dissimilar tones (high-T1 versus dipping-T3) in Mandarin better than similar tones (rising-T2 versus dipping-T3; rising-T2 versus falling-T4). Acoustically dissimilar tones seem resistant to perceptual decline for non-tone-learning infants, such as the Mandarin T1-T4 discrimination (Liu & Kager, 2014).

To better understand the role of acoustic salience, we directly tested non-tone-learning infants' perception of two Mandarin lexical tone contrasts, one acoustically similar (T2-rising vs.

T3-dipping), and another more dissimilar (T1-high vs. T4-falling). We examined how the perception evolved in French-learning infants, whose native language has no lexical tones.

The T2-T3 contrast is widely considered the most acoustically similar in Mandarin (e.g., Li & Chen, 2015; Hao, 2012; Shen & Lin, 1991). In citation form the two tones have similar pitch contours, both starting low and ending higher, although T3 first dips and then moves up less high. Native Mandarin-learning babies can discriminate this contrast across the first year of life (Shi, Gao, Achim, & Li, in press); nevertheless, this contrast is discriminated less well than other tonal contrasts (Tsao, 2008). Mandarin-learning toddlers and young children also make more confusion errors with this contrast during word recognition (Shi, et al., in press; Wong, Schwartz, & Jenkins 2005). Even Mandarin-speaking adults, who are generally perfect in perceiving Mandarin tones, sometimes make identification errors of T2 and T3 (Shen & Lin, 1991; Huang & Johnson, 2010). This contrast is more difficult than other tonal contrasts for second language learners (e.g., Hao, 2012). In Lee, Tao and Bond (2010) English-speaking adults made more errors when identifying T2, confusing it as T3. Non-tone-learning infants' perception of the T2-T3 contrast remains unclear.

The T1-T4 contrast is acoustically more dissimilar. T1 has a simple high-flat F0 trajectory, and T4 has a high-to-low trajectory. They are distinct types of tones (level versus contour), unlike T2 and T3, which are both contour tones and complex (Yip, 2002). Mandarin-learning 3-year-olds showed better processing of the T1-T4 contrast than the T2-T3 contrast during word comprehension (e.g., Singh, Tan, & Wewalaarachi, in press; Wong, Schwartz, & Jenkins 2005). T1-T4 is less difficult than T2-T3 for second language adult learners (e.g., Hao, 2012), although So and Best (2010)'s English-listeners showed comparable performance for these two contrasts.

In light of the previous findings, we hypothesized that for non-tone-learning infants, discrimination of similar tones should decline, and dissimilar tones should remain discriminable, that is, more resistant to decline, during the first year of life.

Method

Participants

Participants were 144 French-learning 4-month-olds (n=48; *M* age: 0;4;28; range: 0;4;01 – 0;5;28; 22 boys), 8-month-olds (n=48; *M* age: 0;8;27; range: 0;7;27 – 0;9;20; 23 boys), and 11 months (n=48; *M* age: 0;11;15; range: 0;11;00 – 0;12;20; 20 boys). Thirty-five other infants were excluded from analyses due to fussiness or crying (28), falling asleep (1), looking time too short (2), uninterested in the task (1), and technical errors (3).

Speech stimuli

The tone-bearing unit in Mandarin is the syllable (Xu & Wang, 2001). We used the syllable *can* (/ts^{h}an/) in T2 (rising) and T3 (dipping), respectively. A female Mandarin speaker recorded many exemplars of the tones (sampling frequency 22050Hz, bit rate 16 bits). To ensure that T2 and T3 stimuli were maximally similar, we selected productions of the two tones with comparable duration values and equated their overall amplitude, while keeping their F0 intact. Thus, the two tones were more similar than usual. The final stimuli were 13 tokens for each tone. The duration ranged from 631ms to 806ms for T2, and from 630ms to 802ms for T3. These stimuli were the same as those used in Shi et al. (in press), in which detailed acoustic measures were reported. These tokens were presented to 6 Mandarin-speaking adults, who identified the tone of each token correctly.

T1 (high) and T4 (falling) stimuli (on the syllable *kui*, /k^{hw}ei/) were recorded in the same way. Final stimuli were 13 tokens for each tone, for which the same adult listeners correctly identified the tones. The duration ranged from 503ms to 645ms for T1, and from 464ms to

536ms for T4. Thus, duration was an additional cue making T1 and T4 distinct. Appendices A and B show T2-T3 and T1-T4 example stimuli, respectively.

Procedure and design

The procedure and design were the same as those in Shi et al. (in press). In an acoustic chamber the child sat on the parent's lap, facing an LCD monitor. Auditory stimuli were played through loudspeakers adjacent to the monitor. Parents heard masking music from headphones. A camera below the monitor filmed the child and sent simultaneous video of the child to another monitor outside, where a researcher, blind to all stimuli, observed and coded the child's looking behavior. The experiment was executed by a computer program, which also recorded the child's eye responses. Each trial started when the child looked at the monitor, and terminated when s/he looked away for 2 seconds or more or when the maximum trial length (21s) was reached. A colorful checkerboard was displayed during each trial. An attention-getter (a jumping star with bird calls) appeared between trials to attract the child back to the monitor.

In a cross-sectional design each infant participated in one experimental session containing two phases, habituation and test. There were 24 infants per contrast (T2-T3; T1-T4) per age (4, 8, 11 months), in total 144 infants. For each contrast group (e.g., T2-T3) within each age, infants were randomly assigned to one of the two habituation conditions: half habituated to one tone in a contrast (e.g., T2), and the other half to the other tone (e.g., T3). While presenting the habituation tone, the experimental program automatically calculated the infant's average looking time for every three consecutive trials and compared it with the average looking time of the first three trials. When the looking time of a later three-trial time window dropped to 50% or lower of the first time window, the habituation criterion was reached, and the test phase began. Each infant heard two types of test trials, Same (the same tone as in the habituation) and Different (the contrasting tone). For example, after T2 habituation, the Same trial type was T2, and the Different trial type was T3 during the test phase; the reverse was the case for infants habituated to T3, with Same being T3, and Different being T2 in the test phase. The design details are in Appendix C.

Seven tokens of each tone served as the habituation stimuli, and the six remaining tokens as the test stimuli. Thus, both the Same and Different test trials presented new tokens. The first test trial was either Same or Different, counterbalanced across infants. Tokens within each trial were presented randomly, separated by a 1sec pause.

After the test phase a post-trial presented distinct visual (a cat picture) and speech (repeated presentations of a distinct syllable, *mao*-T1) stimuli, marking the end of the experiment. Data analysis showed that looking time increased in the post-trial relative to the last test trial ($t(143)=-19.89$; $p<.0005$, 2-tailed), indicating that infants were on task throughout the experiment.

If infants discriminated the tones in a contrast, they should look longer in the Different than the Same test trials.

Results

Looking times in the Same and Different test trials were analyzed. We first checked if there were any effects of habituation tone on the test results. For infants tested with the T2-T3 contrast, an ANOVA with Test Trial Type (Same vs. Different) as the within-subject factor and Habituation Tone (T2 or T3) as the between-subject factor, showed no interaction of the two factors, $F(1,66)=1.286$; $p=.261$. Another ANOVA for the T1-T4 infants also showed no interaction of Test Trial Type and Habituation Tone (T1 or T4), $F(1,66)=2.289$; $p=.135$. That is, results in the test phase were comparable for infants in different habituation conditions of a contrast. The factor of Habituation Tone was thus not included in the subsequent main analysis.

In the main analysis looking times in the Same versus Different test trials were examined in a 2x2x3 ANOVA, with Test Trial Type (Same versus Different) as the within-subject factor, and Contrast (T2-T3; T1-T4) and Age (4, 8, and 11 months) as between-subject factors. The effect of Test Trial Type was highly significant, ($F(1,138)=32.528$; $p<.0005$), with longer looking for Different over Same trials. However, we found no Trial Type x Contrast interaction ($F=(1,138)=1.812$; $p=.18$), no Test Trial Type x Contrast x Age interaction ($F=(2,138)=1.259$; $p=.287$), and no Test Trial x Age interaction ($F=(2,138)=.148$; $p=.862$). There was also no other significant main effect or interaction (all $F \leq .848$; $p \geq .43$). These results suggest that infants of all three ages in both contrast groups (T2-T3 and T1-T4) discriminated the two types of test trials. The lack of any interaction in the 2x2x3 ANOVA does not support any decline for the T2-T3 contrast.

However, the means in Figure 1 seem to indicate a discrimination decline for T2-T3 at 11 months of age, but not for T1-T4. Further analyses were therefore conducted on the differential scores of looking times (i.e., Different minus Same for each baby). Age was entered as the number of days, replacing the three broad categories of months. We first performed linear regression analyses separately for the two contrasts. As predicted, the slope was negative for T2-T3 at -0.00541 per day, $t(70)= -1.542$, $p=.064$ (one-tailed), indicating a tendency of discrimination decline with increasing age. The slope was not significant for T1-T4, $t(70)=0.895$, $p=.374$ (two-tailed), indicating no decline across ages. An ANCOVA further assessed the homogeneity of the slopes of the two contrasts, with Age (in days) by Contrast (T2-T3, T1-T4) interaction added to the test. Recall that the purpose of the experiment was to test the prediction that the T2-T3 contrast should decline in discriminability whereas T1-T4 should not. Accordingly, a one-tailed test was appropriate: the homogeneity of the two slopes was rejected, $F(1,140)=2.808$, $p=.048$ (one-tailed), confirming our prediction.

Overall, although the ANOVA results indicated a strong discrimination of both tonal contrasts across the three ages, linear regression and ANCOVA analyses showed some evidence of decline in infants' discrimination of T2-T3 over age.

[Figures 1, 2]

Discussion

This study examined the perception of Mandarin lexical tones by French-learning infants during the first year of life. We tested an acoustically similar contrast (T2-T3) and a more dissimilar contrast (T1-T4). Based on the literature, we expected that the 4-month-old French-learners, our youngest group, should discriminate both tonal contrasts in Mandarin. The results confirmed this prediction, consistent with other studies of tonal discrimination by English- and Dutch-learning 4-6-month-olds (Liu & Kager, 2014; Mattock & Burnham, 2006; Mattock, et al., 2008; Yeung, et al., 2013). Together, the evidence is clear that infants are born as language-general listeners with acute sensitivity to native and nonnative speech contrasts, including consonants (e.g., Werker & Tees, 1984), vowels (e.g., Kuhl, et al., 2006; Polka & Werker, 1994) and lexical tones.

However, despite hearing a language with no lexical tones, the 8- and 11-month-old French-learning infants in our study showed no decline in their discrimination of T1 and T4 in Mandarin. The continuing sensitivity to this contrast was as predicted and may be attributed to its perceptual salience since it is acoustically dissimilar, as shown in Appendix B. This result differs from the previous finding that non-tone-learning infants declined in their discrimination of the

Thai low versus rising contrast (Mattock & Burnham, 2006; Mattock, et al., 2008) and the Cantonese mid versus high-rising contrast (Yeung, et al., 2013) by 9 months of age. It is possible that those contrasts are less acoustically salient than the Mandarin contrasts that we tested. Task differences might further account for the different findings across the studies. Specifically, Mattock and Burnham (2006) used the Conditioned Headturn Procedure, and Mattock, et al. (2008) and Yeung, et al. (2013) used a stimulus-alternating task. Future research can test the Thai and Cantonese contrasts in a habituation procedure to determine if the task can reveal potential discrimination.

Unlike the strong discrimination of T1 and T4, our non-tone-learning infants' response to T2 and T3 was mixed. On the one hand, the ANOVA showed a highly significant discrimination of the two tones in both contrasts, with a complete lack of 3-way interaction, indicating that the two contrasts were discriminated comparably across ages. These results are surprising given that T2-T3 is the most similar tonal contrast in Mandarin, and that our T2-T3 exemplars were specially selected to be even more similar than usual. This suggests that lexical tone sensitivity in non-tone-learning infants is more robust against decline than had been assumed previously.

On the other hand, we also found that reduced acoustical salience of T2-T3 seem to be costly for older infants. The ANCOVA and linear regression results suggest that the discrimination trajectory of T2-T3 across ages differed from that of T1-T4, with the former contrast group showing a marginally significant negative slope (i.e., a decline), and the latter showing a slightly positive but non-significant slope (i.e., no decline). Thus, while younger infants were successful in discriminating both contrasts, older infants showed a tendency of perceptual decline only for the similar T2-T3 contrast. We note that in another study (Shi, et al., in press) that used the same stimuli and same task as in the present study, we found that T2 and

T3 were discriminated consistently from 4 to 13 months of age by Mandarin-learning infants, in contrast to French-learning infants, who showed evidence of perceptual decline, suggesting that learning a non-tonal language attenuates older infants' sensitivity to acoustically similar tonal contrasts. Likewise, Dutch-learning infants close to one year of age declined in their discrimination only when the T1-T4 tokens were artificially contracted in pitch (Liu & Kager, 2014). In our experiment the T2-T3 contrast was more similar than usual because we deliberately selected tokens with comparable durations and equated their overall amplitude. We can predict that this contrast may remain more discriminable to older infants if the duration and amplitude cues are present in the stimuli as they normally are in natural speech.

In the literature discrimination of non-native contrasts is explained in terms of perceptual assimilation patterns related to the native phonological system (Best, 1995). As infants become attuned to the native phonological system, two nonnative phonetic categories that are non-contrastive in the native language can be assimilated into one existing similar native category, while a nonnative contrast can remain discriminable if the contrasting members can fall into (i.e., assimilated into) distinct categories in the native language. Since French, the native language of our infants, does not have any lexical tone and not even stress or pitch accent, assimilation of any lexical tones to distinct categories would not apply. Infants therefore could not have discriminated the tones at the lexical level. Mattock, et al. (2008) inquired whether learning English, a language with contrastive lexical stress, allows infants to be more sensitive to lexical tone contrasts than learning French, which has no stress. They found, however, that English-learning infants were not more sensitive than their French-learning peers to Thai tones; the two groups showed the same pattern of responses.

Is it possible that our French-learning infants processed the Mandarin tonal contrasts by perceptually assimilating them to their native intonations? Assimilation of lexical tones to native intonation categories was examined with English- and French-speaking adults (e.g., Reid, et al., 2015; So & Best, 2014), but the results were variable, depending on whether the target tone to be assimilated appeared in sentential or single syllable contexts. In sentential context French speakers categorized T1 more as statement (31%) than exclamation (28%), and T4 in the reverse pattern (exclamation 46%, statement 25%) (So & Best, 2014). Despite the differential tendencies, the overlap of both intonations for both tones was still quite high. Assimilation of lexical tones may not be as straight forward. As Hallé, Chang and Best (2004) explained, "there exists a continuum of plausible intonations 'in between' various intonation nuances so that a French listener does not have to classify any given intonation within a finite set of contrastive categories." (P418). It is unclear how easy it would be for infants to assimilate lexical tones to distinct intonations.

The continuing tonal discrimination in our infants is more likely psychoacoustic than phonological. Consistent with this idea, French-speaking teenagers, who matched Mandarin-speaking peers in their discrimination of the T1-T2, T1-T3 and T2-T3 contrasts in Mandarin, showed different areas of brain activation upon hearing the contrasts: the French participants responded only psychoacoustically whereas the Mandarin-speakers responded both psychoacoustically and linguistically (Pierce, et al., 2014). Similarly, French-speaking adults in Hallé, Chang, and Best (2004) also perceived Mandarin tonal distinctions, but did so non-linguistically, unlike the Mandarin participants. Overall, sensitivity to nonnative contrasts, especially acoustically salient ones, remains more acute when the processing is psychoacoustic.

In conclusion, our French-learning infants distinguished the T1-T4 contrast in Mandarin throughout the first year of life, showing no decline. This result is consistent with the tonal discrimination shown in non-tone-learning older infants (Hay, et al., 2015; Liu & Kager, 2014; Singh, et al., 2014) and teenagers (Pierce, et al., 2014). Even for the most similar T2-T3 contrast in Mandarin (plus being more similar in our study due to the removal of duration and amplitude cues), our infants only showed weak evidence of decline in discrimination. This might be because lexical tones are encoded by salient prosodic properties. Moreover, the tone-bearing unit is typically larger (for example, the syllable for Mandarin), which is more acoustically salient than consonants. Nevertheless, the tendency to decline in discriminating the acoustically similar T2-T3 contrast by our French-learning infants suggests that acoustic salience interacts with the acquisition of native phonology: Learning a non-tonal language does attenuate infants' sensitivity to acoustically similar lexical tones towards one year of age. Taken together, our results demonstrate that the perceptual system is quite robust for lexical tones, and that non-tone-learning infants' discrimination of lexical tones, being most likely psychoacoustic, is affected by both acoustic salience and perceptual attunement to the native-language phonology.

References

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

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*(3), 395-421.

Hao, Y.C. (2012). Second language acquisition of Mandarin tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*(2), 269-279.

Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: the contrastive use of lexical tone in early word learning. *Child Development*, *56*(1), 10-22.

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*, *67*(4), 243-267.

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*, F13–F21.

Lee, C. Y., Tao, L., & Bond, Z. S. (2010). Identification of acoustically modified Mandarin tones by non-native listeners. *Language and Speech, 53*(2), 217-243.

Li, X., & Chen, Y. (2015). Representation and processing of lexical tone and tonal variants: Evidence from the mismatch negativity. *Plos One*, *10*(12): e0143097.

Liu, L. Q., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385-394.

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

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*, 1367–1381

Narayan, C. R., 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*, 407–420.

Pierce, L. J, Kline, D., Chen, J. K, Delcencerie, A., & Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences of the USA, 111*(48), 17314-17319.

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

Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Rattanasone, N. X., & Best, C. T. (2015). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Attention, Perception, & Psychophysics, 77*(2), 571–591.

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech*, *34*(2), 145-156.

Shi, R., Gao, J., Achim, A., & Li, A. (in press). Perception of lexical tones in native Mandarin-learning preverbal infants and toddlers. *Frontiers in Psychology*.

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

Singh, L., Tan, A., & Wewalaarachi, T. D. (in press). Lexical tone variation and spoken word recognition in preschool children: Effects of perceptual salience. *Journal of Child Language.*

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

So, C. K., Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, *36*, 195-221.

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

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

Wong, P., Schwartz, R., & Jenkins, J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research*, *48*, 1065-1079.

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*(2), 123–139.

Xu, Y. & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, *33*, 319-337.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Acknowledgement

This research was supported by CFI (Canadian Foundation for Innovation), NSERC (Natural Sciences and Engineering Research Council of Canada), and SSHRC (Social Sciences and Humanities Research Council of Canada) grants to the first author. The third author received a scholarship from the China Scholarship Council for this work. We are grateful for André Achim for his valuable assistance on data analysis. We thank all the families who participated in the study.

Figure Captions

Figure 1. Similar contrast: Tone 2 (rising) versus Tone 3 (dipping) in Mandarin. Four-, eight- and 11-month-old French-learning Infants' looking times (means and standard errors) during the test trials (Same: the same tone as in habituation; Different: the contrasting tone).

Figure 2. Dissimilar contrast: Tone 1 (high-level) versus Tone 4 (falling) in Mandarin. Four-, eight- and 11-month-old French-learning Infants' looking times (means and standard errors) during the test trials (Same: the same tone as in habituation; Different: the contrasting tone).

Figure 1.

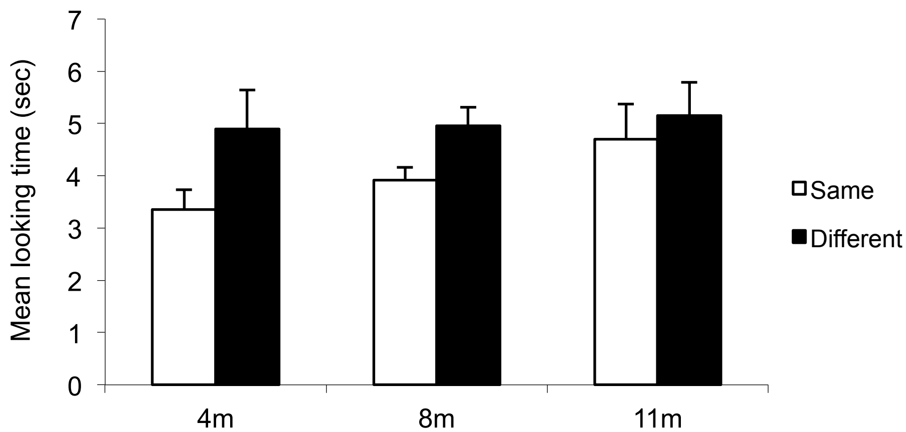
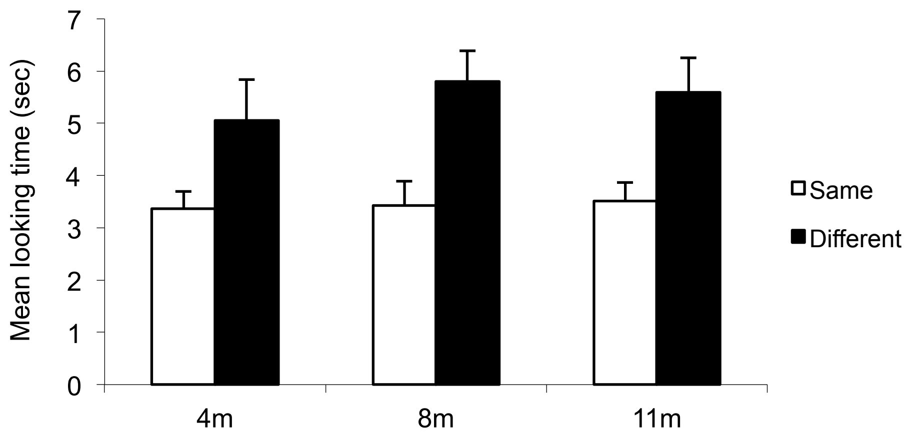
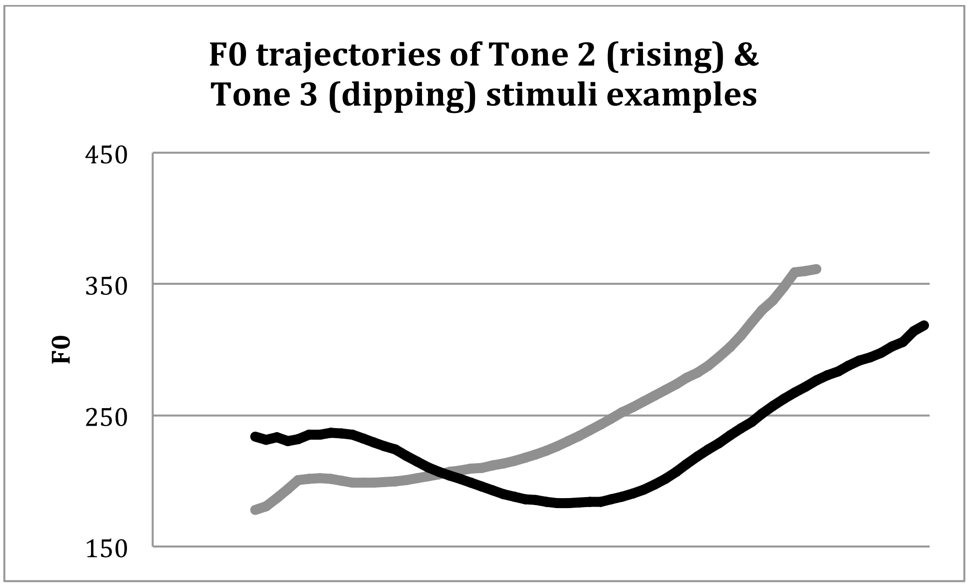


Figure 2.



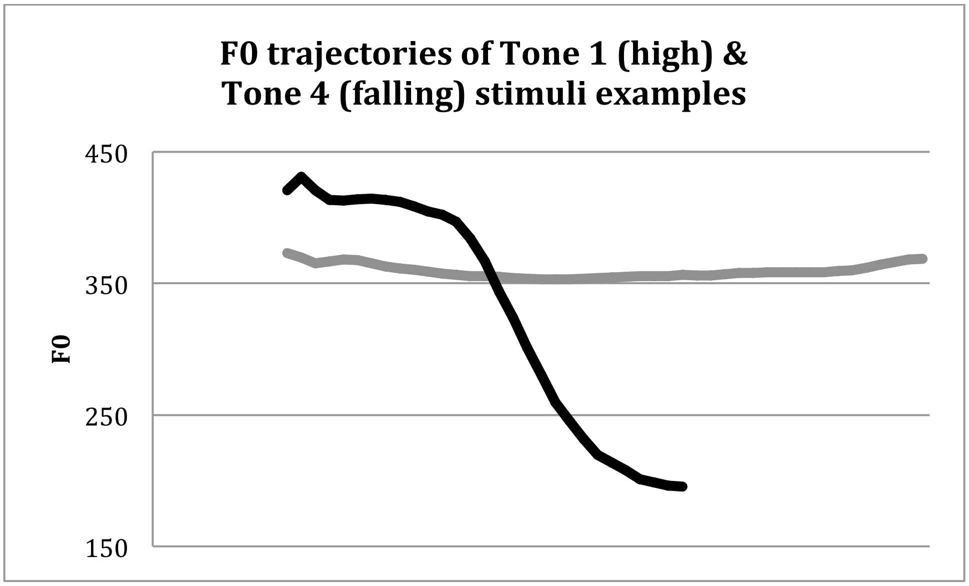
Appendix A

Tone 2 (rising, grey color) and Tone 3 (dipping, black color) Example Stimuli



Appendix B

Tone 1 (high, grey color) and Tone 4 (falling, black color) Example Stimuli



Appendix C

Experimental design (contrast groups and habituation tone sub-groups)

	T1-T4		T2-T3	
Habituation	**T1**	**T4**	**T2**	**T3**
Test trials	Same: T1	Same: T4	Same: T2	Same: T3
	Different: T4	Different: T1	Different: T3	Different: T2