Infant and Child Development
Inf. Child. Dev. (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com). DOI: 10.1002/icd.1944

Discrimination of Lexical Tones in the First Year of Life

Ao Chen* **and René Kager**
Utrecht Institute of Linguistics, Utrecht, The Netherlands

In the current study, we examined the developmental course of the perception of non-native tonal contrast. We tested 4, 6 and 12-month-old Dutch infants on their discrimination of Chinese low-rising tone and low-dipping tone using the visual fixation paradigm. The infants were tested in two conditions that differed in terms of degree of variability. The 4-month-olds did not show discrimination effect in either condition. The 6- and 12-month-old infants, however, discriminated the tones in both conditions. The improvement of perception might be the result of cognitive development carried over from learning the native phonology. Infants can become better listeners in general in the first year of life, as well as get cognitively better equipped in dealing with the variable input in speech in general. Copyright © 2015 John Wiley & Sons, Ltd.

Key words: speech perception; normalization; lexical tones

In the past two decades, the concept of perceptual reorganization (PR) has played a leading role in research on early language development. PR is supported by studies showing that infants are born with sensitivity to both native and non-native phonological contrasts, and, as their exposure to the native language increases, by the end of the first year, their sensitivity to native contrasts improves while their sensitivity to non-native contrasts deteriorates (e.g. Werker & Tees, 1984; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Kuhl et al., 2006). Despite the empirical support for PR, the question of how the establishment of native phonology shapes the perception of non-native speech remains unclear. Native phonology has been assumed to develop primarily through language exposure. Still, the role of general cognitive development as it interacts with the acquisition of phonology remains severely under-investigated. In particular, the role of general cognitive development in early language acquisition cannot be properly assessed by only looking at the perception of native contrasts, as in principle, any improvements observed in the perception of specific contrasts could either be the result of exposure to the native language or be emergent from general cognitive growth. In order to

*Correspondence to: Ao Chen, Utrecht Institute of Linguistics, Trans 10, 3512JK Utrecht, the Netherlands. E-mail: irischen71@hotmail.com

disentangle the factors of input and cognition, the perceptual development of non-native contrasts needs to be studied. If improvement observed in native language reflects and relates to general cognitive ability, then such improvement should be observed for the processing of non-native languages, or even non-language stimuli, as well.

In the current study, we examine the role general cognitive development in language acquisition, by testing infants on their perception of non-native phonological contrasts. We focus on two aspects that are essential to early phonological acquisition: (i) the ability to perceive the phonetic differences between tokens belonging to different categories and (ii) the ability to form a category by normalizing over variable within-category tokens. With regard to the first aspect, we ask how the previously found decline in non-native sound perception (due to a lack of exposure) trades off against the developing skill of perceiving acoustical details (due to general cognitive growth). That is, can infants become more (rather than less) accurate at discriminating acoustic details, irrespective of whether the contrast belongs to the native or non-native language? For the second aspect, we ask whether the development of normalization ability is restricted to the native language or may be observed in non-native language as a result of general cognitive development. We tested Dutch 4-, 6-and 12-month-old infants on their perception of non-native lexical tone contrasts. Pitch variations abound in all languages, yet pitch does not play a lexically contrastive role in Dutch as it does in Chinese. Hence, the perception of lexical tones by non-tone language-learning infants can be used to pinpoint whether developmental change emerges from growth in general cognitive or linguistic ability or from exposure to the contrastive function of tones *per se*.

Based on previous studies, three developmental trends have been found in early speech processing. First, in the second half of the first year, the discrimination of non-native phonological contrasts deteriorates for both consonants and vowels (e.g. Werker & Tees, 1984; Kuhl et al., 1992); second, the perception of native contrasts improves (e.g. Kuhl et al., 2006) until childhood; third, infants get better at normalizing variable speech tokens and even benefit from variations in input when learning words (e.g. Rost & McMurray, 2009). With regard to lexical tones, Mattock and Burnham (2006) and Mattock et al. (2008) tested non-tone language-learning infants and found that native English infants and native French infants were able to discriminate Thai rising and low tones at 4 and 6 months, but they lost this sensitivity at 9 months. Native Chinese infants, on the other hand, were able to discriminate the contrast at both 6 and 9 months. These two studies serve as the basis for the claim that the PR of lexical tones is in line with the PR of consonants and vowels. Recently, Yeung, Chen, and Werker (2013) tested native English infants on their discrimination of Cantonese rising tone and low-level tone and found a similar decline from 4 to 9 months. In addition, the authors also tested Chinese infants on the same tonal contrast. Although Chinese is a tone language, there is no contrast between rising and low-level tone. They found that at 4 months of age, the Chinese infants and Cantonese infants already showed different discrimination patterns. Cantonese infants detected a change from rising to rising low-level alternating tones as well as from dipping to the alternating tones, whereas Chinese infants only detected a change from low-level to alternating tones. These studies seem to show that language-specific perception of tones can occur as early as 4 months. They also argue for a similar decline in the discrimination of lexical tones between 6 and 9 months among infants learning a non-tone language. Contradictory evidence, however, has also been shown in previous studies. Liu and Kager (2014) found that Dutch infants succeeded at discriminating

DOI: 10.1002/icd

the Chinese high-level (T1) and high-falling (T4) contrast at 4–5, 8–9, 11–12 and 14–15 months, with no decline in sensitivity. Liu and Kager interpreted this maintained discrimination ability as the result of the high salience of the Chinese T1–T4 contrast.

Infants' perception of native contrasts improves in the first year of life. Their auditory system gets sharpened in the first months (Tharpe & Ashmead, 2001; Marcoux, 2011), and exposure to the native language helps to tune their perception to native phonology. Narayan, Werker, and Beddor (2010) found that Filipino-learning infants only succeeded at discriminating the acoustically not-so-salient native initial /n/-/ŋ/ contrast at 10–12 months but not at 6–8 months. The English infants, on the contrary, remained unable to discriminate this contrast. Kuhl et al. (2006) found a significant improvement in the discrimination of the English /r/-/l/ contrast among English-learning infants between 6 and 12 months, whereas infants learning Japanese, where the /r/-/l/ contrast is not phonemic, showed a decline in discrimination. A similar improvement has been observed for the English /d/-/ð/ contrast. When English 10- to 12-month-old infants, children and adults were tested, a significant improvement in discrimination was observed from infancy to childhood and adults surpassed both infants and children (Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006). These studies lend evidence to the hypothesis that experience with certain contrasts is beneficial for discrimination. Also, they suggest that learning might be necessary for discriminating the not-so-salient contrasts. In other words, if the infants are not exposed to languages where these contrasts are present, they may remain insensitive to such contrasts regardless of age. Yet, so far, the questions of whether lexical tone perception also improves with age and whether input is the foundation for (later) success remain unanswered. Interestingly, non-tone language adults can be quite accurate at discriminating lexical tones. Chen, Liu, and Kager (2015) tested Dutch adult listeners on their discrimination of bisyllabic Chinese lexical tones and found that overall the Dutch listeners reached an accuracy higher than 75%. Similarly, So and Best (2010) found that after a 2-min familiarization phase, English listeners were able to identify and discriminate Chinese lexical tones fairly accurately. In another study, native English speakers were found to be highly accurate at discriminating Thai lexical tones (Burnham et al., 1996). These findings lead to the question whether lexical tones have to be acquired, as suggested by Narayan et al. (2010) and Kuhl et al. (2006), or the discrimination simply improves as infants develop into better perceivers.

Besides the perception of acoustics, infants need to cope with variation in speech. The PR studies have not focused on this question, yet it has received much attention in word learning studies. It has been found that young infants (7.5-month-old) tended to overlisten to phonetic details, and they failed to generalize across the same words produced with different affects and voices (Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004). Older infants (14-month-old), however, benefited from training with extensively variable tokens of the same word when learning new words (Rost & McMurray, 2009). Singh et al. (2004) also showed that for 7.5-month-olds, familiarization with highly variable tokens of the target words was beneficial for recognizing these words later in fluent speech, whereas low variability impeded recognition. Hence, variable tokens can work as a beneficial factor for teasing apart phonetic details from the core phonological representation of a word. In light of this facilitation effect in word learning, presumably, exposure to variable input may also enhance the specification of phonological categories, especially for older infants. Looking at adults, although the Dutch listeners reached high accuracy when discriminating Chinese lexical tones,

DOI: 10.1002/icd

they perceived the tones acoustically rather than linguistically (Liu, Chen, & Kager, under review). For PR studies, a distinction should be made when infants are presented with one single token of a category and when they are presented with variable tokens of one category. In the former case, the infants are discriminating pure acoustics, whereas in the latter case, they are encouraged to ignore the inter-token variation and treat the variable tokens as one phonological category. Yet, the influence of variability has not received much attention in studies focusing on PR. The degree of variation in stimuli differed from one study to another. For example, in Werker and Tees (1984), four tokens of English /ta/-/da/ contrast were used, but only three tokens of the Thompson / ḱi/-/ q́i/ contrast were used. Kuhl et al. (1992) used 32 tokens of each vowel in the discrimination task. In the lexical tone perception studies, Mattock and Burnham (2006) used five different tokens as stimuli, whereas Liu and Kager (2014) and Yeung et al. (2013) used four lexical tone tokens. Hence, it remains unclear whether the specific variation in the stimuli caused the difference in performance. To our knowledge, Shi (2010) is the only study that has focused on the generalization of lexical tones, yet only by native infants. She found that 8- to 11-month-old, but not 4- to 5-month-old, Chinese infants were able to generalize between variable tokens of Chinese rising tone (T2) and Chinese falling tone (T4). The findings of Shi (2010) suggest that older native infants get better at categorizing lexical tones. Based on the latter studies, one possible cause of the perceptual decline found in previous PR studies is that non-tone language-learning infants become less capable of differentiating the tonal categories *phonologically* when presented with variable tokens, rather than lose the ability to perceive the difference between the lexical tones *phonetically*.

So far, the normalization of variable tokens has only been studied in the context of native language. Yet what underlies the improvement of normalization has not received much attention. When considering only speech perception, categorical perception of non-native speech tokens seems to be not motivated. Categorical perception, which is the basis for phonological knowledge, has only been observed for native language (Francis, Ciocca, & Ng, 2003; Hallé, Chang, & Best, 2004). The ability to normalize variable tokens, however, may not be solely relevant to the native language. Three hypotheses can be made regarding infants' normalization ability. First, it may emerge solely as the result of learning the native language. If this is the case, then infants should only be able to normalize within native, not non-native, phonological categories. Second, it may be driven by the experience with native language, yet such ability can be transferred to the perception of a non-native language. Third, such ability may reflect emergent cognitive skills that may not be specific to language. To our knowledge, only Frota, Butler, and Vigário (2014) incorporated the issue of normalization into the study of early perception of native versus non-native prosody. She tested Portuguese infants on their discrimination of Portuguese intonation as well as Chinese lexical tones. Portuguese 5- and 8-month-old infants were tested on their discrimination of question versus statement intonation in Portuguese, as well as the discrimination of bisyllabic Chinese tones, all realized with multiple tokens. The tonal contrast always occurred on the second syllable. Both groups of infants succeeded at discriminating the Portuguese intonation, whereas no lexical tone discrimination was found in either group. These findings suggest that variability might be a factor that hinders infants from discriminating non-native contrasts, yet they did not include a single token version to tackle this hypothesis.

To sum up, previous studies suggest that on one hand, infants become less sensitive to non-native contrasts, as these are irrelevant for learning the native phonology; on the other hand, their auditory perception is better tuned and they

DOI: 10.1002/icd

improve at dealing with more complex and variable speech signals. These findings lead us to ask how the development of speech perception may reflect and interact with general cognitive development. Specifically, we are interested in the question whether the improvement observed in the perception of the native language in previous studies may reflect development that serves a broader purpose than native language *per se*. Therefore, in the current study, we tested Dutch infants on non-native contrasts, thus ruling out native language exposure as a source of perceptual development.

Acoustical salience has been found to influence speech discrimination (Aslin & Pisoni, 1980). Some non-native contrasts remain robustly discriminable from infancy into adulthood (Best, McRoberts, & Sithole, 1988). On the contrary, some phonetically more similar contrasts, even native ones, can only be discriminated at a later age (Narayan et al., 2010). With regard to Chinese, among the four lexical tones, the T2–T3 contrast has been found to be perceptually the most similar contrast for both native and non-native adults, as well as native children (Hume & Johnson, 2003; Hua & Dodd, 2000). Tsao (2008) found that Chinese infants were able to discriminate T1–T3, T2–T4 and T2–T3 contrasts, but they were more accurate at discriminating the acoustically less similar T1–T3 contrast than the other two contrasts. Gao, Shi, and Li (2011) also found that, although the Chinese infants failed to discriminate the T1–T4 contrast, they did succeed at discriminating T1–T3 contrast, which she claimed to be acoustically more distinct. The maintained success of Dutch infants at discriminating Chinese T1–T4 contrast is also attributed to the high acoustic salience of these tones (Liu & Kager, 2014).

In the current study, we chose the not-so-salient T2–T3 contrast as stimuli, as the discrimination of the salient T1–T4 contrast remains successful regardless of age (Liu & Kager, 2014), which might mask the possible developmental change in the perception of non-native lexical tones. We manipulated the variability in the stimuli in two conditions to assess the development in perceiving phonetic differences and phonological contrasts. Seeing the success of Dutch adults at perceiving Chinese tones, we hypothesize that Dutch infants will ultimately succeed at discriminating the lexical tones acoustically beyond the assumed PR window. If token normalization depends on and is restricted to the native language, then discrimination of non-native speech tokens should not benefit from developing normalization skills, and the older infants are expected to be unable to discriminate the lexical tones when presented with variable tokens. On the other hand, if token normalization depends on general cognitive ability or can be transferred to non-native language from native language, then we expect to see an improvement in the discrimination of variable non-native speech tokens during the first year of life.

EXPERIMENTS

Participants

Twenty-four 4-month-old infants (age range 4:02–4:29), 25 6-month-old infants (age range 6:01–7:02) and 24 12-month-old infants (11:01–12:10) participated in the study. All infants were born in the Netherlands and had been raised in Dutch by the time of the experiment. None of them had been exposed to a tone language by the time of the experiments. All were healthy infants, and no parents reported hearing problems or other disorders. Another 14 4-month-old infants were tested but excluded from the analysis due to crying ($N = 3$), not finishing the experiment ($N = 3$) and failure to meet the habituation criterion ($N = 8$). Another nine 6-month-

DOI: 10.1002/icd

old infants were tested but excluded from analysis due to not finishing the experiment ($N = 2$) and failure to meet the habituation criterion ($N = 7$). Another 10 12-month-old infants were tested but excluded from analysis due to equipment failure ($N = 3$), fussiness ($N = 2$), experimenter's error ($N = 2$), parental interferences ($N = 2$) and failure to finish the experiment ($N = 1$).

For each age group, the same infants participated in the two conditions on two different days, with at least 2 days in between. Within each age group, the order of condition A and condition B was counterbalanced.

Stimuli

The Chinese rising tone (T2) and dipping tone (T3) were used as stimuli. The syllable /ma/ was chosen as the tone-bearing syllable. The infants' task was to discriminate between /ma/ carrying T2 and /ma/ carrying T3. The experiment involved two conditions, with stimulus sets differing per condition in terms of inter-token variability. In condition A, one single token of T2 and one single token of T3 produced by a single female native speaker were used. In condition B, 13 different tokens of T2 and 13 different tokens of T3 produced by a single female native speaker were used.

For condition A, a female Chinese speaker recorded /ma2/ and /ma3/ in isolation, together with other syllables. Then the pitch contours of naturally produced /ma2/ and /ma3/ were extracted by the software PRAAT (Boersma & Weenink, 2001). After normalizing the duration of these two contours (450 ms), the pitch contours of the T2 after time normalization were re-synthesized onto the original T3 syllable using the PSOLA method (Moulines & Laroche, 1995). The pitch contour of T2 and T3 in condition A is depicted in Figure 1. Time normalization ruled out the possibility of interference from duration as a potential confounding factor in the experiment. Five native Chinese speakers listened to the stimuli and were all in agreement that all the stimuli sounded like natural, normal speech.

For condition B, multiple tokens of either T2 or T3 were used as habituation stimuli. The same native Chinese female speaker as in condition A recorded multiple tokens of /ma/ carrying either T2 or T3, and 12 tokens that shared a similar intensity and duration were chosen as habituation stimuli. The mean duration of

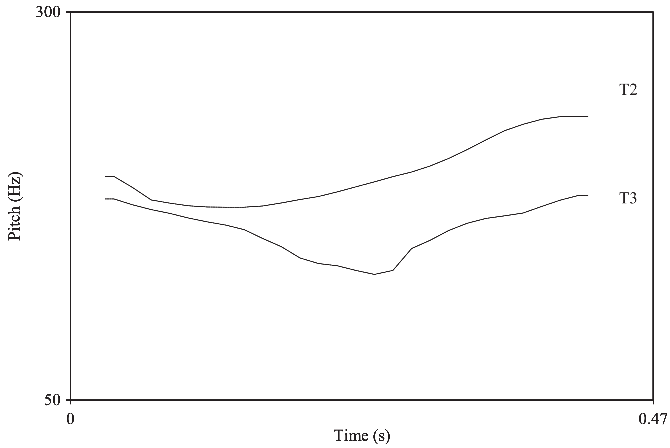

Figure 1. The pitch contour of T2 and T3 used in condition A.

DOI: 10.1002/icd

the syllable /ma/ carrying T2 was 431 ms (SD = 20.5 ms, range 396–459 ms), and the mean duration of /ma/ carrying T3 was 431 ms (SD = 20.7 ms, range 400–453 ms). For each tone, four chains were generated, each containing three different tokens out of the 12. The pitch contours of the T2s and T3s used in condition B are depicted in Figure 2. As creaky voice in the middle of the T3 contour is common in Chinese (Belotel-Grenié & Grenié, 2004), we did not exclude those tokens that had creaky voice in condition B. Within each chain, the inter-stimulus interval was set at 1 s, identical to the inter-stimulus interval in condition A. Within each trial of the habituation phase, the infants were presented with one of the chains, and all four chains occurred in the habituation phase. The order of occurrence of the chains in the habituation phase was randomized for each infant. By manipulating the stimuli in this manner, we created a semi-random presentation of the multiple tokens of each tone in the habituation phase. In the test phase, we selected another token of each tone produced by the same female speaker. These two tokens were manipulated in PRAAT to have an equal duration of 450 ms. The visual stimuli were the same as in condition A.

Procedure

The visual fixation paradigm was selected for the current study. The lab consisted of a test cabin and a separate control room for the experimenter. During the experiment, infants sat on their parent's lap in the test cabin; a 14 in computer screen displaying the visual stimuli was placed in front of the baby at a distance of about 1 m. The auditory stimuli were presented at a comfortable volume through a hidden speaker in front of the baby. The parent listened to background music through headphones to prevent possible interaction with the infants. A hidden camera mounted above the screen recorded the infants' looking behaviour; the live video stream was transferred to the experimenter's computer in the control room. The experimenter observed the video and recorded the visual fixation of the infants by pressing the 'looking' and 'non-looking' button on a button box connected to the control computer. The experiment included a habituation phase followed by

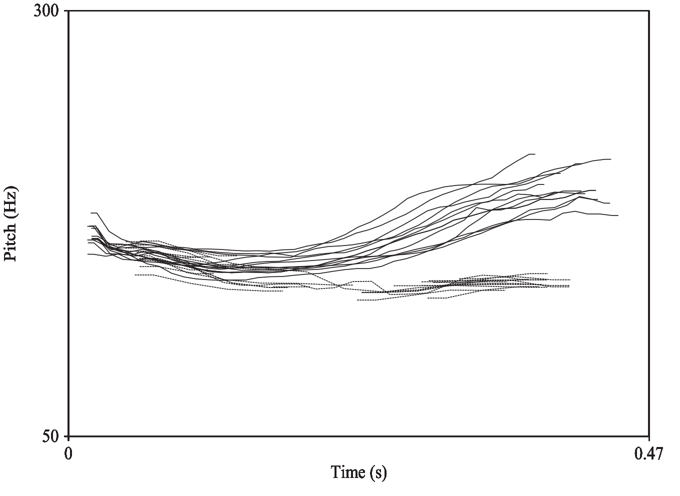

Figure 2. The pitch contours of T2s and T3s used in condition B.

DOI: 10.1002/icd

a test phase. When the pretest finished and once the infant focused on the screen, the experimenter initiated the first habituation trial by pressing the 'looking' button, and once the infant looked away, the experimenter pressed the 'non-looking' button. The looking and non-looking of the infants were always recorded by these two buttons. If the infant looked away and then looked back at the screen within 2 s, the same trial continued, while if the infant looked away for more than 2 s, the current trial ended and the attention getter appeared on the screen to get the attention of the participant back. After the attention getter was activated and the infant looked back at the screen, the experimenter started the next trial by pressing the 'looking' again. All the trials were started and finished as described earlier. The looking time of each look and the total looking time of one single trial were recorded automatically by the experimenter's control computer. In this way, the infants' looking time to the visual stimuli was used as an indicator of their attention to the auditory stimuli.

The total looking time of the first three trials in the habituation phase was used as a baseline for measuring habituation. Starting from the fourth trial, the total looking time of each three consecutive habituation trials was calculated, and once this looking time is less than 65% of the total looking time of the first three habituation trials, the habituation criterion was met, and the test phase started automatically. In this way, the habituation phase had minimally six trials. In the test phase, the infants were presented with one 'old' trial, which was another token of /ma/ with the same tone as they had heard in the habituation phase, and another 'novel' trial, which was /ma/ carrying the other tone that they had not previously heard in the habituation phase. One single token of each tone was used in the test phase. The tones that were used in the habituation phase were counterbalanced. In the test phase, if the infants were able to detect the difference between the two tones, then upon hearing the novel trial, their listening time should be recovered due to hearing something new, hence a longer looking time is expected to the novel trial than to the old trial. Specifically regarding this procedure, there is no reason to suspect that the looking time of the novel trial would be significantly shorter than the looking time to the old trial. In other words, only increase in looking time can serve as an indicator of successful discrimination, while equal or decrease of listening time may just be the result of losing interest in general.

In the visual fixation procedure, the habituation phase works as a training phase for the infants to learn the tones, and in the test phase, they needed to discriminate between the 'learned' tone and a novel tone. Hence, it gives us the opportunity to manipulate how much variability is provided in the training phase. In the current study, in condition A, the infants were habituated with one single token of either T2 and T3, and the same token of the habituated tone together with the unheard tone were used in the test phase. In condition B, the infants were habituated with 12 different tokens of one tone, and in the test phase, they were presented with another token of the habituated tone together with one token of the other tone.

Results and Discussion

As we consider the habituation phase as a learning phase, only those who 'learned' the tones, that is, those who met the habituation criteria were taken into analysis. After excluding the infants who failed to reach the habituation criteria, there were 25 4-month-old, 24 6-month-old and 24 12-month-old infants.

DOI: 10.1002/icd

The raw looking times of the 'old trial' and the 'novel trial' in the test phase were first log transformed to correct skewness, and the log-transformed looking times fit normal distribution. We submitted the log-transformed looking time to a 2 condition (condition A or condition B) × 2 trial type (old or novel) repeated measures ANOVA, and age and order (condition A first or condition B first) were put into the model as between-subject factors. A significant effect of trial type was found, $F_{\text{trial type}}$ (1, 67) = 14.68, $p < 0.05$, partial $\eta^2 = 0.18$, and age showed marginally significant effect, F_{age} (2, 67) = 2.78, $p = 0.07$, partial $\eta^2 = 0.08$. There was significant interaction between age and trial type, $F_{\text{age*trial type}}$ (2, 67) = 4.89, $p < 0.05$, partial $\eta^2 = 0.13$. No significant effect of condition was found $F_{\text{condition}}$ (1, 67) = 0.48, ns, and condition did not interact with any other factor significantly. Neither did order show any significant effect, F_{order} (1, 67) = 0.29, ns. Figure 3 plots the log-transformed looking time of each trial type of each age group. As can be seen in Figure 3, the 6- and 12-month old infants had significantly longer looking time to the novel trial whereas the 4-month-old infants did not, and the main effect of trial type came from the two older groups. Overall, the 6-month-old infants had slightly shorter looking time than the 4- and 12-month-olds, which accounts for the marginal significant effect of age.

Next, we look into each age group, again using the 2 condition × 2 trial type condition repeated measures ANOVA. As order failed to introduce any significant effect, both orders were pooled together for the analysis. For the 4-month-old infants, a marginal effect of condition was found, $F_{\text{condition}}$ (1, 24) = 3.73, $p = 0.065$, whereas trial type did not show any significant effect, $F_{\text{trial type}}$ (1, 24) = 0.13, ns. No significant interaction between the factors was observed. For 6-month-old infants, trial type had a significant main effect, $F_{\text{trial type}}$ (1, 23) = 7.82, $p < 0.05$, but condition did not, $F_{\text{condition}}$ (1, 23) = 0.01, ns, and no significant interaction between these two factors was found. Twelve-month-old infants showed a similar pattern to the 6-month-old group, namely that trial type had a significant main effect, $F_{\text{trial type}}$ (1, 23) = 9.92, $p < 0.01$, but condition did not, $F_{\text{condition}}$ (1, 23) = 0.55, ns, and no significant interaction between the two factors was found.

As can be seen from Figure 3, both the 6- and the 12-month-old groups had increased looking time to the novel trial in both conditions, which indicates successful discrimination between the two tones. In comparison, the 4-month-old infants

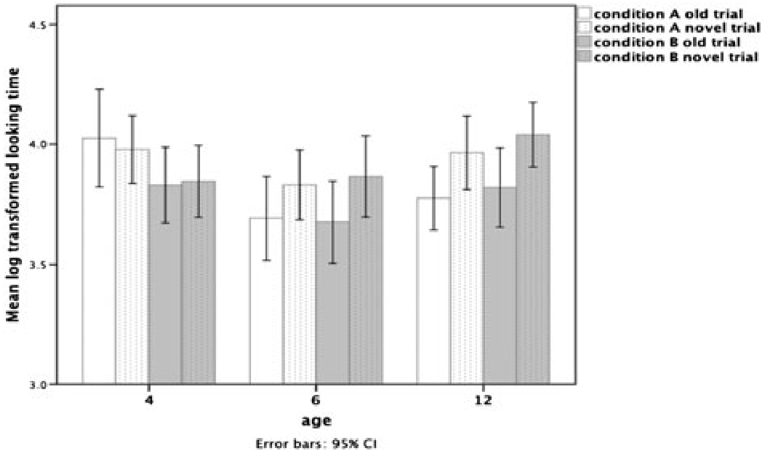

Figure 3. The log-transformed looking time of the old and novel trials of each age group in both conditions.

DOI: 10.1002/icd

had almost equal looking time to the old and the novel trial in both conditions and failed to show the discrimination effect. These results demonstrate that 4-month-old Dutch infants are not able to discriminate the acoustically not-so-salient Chinese T2–T3 contrast. However, as they grow older, they become better at perceiving this contrast and show successful discrimination at both 6 and 12 months of age. In addition, for 6- and 12-month-old Dutch infants, no effect of variability is found. They succeeded at discriminating Chinese T2 and T3 both when they were habituated with one single token of the tone or with multiple tokens of the tone. The success of the 6- and 12-month-old Dutch infants demonstrates that they are able to perceive the acoustical differences between Chinese T2 and T3. In addition, when trained with multiple tokens of the same tone, they were able to ignore the inter-token variations and were able to treat variable tokens as one single type and later discriminate a different tone from it. In summary, Dutch infants get better at discriminating Chinese T2 and T3 in the first year of life, and 6- and 12-month-olds outperform 4-month-old infants.

GENERAL DISCUSSION

In this study, we tested Dutch infants' discrimination of non-native Chinese rising (T2) and dipping (T3) tone at 4, 6, and 12 months of age using a visual fixation paradigm. We manipulated the amount of variation present in the habituation phase. In two different conditions, the infants were habituated with a single token of each tone and multiple tokens of each tone produced by one speaker. We found that the 6- and 12-month-old infants outperformed the 4-month-old infants: no discrimination effect was found among the youngest group in either condition, whereas the 6- and 12-month-old infants succeeded in discriminating the tones in both conditions. It can be seen that the overall discrimination of T2 and T3 improved between 4 and 12 months. Moreover, both the 6- and 12-month-old infants, but not the 4-month-old infants, succeeded at discriminating the two tones when they were habituated with multiple tokens of one tone in condition B. This improvement indicates that not only can infants get better at perceiving the non-native tones acoustically but they also become more capable of normalizing variable non-native tokens when they grow older.

The current study demonstrates that at least for some non-native contrasts, infants become more accurate at discriminating them in the course of the first year of life. We chose the not-so-salient T2–T3 contrast (Hume & Johnson, 2003) in order to observe how the perception of non-native acoustics and the perception of non-native phonology develop and interact through time. Narayan et al. (2010) has shown that for the native not-so-salient contrasts, it takes time for infants to learn them, and older infants outperform younger ones. Here, we find that not only can the perception of not-so-salient native contrast improve when infants grow older but the perception of not-so-salient *non-native* contrast can also improve. Such improvement cannot be attributed to being exposed to this tonal contrast *per se*. What it suggests instead is that infants perceive acoustical information better in general when they grow older and that the enhanced acoustical perception needs not to be specific to native speech. On the one hand, such enhancement may be the bonus of learning native language, namely that acquaintance with native speech helps infants to pick up acoustical details in general. On the other hand, the current study suggests that the causal relationship can be reverse: infants become more accurate at perceiving acoustical details, which facilitates their perception of native contrasts. Most likely, the general auditory development and the learning of

DOI: 10.1002/icd

language-specific acoustics complement each other. How general cognitive development and native language acquisition interact needs further investigation.

Variations abound in natural speech, and learning phonology requires the organization of acoustics into structures. We have explored how infants deal with variable non-native tokens when they learn their native language. Young infants tend to overlisten to phonetic details at an early age, which may hinder them from splitting the core phonological representations from phonetic variations (Houston & Jusczyk, 2000; Singh et al., 2004; Singh, L. 2008). Older infants, on the other hand, take advantage of variation when learning new words (Rost & McMurray, 2009; Hollich, Jusczyk, & Brent, 2002; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). A similar developmental pattern in normalizing variant non-native tokens has been found in our study. In our case, the infants were to discriminate between two non-native contrasts without any meaning involved. The habituation phase in the visual fixation paradigm encourages the infants to 'learn' a sound. Depending on the variations present, the infants were either encouraged to learn the phonetic difference (condition A) or to represent the tones phonologically (condition B). Our results suggest that, after a short exposure to variable tokens, 6- and 12-month-old infants were able to build up a representation of a non-native category to some extent and to discriminate a token from a different category from it, which suggests successful phonological processing of lexical tones by non-tone language-learning infants to a certain extent. Lexical tones are notoriously difficult for non-native speakers to learn (Kiriloff, 1969; Bluhme & Burr, 1971). Nevertheless, adult non-tone language speakers' ability to associate pitch pattern to word meaning, which is similar to learning lexical tones, improves after a short training (Wong & Perrachione, 2007; Song, Skoe, Wong, & Kraus, 2008). It seems that the ability to process non-native lexical tones phonologically is not completely lost, yet what causes such skills to be maintained needs to be future investigated.

The 6- and 12-month-olds succeeded at discriminating the lexical tones in condition B where they needed to normalize variable tokens. These results clearly reject the first hypothesis, namely that normalization is restricted to the native language. Our findings lead to new questions such as how the perception of non-native contrasts changes in early life and what exactly differs between the perception of native and that of non-native speech. The ability to normalize non-native tokens cannot be attributed to learning these contrasts. Our finding that the older infants outperform the younger ones in condition B argues for a development serving broader purposes than native language alone. For now, it is hard to ascertain whether normalization ability is language specific or domain general. Categorization is common in multiple domains, such as the perception of facial expressions in the visual domain (Ludemann, 1991; Grossmann, Gliga, Johnson, & Mareschal, 2009; Batty & Taylor, 2002) and pitch perception in music (Trehub & Hannon, 2006; Stalinsky & Schellenber, 2010). Hence, it is probable that the observed development reflects general cognitive enhancement. Yet again, how the acquisition of native phonology, where categorization is essential, contributes to such general cognitive development or vice versa needs more investigation. One possible way to calibrate the general development would be to investigate whether the improvement in categorization abilities of different stimuli (possibly from different perceptual modalities) tends to occur at a similar age.

It has been assumed that by the end of the first year, infants have already tuned into the phonology of their native language (e.g. Werker & Tees, 1984; Kuhl et al., 1992; Mattock & Burnham, 2006). Yet, the observations in PR studies generally focused on the age window between 6 and 9 months, whereas the performance beyond 9 months has not been well studied (e.g. Mattock & Burnham, 2006;

DOI: 10.1002/icd

Mattock et al., 2008). Liu and Kager (2014) found a 'U curve' in the discrimination of non-native tonal contrasts, such that the discrimination is less accurate at 9 months compared with both 6 and 12 months. We also found successful discrimination of non-native contrast at 12 months. Recent studies show evidence that infants maintain sensitivity to certain non-native contrasts at an older age after the presumed PR window (Mazuka, Hasegawa, & Tsuji, 2014; Tyler, Best, Goldstein, & Antoniou, 2014), and meta-analysis on infant vowel perception found increased effect size in native vowel discrimination after 6 months, but no statistical evidence for developmental change regarding non-native vowels (Tsuji & Cristia, 2014). What our results, together with other studies, suggest is that what we observe in infants' native language development may not simply reflect the acquisition of specific structures in the native language, either in terms of acoustical perception or in terms of categorization. At least by the end of the first year, these abilities can be observed in the processing of non-native speech. Undoubtedly, adults are less capable of learning a non-native language compared with infants. Yet language and cognition are both complex and flexible, and it is not surprising that 1-year-olds have not completely lost their ability to discriminate non-native sounds. In light of the variable developmental patterns observed so far, it is now time to ask what factors determine the developmental trajectory of the perception of different speech sounds, both within and outside the native language. Also, an old paradox is reinforced: although we are equipped with better cognitive abilities when we grow up, we nevertheless lose the ease to acquire a language. A more articulated view of interactions of cognitive and linguistic development will allow a deeper exploration of what restrains us from applying our cognitive abilities to second language acquisition, which will be highly informative for understanding how the human mind functions, as well as learning.

REFERENCES

Aslin, R. N. & Pisoni, D. B. (1980). Effects of early linguistic experience on speech discrimination by infants: A critique of Eilers, Gavin, & Wilson. *Child Development*, *51*, 107–112.

Batty, M. & Taylor, M. J. (2002). Visual categorization during childhood: An ERP study. *Psychophysiology*, *39*(4), 482–490. DOI: 10.1111/1469–8986.3940482.

Belotel-Grenié, A. & Grenié, M. (2004). The creaky voice phonation and the organisation of Chinese discourse. *Proceedings of tonal aspects of languages*. Beijing.

Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology*, *14*(3), 345–360.

Bluhme, H. & Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics*, *22*, 51–57.

Boersma, P., & Weenink, D. (2001) *Praat*: *Doing phonetics by computer* (*version 5.1.05*).

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F. & Keller, P. (1996). Perception of lexical tones across languages: Evidence for a linguistic mode of processing. *ICSPL*, *1996*, 2514–2517.

Chen, A., Liu, L., & Kager, R. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences*, *48*(0), 62–69. DOI: 10.1016/j.langsci.2014.12.002

Francis, A. L., Ciocca, V. & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Attention, Perception, & Psychophysics*, *65*(7), 1029–1044.

Frota, S., Butler, J. & Vigário, M. (2014). Infants' perception of intonation: Is it a statement or a question? *Infancy*, *19*(2), 194–213. DOI: 10.1111/infa.12037.

Gao, J., Shi, R., & Li, A. (2011). Lexical tone perception in non-tone-learning infants. 12th International Congress for the Study of Child Language, Montreal.

DOI: 10.1002/icd

Grossmann, T., Gliga, T., Johnson, M. & Mareschal, D. (2009). The neural basis of perceptual category learning in human infants. *Journal of Cognitive Neuroscience*, *12*, 2276–2286. DOI: 10.1162/jocn.2009.21188.
Hallé, P. A., Chang, Y., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*, (3), 395–421. DOI: 10.1016/S0095–4470(03)00016–0
Hollich, G., Jusczyk, P. & Brent, M. (2002). Talker variation and word learning. *Acoustical Society of America Journal*, *111*, 2455–2455.
Houston, D. M. & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570–1582.
Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, *27*(01), 3–42. Retrieved from DOI: 10.1017/S030500099900402X
Hume, E., & Johnson, K. (2003). The impact of partial phonological contrast on speech perception.
Kiriloff, C. (1969). On the auditory discrimination of tones in Mandarin. *Phonetica*, *20*, 63–67.
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), 13–21.
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*, *31*, 606–608.
Liu, L., Chen, A. & Kager, R. (under review). Perception of tones in Mandarin and Dutch adult listeners, submitted to Language and Linguistics.
Liu, L. & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385–394. DOI: doi.org/10.1016/j.cognition.2014.06.004
Ludemann, P. M. (1991). Generalized discrimination of positive facial expressions by seven- and ten-month-old infants. *Child Development*, *62*(1), 55–67. DOI: 10.1111/j.1467–8624.1991.tb01514.x.
Marcoux, A. M. (2011). Maturation of auditory function related to hearing threshold estimations using the auditory brainstem response during infancy. *International Journal of Pediatric Otorhinolaryngology*, *75* (2), 163–170. DOI: 10.1016/j.ijporl.2010.10.027
Mattock, K. & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241–265. DOI: 10.1207/s15327078in1003_3.
Mattock, K., Molnar, M., Polka, L. & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*(3), 1367–1381. DOI: 10.1016/j.cognition.2007.07.002.
Mazuka, R., Hasegawa, M. & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology*, *56*(2), 192–209. DOI:10.1002/dev.21193.
Moulines, E., & Laroche, J. (1995). Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Communication*, *16*(2), 175–205. DOI: 10.1016/0167–6393(94)00054–E
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*(3), 407–420. DOI: 10.1111/j.1467–7687.2009.00898.x.
Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, *109*, 2190—2201.
Rost, G. C. & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349. DOI: 10.1111/j.1467–7687.2008.00786.x.
Shi, R. (2010). Contextual variability and infants' perception of tonal categories. *Chinese Journal of Phonetics*, *2*, 1–9.
Singh, L., Morgan, J. L. & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*(2), 173–189. DOI: 10.1016/j.jml.2004.04.004.

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*(2), 833–870. DOI: 10.1016/j.cognition.2007.05.002.

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. DOI: 10.1177/0023830909357156.

Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, *20*(10), 1892—1902. DOI: 10.1162/jocn.2008.20131

Stalinsky, S. & Schellenber, G. (2010). Shifting perceptions: Developmental changes in judgments of melodic similarity. *Developmental Psychology*, *46*, 1799–1803.

Sundara, M., Polka, L. & Genesee, F. (2006). Language-experience facilitates discrimination of /d-/ in monolingual and bilingual acquisition of English. *Cognition*, *100*(2), 369–388. DOI:10.1016/j.cognition.2005.04.007.

Tharpe, A. M. & Ashmead, D. H. (2001). A longitudinal investigation of infant auditory sensitivity. *American Journal of Audiology*, *10*(2), 104–112. DOI: 10.1044/1059–0889(2001/011).

Trehub, S. E. & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*(1), 73–99. DOI: 10.1016/j.cognition.2005.11.006.

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

Tsuji, S. & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, *56*(2), 179–191. DOI: 10.1002/dev.21179.

Tyler, M. D., Best, C. T., Goldstein, L. M. & Antoniou, M. (2014). Investigating the role of articulatory organs and perceptual assimilation of native and non-native fricative place contrasts. *Developmental Psychobiology*, *56*(2), 210–227. DOI: 10.1002/dev.21195.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M. & Stager, C. L. (1998). Acquisition of word-object associations by 14-month old infants. *Developmental Psychology*, *34*, 1289–1309.

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*(1), 49–63. DOI: 10.1016/S0163–6383(84)80022–3.

Wong, P. C. M. & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*, 565–585.

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. DOI: 10.1016/j.jml.2012.09.004