The Development of Categorical Perception of Lexical Tones in Mandarin-speaking Preschoolers

Fei Chen 1, Nan Yan 1, Lan Wang 1, Tao Yang 1, Jiantao Wu 1, Han Zhao 1, Gang Peng 1, 2

1 Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
2 Department of Linguistics and Modern Languages, and Joint Research Centre for Language & Human Complexity, The Chinese University of Hong Kong, Hong Kong SAR

Abstract
The present study investigates how categorical perception of Mandarin tones develops along age from four- to six-year-old preschoolers. The results showed that Mandarin tones could be perceived categorically by children as young as four years of age, who were able to process tones as linguistic categories. The positions of the identification boundaries did not differ significantly between children and adults, but the boundary widths did differ significantly, with much narrower boundary widths (i.e., sharper boundaries) for the six-year-old group and the adult group. Moreover, the overall discrimination accuracies for all the children groups were lower than that for the adult group. However, children as early as five years of age could discriminate between-category tone pairs almost equally well as the matching adults. These findings lead us to conclude that five to six years of age seems to be a critical period for the development and refinement of Mandarin tone perception.

Index Terms: categorical perception, Mandarin tones, preschoolers, developmental trajectory

1. Introduction
Newborns have the ability to discriminate most phonetic contrasts of all languages [1]. However, as a result of consistent exposure to a native language, discrimination of most non-native contrasts gradually deteriorate at around 10 months, whereas native phonological contrasts are maintained and enhanced stage by stage [2, 3]. Studies about tone perception in infants have found that the discrimination performance of tonal and non-tonal language listeners started to diverge as early as 6-9 months, when tonal speech perception was reorganized due to the exposure to ambient native language [4]. Another longitudinal study [5] about tone acquisition in Cantonese indicated that perceptual discrimination of linguistic tones began as early as the 10th month. Such language-dependent reorganization of speech perception during infancy with a strong preference for phonological contrasts in their mother tongue is considered to be largely attributed to the use of statistical learning to track the frequencies of the exposed sound tokens [2].

The acquisition of tones occurs well in advance of the mastery of segmental phonemes [5, 6]. Studies on tone production have showed that children mastered most of the Mandarin tones as early as 3 years of age, and Tone 1 and Tone 4 were mastered earlier than Tone 2 and Tone 3 [7 - 9]. Studies on tonal perception showed a similar developmental course as well. In a picture-pointing task [9], 3-year-old children showed high perceptual accuracy in perceiving Tone 1, Tone 2, Tone 3 and Tone 4 (91%, 95%, 89% and 88%, respectively). It is reasonable, based on the previous research of tonal perception and production, to conclude that most Mandarin-speaking children have already acquired Mandarin tones before four years of age.

However, the aforementioned mastery of Mandarin tones in children before four years of age does not necessarily mean that they have already achieved the equal competence of tonal ability in comparison to matching adults. For instance, Yang, Diehl, and Davis [10] investigated the developmental changes in the duration of Mandarin tones produced by 5-, 8-, and 12-year-old monolingual Mandarin-speaking children, and found that tone duration decreased with age. For each tone category, 5- and 8-year-old children showed significantly longer duration than adults, while the tone duration for 12-year-old children approximated to adult values. Although 5- and 8-year-old children have already established lexical contrasts of tones, adult-like tonal production is still in the stage of development. So far, the development of tonal perception in children has been barely studied, especially using the classical paradigm of categorical perception (CP).

For more than 50 years, categorical perception (CP) has been one of the most extensively studied phenomena. For Mandarin tone perception, adult Mandarin listeners showed not only a sharp boundary between the level tone and the rising or falling tone in the identification function, but also exhibited a discrimination peak around the categorical boundary, indicating a typical categorical perception of Mandarin tones [11 - 17]. However, development of CP of tones in young children has not been extensively examined [18, 19]. It is still unknown whether lexical tones can be perceived categorically in young children who are still developing their perceptual strategy for lexical tone perception. In a behavioral study by Xi et al. [18], healthy 5- to 7-year-old children and adults were tested only on the identification task of a tone continuum ranging from Mandarin Tone 1 to Tone 2. Results showed that 6-year-old children have already acquired adult-like identification competence of Mandarin tones. Due to the lack of a discrimination test, their conclusion may not be fully conclusive. In another study [19], eight Mandarin monolingual children aged 6 to 8 years participated in tone identification and discrimination tasks. Results indicated that Mandarin monolingual children aged 6 to 8 years showed a typical categorical boundary in a tone identification task, while no
prominent discrimination peak was found near the observed identification boundary.

In summary, compared to the extensive research on the CP of tones in adults, CP of tones in young children of different ages, especially its developmental trajectory, has not been systematically studied. It is crucial to find out the critical period of the developmental trajectory and, in particular, the time point when children can perceive tones categorically in the same way as the adults.

2. Methods

2.1. Participants

Forty-two preschoolers in Shenzhen, with Mandarin as native language for the child participants and their parents, aged from 4 to 6 years, participated in this study. Fourteen children were in the 4-year-old group (seven girls; mean age = 4.5), 14 children were in the 5-year-old group (nine girls; mean age = 5.6) and 14 children were in the 6-year-old group (eight girls; mean age = 6.3). Additionally, the control group consisted of sixteen young adults (eight females; mean age = 26.2) from North China. All the 58 participants were free of neurological disease, psychiatric disorder, or hearing deficits. The consent form approved by the Behavioral Research Ethics Committee of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences was obtained from each participant.

2.2. Materials

The Mandarin syllable /i/ with the high level tone (around 290 Hz) was derived from natural recordings of a native female speaker. This female speaker was chosen because her F0 range was close to that of the female average voice [20]. Fig. 1 shows the schematic diagram of the pitch contours of the 11 stimuli (following Wang [11]; Peng et al. [17]). All the 11 tone stimuli were re-synthesized by applying the pitch-synchronous overlap and add method [21] implemented in Praat [22]. The major procedures for synthesizing the stimuli followed those described in [17]. The different F0 onsets for the target stimuli were determined by the formula “\(230 \text{ Hz} + 6 \text{ Hz} \times (\text{Stimulus Number} - 1)\)”. In Mandarin, the syllable /i/ with the high-level tone means “clothes”, and is coded as stimulus #1. However, when /i/ is spoken with the mid-rising tone, it means “aunt” coded as stimulus #1. All the tone stimuli were presented binaurally at 70 dB SPL.

2.3. Procedure

To induce young children to complete the whole identification task and discrimination task which lasted 30 min and 50 min respectively, the experimental designs were modified, and the two tasks were divided into two successive days.

2.3.1. Identification task

The identification task was conducted on the first day. Firstly, a training part was given. The experimenter instructed them to point at the left picture on a computer screen (a car driving on a level road) representing the F0 direction of stimulus #11 after playing the sound #11 several times, and point at the right picture (a car driving on a rising road) representing the F0 direction of stimulus #1 after playing the sound #1 several times. Only after children succeeded in pointing at the matched picture each time they heard stimulus #11 or #1 would they progress to the next step. The practice and test parts were monitored by E-prime. The experimenter logged the child’s pointing responses by pressing key ‘1’ (Tone 1) on the keyboard when children pointed at the left picture and key ‘2’ (Tone 2) when children pointed at the right picture (two-alternative forced choice, 2AFC). The 11 stimuli were repeated four times in a block, with two such testing blocks, totaling 88 testing trials in a random order. There was an additional practice block (including stimuli #1, #2, #5, #6, #7, #10, #11, repeating 2 times randomly) before the two testing blocks to familiarize children with the identification procedure.

2.3.2. Discrimination task

The discrimination task was conducted on the second day. In the training part, experimenter instructed them to point at the left picture (a happy face with two identical eyes) representing the same pairs after playing sounds (11-11 or 1-1) several times, and point at the right picture (a sad face with two different eyes) representing the different pairs after playing the sounds (1-1 or 1-11) several times. The practice part contained 12 pairs (5-7, 6-8, 7-5, 8-6, 1-1, 11-11, repeating two times randomly). For the testing part, 29 pairs were presented in a random order, with a 500 ms inter-stimulus interval (ISI). Among the 29 pairs, 18 pairs consisted of two different stimuli separated by two steps (i.e., 12 Hz) (different pairs), in either forward (1-3, 2-4, 3-5 … 8-10, 9-11) or reverse order (3-1, 4-2, 5-3 … 10-8, 11-9), and 11 pairs consisted of the 11 stimuli on the continuum each paired with itself (same pairs). The above 29 pairs were repeated five times, resulting in 145 testing pairs in total. Experimenter pressed key ‘v’ (for ‘same’) on the keyboard when children pointed at the left smiling picture and press key ‘n’ (for ‘different’) when children pointed at the right sad picture.

2.4. Data analysis

To investigate the developmental trajectory of the identification and discrimination performance, we calculated individual measures based on three basic characteristics of CP: position of category boundary, width of category boundary, and discrimination accuracy.

2.4.1. Identification scores

The identification score was calculated as the percentage of responses with which participants identified that stimulus as being either ‘Tone 1’ or ‘Tone 2’. The boundary position, defined as the 50% crossover points, and the boundary width as the linear distance between the 25th and 75th percentiles, were assessed by Probit analyses [23] and analyzed according to the procedures described in Peng et al. [17].
2.4.2. Discrimination scores

In order to calculate the obtained discrimination scores, we divided the 145 discrimination pairs into nine comparison cohorts (groups), each consisting of 20 pairs in four types of pairwise comparisons (AB, BA, AA, and BB) [17]. The discrimination accuracy for each comparison cohort was calculated based on the formula described in Xu et al. [15].

3. Results

3.1. Identification and discrimination curves

Identification and discrimination curves for the different age groups are shown in Fig. 2.

3.2. Position and width of categorical boundary

The estimated boundary position and width, obtained by Probit analysis, are shown in Table 1. Moreover, the distribution of boundary width among different age groups is shown in Fig. 3.

Table 1: Derived categorical boundary position and width for each age group.

<table>
<thead>
<tr>
<th>Group</th>
<th>Position</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>4 yr</td>
<td>5.81</td>
<td>0.98</td>
</tr>
<tr>
<td>5 yr</td>
<td>5.63</td>
<td>0.59</td>
</tr>
<tr>
<td>6 yr</td>
<td>5.66</td>
<td>0.56</td>
</tr>
<tr>
<td>Adult</td>
<td>5.90</td>
<td>0.93</td>
</tr>
</tbody>
</table>

The mean boundary positions for the groups of the four years of age, five years of age, six years of age and adults were 5.81, 5.63, 5.66, and 5.90, respectively. Results of one-way ANOVA revealed that the perceptual boundary position was not significantly different among different age groups ($F (3, 54) = 0.401; p = 0.752$). Moreover, the average boundary widths for the four groups were 2.03, 1.93, 1.22, and 1.01, respectively. One-way ANOVA showed that the identification widths among different age groups were significantly different ($F (3, 54) = 12.133; p < 0.001$). Tukey’s HSD post hoc pairwise comparisons of the four groups indicated that both the 6-year-old group (both $p < 0.01$) and the adult group (both $p < 0.001$) had significantly narrower boundary widths than the 4-year-old group and 5-year-old group, but that the boundary widths between the 4- and 5-year-old groups ($p = 0.961$) were not significantly different, nor the boundary widths between the 6-year-old and adult groups ($p = 0.731$).

3.3. Discrimination accuracy

The discrimination accuracies of nine comparison cohorts among four groups are shown in Fig. 4. A 9 (comparison cohort) × 4 (group) repeated measures ANOVA was conducted to examine whether the four age groups showed different discrimination accuracies, with the comparison cohort as a within-subject factor and with the group as a between-subject factor. Where appropriate, the Greenhouse-Geisser method was used to correct violations of sphericity. The analysis confirmed a significant main effect of cohort ($F (8, 408) = 31.866; p < 0.001$), and a significant main effect of group ($F (3, 51) = 15.76; p < 0.001$). Post hoc multiple comparisons of the four groups indicated that the overall discrimination accuracy in adults was significantly higher than the other three children groups (all $p < 0.001$), while the overall discrimination accuracies between any two of the three children groups were not significantly different (all $p > 0.05$).

As shown in Fig. 4, the accuracy for the four groups reached their maxima around the pair 5-7, which also corresponded to the respective categorical boundary. To test the between-category discrimination sensitivity, one-way ANOVA was conducted and revealed a significant difference in accuracy at pair 5-7 across the four age groups ($F (3, 51) = 3.625; p < 0.05$). Tukey’s HSD post hoc comparison revealed that the between-category accuracy at pair 5-7 only for the 4-year-old group, 58.8%, was significantly lower than that for the adult group, 73.33% ($p < 0.05$). No significant difference of between-category accuracy at pair 5-7 between the 5- or 6-year-old group and the adult group ($p = 0.266$ and 0.707 respectively) was observed.

Figure 2: Identification curves (solid lines, with round dots representing Tone 2 responses and square dots representing Tone 1 responses respectively) and discrimination curves (dashed lines with triangle dots) across different age groups. The left y-axis indicates the percentage score of Tone 2 or Tone 1 responses, while the right y-axis indicates the discrimination accuracy.
The development of tone identification

In the present study, different age groups did not show significantly different boundary positions, which were ranging from 5.63 to 5.90 (see Table 1). However, both the 6-year-old group and the adult group exhibited significantly narrower boundary widths (i.e., sharper boundaries) than the 4- and 5-year-old groups (see Table 1 and Fig. 3), indicating that 6-year-old children have already acquired adult-like identification competence of Mandarin tones. This finding was consistent with another identification study by Xi et al. [18], although the 6-year-old children in their research have explicitly learned about the Mandarin tones in the first grade of primary school. What we would like to emphasize here is that all the preschoolers in our research did not explicitly study the conception of Mandarin tones in kindergarten, and, in the identification task, they were just asked to judge the sound as either a level or rising one. Consequently, it is reasonable to conclude that around six years of age is the critical period for the development of tone identification. During this developmental period, children’s competence in Mandarin tone identification started to reach the adult level.

The development of tone discrimination

As shown in Fig. 4, four-to-six-year-old preschoolers lag far behind adults in the overall performance of the discrimination accuracy of the nine comparison cohorts. On one hand, the discriminability may develop more slowly than the identification ability, since the discrimination task involves a more complicated multi-store model consisting of unanalyzed and analyzed sensory memory, short-term and long-term categorical memory [15]. On the other hand, as put forward by a recent study [24] exploring the just noticeable difference (JND) of F0 contour, Mandarin adult listeners required nearly 10 Hz to detect a tonal pitch change. Given that young children generally require greater JND in F0 discrimination than adult listeners [25, 26], preschoolers may need a F0 difference of more than 12 Hz used in this study to facilitate the discrimination task. However, children as young as five years of age showed similar between-category accuracy as the adults at the pair 5-7, indicating that the overall weaker discriminability in young children had no decisive effect on the between-category discrimination sensitivity, which is crucial for distinguishing different tone categories.

4. Discussion

4.1. The development of tone identification

4.2. The development of tone discrimination

As defined by Liberman et al. [27], typical categorical perception should exhibit a sharp boundary between two categories and a corresponding discrimination peak around the identification boundary. In this study, prominent peaks were aligned with the corresponding identification crossovers for all the four groups (see Fig. 2). Consequently, Mandarin tones can be perceived categorically in children as young as four years of age, who were able to process lexical tones as linguistic categories. The development and refinement of CP from infancy to young children is presumably influenced by both physiological maturation and accumulated exposure to the spoken language, which cause within-category differences to become less salient, thereby preventing non-relevant acoustic information, e.g. pitch variations of the same tone category, from reaching the mental lexicon [17]. Therefore, CP, which develops as the child ages, helps them learn the location of tone category boundaries to master the many-to-one mapping between acoustic patterns and phonological categories, and accordingly enhances spoken communication.

5. Conclusions

We have examined categorical perception of Mandarin tones in 4- to 6-year-old preschoolers. From the identification curves and the distribution of boundary widths shown in Figs. 2 and 3, we can see clearly that changes from one tonal category to another are more abrupt from six years of age. Around six years of age is the critical period for the refinement of tone identification. From the discrimination scores shown in Fig. 4, although the overall discrimination accuracy of young children is much lower than that of adults, the 5-year-old children can perform equally well as matching adults do in the discrimination performance of the between-category pair 5-7. Moreover, Children as early as four years of age, with plenty of exposure to Mandarin, can perceive Mandarin tones categorically.

6. Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (NSFC: 61135003, 11474300, 91420301 and 61401452), Shenzhen Speech Rehabilitation Technology Laboratory, and Shenzhen Fundamental Research Program JCJ20130401170306806. We thank all the preschoolers, from Guangming Central Kindergarten and Taoyuan Secondary Kindergarten, and their parents for their participation and cooperation.
7. References


