<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Frequency discrimination of speech and nonspeech sounds by children</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Other Contributor(s)</strong></td>
<td>University of Hong Kong.</td>
</tr>
<tr>
<td><strong>Author(s)</strong></td>
<td>Kung, Wan-sum, Anita; 龔韻心</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Kung, W. A. [龔韻心]. (2005). Frequency discrimination of speech and nonspeech sounds by children. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2005</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/48743">http://hdl.handle.net/10722/48743</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; The author retains all proprietary rights, such as patent rights and the right to use in future works.</td>
</tr>
</tbody>
</table>
Frequency discrimination of speech and nonspeech sounds by children

Kung Wan-Sum Anita

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (speech and Hearing Sciences), The University of Hong Kong, May 6, 2005
Abstract

This study investigated the effect of age and type of stimuli on children’s performance in fundamental frequency (F₀) discrimination. Sixteen subjects were tested in each of 6 age groups: 3, 4, 5, 6, and 10 years old and adults. The stimuli were the syllable [lou22] and a hum with the same F₀ and amplitude characteristics as the speech stimulus. Subjects performed a 3-interval, 2-alternative forced choice task. Fundamental frequency difference limens (F₀DLs) were estimated using the 2-down, 1-up adaptive procedure. Results showed that F₀DLs decreased with age. Adult performance was achieved at age 10. The results were discussed in relation to previous studies on frequency discrimination and tone perception. No significant difference was found between speech and nonspeech tasks, meaning that F₀ difference was perceived equally well in linguistic and nonlinguistic stimuli.
Frequency discrimination of speech and nonspeech sounds by children

There are a number of functions associated with the F₀ of speech. Being able to perceive changes in F₀ enables listeners to identify the gender and age of speakers; and to identify intonation information such as imperative and interrogation sentences (Vance, 1976). For the linguistic purpose, F₀ can be used for recognizing lexical tones in tone languages (Bauer & Benedict, 1997). Therefore, researchers have been interested in finding out the smallest difference in frequency or F₀ (frequency difference limens FDLs, or F₀DLs) human could perceive (Klatt, 1973; Maxon & Hochberg, 1982; Olsho, Koch & Halpin’s, 1987; Olsho, Schoon, Sakai, Turpin & Sperduto, 1982; Sinnott & Aslin, 1985; Thompson, Cranford and Hoyer, 1999).

Pure tone frequency discrimination in infants

There has been research that investigated pure tone frequency discrimination in human infants. It was found that infants showed significantly larger FDLs than adults at 1000Hz, 2000Hz, and 3000Hz (Olsho et al., 1982; Sinnott & Aslin, 1985). Olsho et al. found average FDLs for infants aged 5-8 months to be 21.6Hz, 36.3Hz, and 54.9Hz at 1000Hz, 2000Hz, and 3000Hz respectively. The adult group attained FDLs of 7.4Hz, 16.7Hz, and 32.3Hz respectively. In Sinnot and Aslin’s study (1985), FDLs for 7- to 9-month-old infants at 1000Hz were comparable with those found in Olsho et al.’s study. In another study, Olsho et al. (1987) investigated the FDLs for 3-, 6-, and 12-month-old infants, and adults at 500Hz, 1000Hz, and 4000Hz. The results showed that at 500Hz and 1000Hz, FDLs for all infant groups were significantly larger that those of adults. At 4000Hz, 3-month infants had significantly larger FDLs than the 6- and 12-month infants, who had comparable FDLs with those of adults.
Although Olsho et al.’s study (1987) evidenced early maturation of frequency discrimination ability at 4000Hz in 6- and 12-month-old infants, results from various studies indicated that frequency discrimination ability at infancy had not reached adult performance (Olsho et al., 1982; and Olsho et al., 1987; Sinnott & Aslin, 1985). Therefore, the ability to discriminate frequency was at an emerging stage during the infancy, for frequencies below 4000Hz.

**Pure tone frequency discrimination in children**

Thompson et al. (1999) investigated the pure tone FDLs at 1000Hz with various tone durations in children aged 5 to 11. In their study, each trial consisted of two intervals, each with two tones. The subjects responded by choosing the interval which contained two tones with different frequencies. Eleven out of 16 5-year-old children were not able to learn the task or perform it consistently; therefore statistically analysis was not applied to this group. The results showed that the 7-year-old children obtained significantly larger FDLs than the 9- and 11-year-old children, and adults. The estimated FDLs of the 9- and 11-year-old children were comparable with those of adults, meaning that adult performance was achieved at about 9 year of age. FDLs for the 5-, 7-, 9-, 11-year-old children, and adults were estimated to be about 25.2Hz, 10.6Hz, 4.1Hz, 3.4Hz, and 4.8Hz respectively, for 200ms stimuli. However, the FDL of the 5-year-old group was not representative for this age group. It was obtained from only 5 out of 16 of the subjects in this group, and the reason for not including the other 11 subjects was related to the nature of the test. Therefore, the 5-year-old group with the 5 remaining subjects was considered to be a biased sample (Shaughnessy & Zechmeister, 2000). The authors commented that the 5-year-old children who failed to learn the task were able to perform a similar visual
discrimination task during the training, but they were unable to perform the task with auditory stimuli. Since four tones were presented in each trial, children had to memorize them within a short period of time before making the required judgment. It was possible that the 5-year-old children had less mature memorizing ability, which limited their performance in that task.

Maxon and Hochberg (1982) investigated FDLs of children aged 4 to 12. They found that FDLs at 500Hz, 1000Hz, and 4000Hz decreased with age. FDLs for 4-, 6-, 8-, 10-year-old children were about 15.9Hz, 11.6Hz, 7.4Hz, 5.1Hz, and 4.7Hz at 500Hz respectively. FDLs were estimated to be about 12Hz, 9.6Hz, 7.1Hz, 4.4Hz, and 3.7Hz at 1000Hz respectively. In their study, subjects were presented two tones. They were required to determine whether the tones were the same or different in pitch. However, the concept of same or different was difficult for young children. Also, requiring them to give a verbal labeling in their response further complicated the task.

From the results of various studies on pure tone frequency discrimination in children, it is observed that FDLs decrease steadily with age. The frequency discrimination ability did not mature, or reach adult’s level, until late childhood.

*Complex tone F₀ discrimination*

There has been limited number of researches documenting developmental data on F₀ discrimination of complex tones. Although some researchers studied F₀ discrimination in children with learning impairment and dyslexia (Marshall, Snowling & Bailey, 2001; Waber et al., 2001), the focus of those studies was to investigate the ability to process rapid auditory stimuli. The discrimination pair used in those studies had large F₀ differences (i.e. 80Hz, 205Hz).
Therefore, no information for FøDLs in children was provided.

For adults, FøDLs were found to be 0.3Hz and 0.5Hz for vowel (V) [ε] and consonant-vowel (CV) syllable [ya] respectively with steady level of Fø at 120Hz (Klatt, 1973). In the task, a standard tone was presented, followed by a comparison tone in each trial. The subjects were required to determine whether the comparison tone was higher or lower in pitch than the standard tone. This task was only suitable for adults as it involved the judgment of high or low pitch. However, the adjective concept of high and low has not been acquired until the age of 3; 06 (Opper, 1996).

*Significance of Fø discrimination ability for tone language speakers*

The ability to discriminate small Fø difference is especially important for speakers of tone languages in which the change of tone in a segment brings about lexical contrast (Bauer & Benedict, 1997). According to Bauer and Benedict (1997), there are six contrastive tones in Cantonese including level tones, rising tones, and falling tones. Tone contrasts are perceived through the dimension of Fø level and contour. There has been research that investigated tone perception using an identification task (Ching, 1984; Ciocca & Lui, 2002; Sze, 2004). In Ciocca and Lui’s study (2002), children and adult subjects were presented a target word in an embedded phrase. They were required to match the spoken word with one of the two pictures representing a minimal monosyllabic word pair differed by tone. The results showed that children’s performance in tone identification improved with age up to the age of 10, at which adult’s level of performance was acquired. Also, they found that one of the most difficult tone contrasts to be identified was the mid level-low level pair, for both children and adults. Such difficulty was
attributed to the closeness in F₀ of the tone pair (Ching, 1984; Ciocca & Lui, 2002; Sze, 2004). Besides, there was significant improvement in identifying the mid level-low level contrast between the 4- and 6-year-old groups, and between the 6-year-old and the adult groups (Ciocca & Lui, 2002). Since the contour of fundamental frequency was not a relevant cue in identification of the mid level-low level pair, the difference in F₀ level was the sole basis of tone identification. Therefore, it is possible that the improvement in tone identification of mid level-low level contrast reveals the development in the basic auditory ability of perceiving the F₀ level of sounds.

Sze (2004) replicated Ciocca and Lui’s study (2002) to investigate the word familiarity effect on tone identification. The age group 2;09-3;03 was added. The results showed that word familiarity only affected performance of the 2;09-3;03 group. Major results were in general agreement with those of Ciocca and Lui’s study (2002). Additionally, there was significant improvement in ability to identify the high level-mid level tone pair between the 2;09-3;03 and the 4;00-4;11 groups. This result provided additional evidence of development of tone perception ability in early childhood.

*Rationale of current study*

While there are plenty of studies providing information on FDLs, much less research has been explored the F₀ discrimination of complex tones. Moreover, studying the developmental change in F₀DLs at F₀ similar with that of human’s voice can provide information on the development of children’s tone perception ability, especially for tokens which have close level of F₀.
Previous studies on frequency discrimination were conducted with infants (Olsho et al., 1982; Olsho et al., 1987; Sinnott & Aslin, 1985) and 4- to 12-year-old children (Maxon & Hochberg, 1982; Thompson et al., 1999). On the other hand, recent studies on auditory judgment showed that children as young as 2;09-3;03 were able to give reliable response (Lee, Chiu, & Hasselt, 2002; Sze, 2004). It was also found that between the age of 2;09-3;03 to 4;00-4;11, there was significant improvement in tone perception ability (Sze, 2004). Therefore, in the current study, 3-year-old children were included as the youngest group in order to investigate F0 discrimination ability in early childhood. A 3-interval, 2-alternative forced-choice task was adopted from McArthur and Bishop (2004). This test procedure minimized the need of memorization of the long stimuli. Also, it did not require the understanding or verbal labeling of high and low or same and different concepts.

Comparison of speech and nonspeech stimuli

Speech and nonspeech stimuli have been used in Francis and Ciocca’s study (2003), which investigated the selective sensitivity in stimuli presentation order in F0 discrimination task. The study showed that Cantonese speakers were better at discriminating the small F0 difference in speech stimuli of about 4Hz when the second token had a higher F0 than the first one (presentation order effect). However, this effect did not occur in the discrimination of nonspeech stimuli. Neither did this effect occur for American-English-speaking listeners in both speech and nonspeech discrimination. The authors suggested that the experience in perceiving Cantonese tone accounted for the occurrence of the stimuli presentation order effect. Since there is fall in pitch towards the end of an utterance, Cantonese-speaking listeners have to consider the position
Frequency discrimination of a word in an utterance, and carry out perceptual adjustment in order to achieve accurate lexical tone identification in that part. In Francis and Ciocca’s study (2003), the Cantonese-speaking subjects were thought to adjust the pitch of the second token as if they were required to identify the tone of a word appeared at the end of the utterance. The adjustment was made by raising the pitch of the second token during the perception. Therefore, when the second token had a higher Fø than the first one, the adjustment made the perceptual Fø difference between the tokens more obvious, which allowing the subjects to be more able to discriminate the small Fø difference. This effect indicated that the Cantonese-speaking subjects perceived the Fø level of the speech stimuli in the same way as perceiving lexical tones in real speech.

However, the stimuli presentation order effect was not observed in discrimination of nonspeech stimuli. Thus, the Fø of the nonspeech stimuli was not perceived as lexical tones. Besides, when comparing the performance in Fø discrimination of speech and nonspeech stimuli, Francis & Ciocca’s study (2003) showed that Cantonese speakers were less sensitive to Fø difference in speech stimuli than nonspeech stimuli. A similar pattern was also found in the performance of the American English-speaking subjects. Therefore, the authors proposed that the difference in sensitivity may be due to the difference in the complexity of stimuli rather than the nature of stimuli.

Another rationale of current study

Since Fø is one of the perception dimensions of Cantonese tones which is used to mark lexical contrasts, FøDLs in speech stimuli has a linguistic status for Cantonese-speaking listeners. As revealed by Francis and Ciocca’s study (2003), Fø level of the speech stimuli was processed
in the same way as lexical tones, that meant it was perceived and interpreted as linguistic materials. In contrast, the $$F_0$$ level of the nonspeech stimuli was not processed in the same way, it was interpreted as nonlinguistic materials. Therefore, comparing $$F_0$$DLs of speech and nonspeech stimuli allows making a comparison between linguistic and nonlinguistic pitch processing. Since conclusion on this issue could not be obtained from Francis and Ciocca’s study (2003), this issue was addressed in the current study.

**Research questions**

In the current study, two research questions were focused. The first research question was to investigate the age effect on Cantonese-speaking children’s ability to discriminate $$F_0$$ of complex tones with $$F_0$$ similar to that of human speech. The age at which children’s ability to discriminate $$F_0$$ of complex tones reach adult’s performance was investigated. Kent (1997) suggested that there was continuing refinement in auditory discrimination ability beyond the infancy. This claim was confirmed by the finding of decreasing FDLs with age in children (Maxon & Hochberg, 1982). Therefore, it was hypothesized that the estimated FDLs would decrease with age. The second research question was to investigate the effect of type of stimuli, speech and nonspeech, on the ability to discriminate $$F_0$$ of complex sounds. That is, to compare the linguistic and nonlinguistic processing of $$F_0$$ level of complex tones.

**Methodology**

**Participants**

One hundred and sixteen participants (54 males, 62 females) were recruited. The participants were selected from 6 age groups: 3;00 to 3;11; 4;00 to 4;11; 5;00 to 5;11, 6;00 to
6;11, 10;00 to 10;11, and adults (>18;00), there were 23, 23, 17, 19, 16, 18 subjects in these groups respectively. The first age group was chosen since children as young as 2;09-3;03 were found to be able to give reliable response in auditory task (Lee, Chiu, & Hasselt, 2002; Sze, 2004). The 4-, 6-, and 10-year-old groups were chosen since significant improvement in tone identification was found between these ages (Ching, 1984; Ciocca & Lui, 2003; Sze, 2004). Since the most remarkable improvement was found between the age of 4 and 6, the inclusion of the 5-year-old group revealed more information about the improvement. Adult was chosen to serve as a comparison group to indicate the age at which children achieve adult performance.

All participants were native Cantonese speakers. Child participants were selected randomly from one normal kindergarten, and one normal primary school. The participants received a hearing screening of pure tone at octave frequencies between 250Hz and 8000Hz. The screening level was set at 35dB HL due to background noise at the testing environment (average noise level at 40dB SPL). Participants should pass a training before taking the test.

Among the participants recruited (n=116), 9 (8%) failed the hearing screening. There was 1 such participant from each of the 3- and 5-year-old groups, 2 from each of the 6-year-old and adult groups, and 3 from the 4-year-old group. Besides, 6 participants from 3-year-old group, 4 participants from 4-year-old group, and 1 participant from 6-year-old group did not pass the training. Ninety six participants were remained. There were 8 males and 8 females in each age group.

**Stimuli**

Speech and nonspeech stimuli with an average F₀ of 104Hz, and duration of 300ms were
used as the standard tones. This Fø was chosen since it represents typical Fø of males (Bauer & Benedict, 1997). The stimuli were synthesized from the syllable [lou22] produced in a carrier phrase /ngoh23 wui33 duk2 – bei35 nei23 teng55/ (‘I will read – for you to hear’) by a native male speaker of Cantonese. For speech stimuli, 11 tokens were produced by increasing the Fø of the standard tone by the step sizes shown in Table 1 using the Praat software. Table 1 shows the Fø values for these tokens. The PSOLA algorithm was used to resynthesize the stimuli. The nonspeech stimuli were humming sounds resynthesized with the same Fø and intensity contour as the speech stimuli. The stimuli were then low pass filtered at a cutoff frequency of 1900Hz. After that, the stimuli undergone pre-emphasis twice in order to make them sound non-speech like. Both types of stimuli were normalized to get the maximum amplitude without clipping.

Procedure

A 3-interval, 2-alternative forced-choice task (McArthur & Bishop, 2004) was used. Each trial was composed of 3 consecutive intervals. The interstimulus intervals were 500ms. The first interval was always the standard tone of 104Hz Fø. Either the second or the third interval contained the standard tone. The remaining interval was the comparison tone with a Fø higher than the standard tone.

A computer program was created using the Revolution 2.2.1 software development package (Runtime Revolution Ltd, 2000-2003) for the presentation of the visual and auditory stimuli. The standard tone was represented by a jumping chick on the computer monitor. The second the third intervals were represented by two identical eggs, which appeared on the left side and right side of the computer monitor respectively. The chick and eggs appeared successively during the
Table 1

*Step size, fundamental frequency values for stimuli token/comparison tone*

<table>
<thead>
<tr>
<th>Step size (Hz)</th>
<th>F₀ of stimuli token/comparison tone (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>76.8</td>
<td>180.8</td>
</tr>
<tr>
<td>38.4</td>
<td>142.4</td>
</tr>
<tr>
<td>19.2</td>
<td>123.2</td>
</tr>
<tr>
<td>9.6</td>
<td>113.6</td>
</tr>
<tr>
<td>4.8</td>
<td>108.8</td>
</tr>
<tr>
<td>2.4</td>
<td>106.4</td>
</tr>
<tr>
<td>1.2</td>
<td>105.2</td>
</tr>
<tr>
<td>0.6</td>
<td>104.6</td>
</tr>
<tr>
<td>0.3</td>
<td>104.3</td>
</tr>
<tr>
<td>0.15</td>
<td>104.15</td>
</tr>
<tr>
<td>0.075</td>
<td>104.075</td>
</tr>
</tbody>
</table>

presentation of the corresponding intervals, their presentation was synchronized with the presentation of the auditory stimuli.

The listeners heard the auditory stimuli at 70dBA bilaterally through Sennheiser HD-590 headphones. They were required to determine whether the second or third interval contained the signal tone by pointing to the egg which had the same sound as the standard tone presented with the jumping chick. Then, the experimenter clicked on the egg with the computer mouse to record
the response. Visual feedback was provided. A correct response was rewarded with an animation showing the chick coming out of the egg. For wrong responses, the egg shook from side to side but no chick came out of the egg.

The F0 differences were adjusted between successive trials using an adaptive procedure (Levitt, 1970). The adaptive procedure was chosen since it was sensitive to changes in response parameter and enabled the test to be finished in relatively small number of trials (Levitt, 1970). In particular, the 2-down, 1-up adaptive procedure was adopted. Two correct responses lead to 1 step decrease in F0 difference, while 1 step increase in F0 difference was brought about by 1 incorrect response, or 1 correct response followed by 1 incorrect response. Table 1 shows the step size and the F0 values of the corresponding comparison tone. The F0 difference between signal and comparison tones was set at 38.4Hz initially. For example, if a participant got a trial wrong, then the F0 difference of next trial was the double of that trial. If a participant got 2 correct trials, the F0 difference of the following trial was the half of those trials. The test was terminated after eight response reversals.

The participants were required to take 2 sessions: speech and nonspeech. The order of presentation of these sessions was counter balanced across participants for each age group. Listeners received the test individually in a quiet room of the kindergarten or school they attended. The adult participants were tested in a quiet room.

Each participant was given training before the test to familiarize them with the test procedure. In the training, the F0 difference between signal tones and comparison tones was fixed at 45Hz. The training began by a maximum number of 5 demonstrative trials done by the
experiment, and then followed by training trials with visual feedback. The participants were required to achieve 7 consecutive correct responses within 20 trials to pass the training. The participants taking speech session first received training with speech stimuli, while those who took nonspeech session first had been given nonspeech stimuli in training. Before taking the second session, the participants were allowed to practice the training of the other type of stimuli for at least 1 trial to familiarize themselves with the stimuli. The same visual feedback was given.

Results

The FøDLs were estimated by calculating the average of the 4 smallest consecutive reversals within 3 steps in Fø difference. In 2 types of case, the 4 smallest consecutive reversals within 4 steps were accepted. First, for 4 subjects, the smallest reversals obtained included 0.075Hz and 0.3Hz, 4-step increment from these reversals were 1.2Hz and 4.8Hz respectively. Since the fluctuation from 0.075Hz to 1.2Hz, and from 0.3Hz to 4.8Hz were reasonable, 4-step fluctuation was accepted in these cases. Second, subjects got reversals of 38.4Hz or 76.8Hz for occasional inattentiveness, adjustment was made in order to avoid the inclusion of these reversals or to include the least amount of them. Therefore, in 2 subjects, 4-step fluctuation from 1.2Hz to 19.2Hz was preferentially selected rather than 3-step fluctuation from 4.8Hz to 38.4Hz for calculating FøDLs. In another 2 subjects, 4-step fluctuation from 2.4Hz to 38.4Hz was selected rather than 3-step fluctuation from 4.8Hz to 38.4Hz with two reversals of 38.4Hz. After making this adjustment, data from either speech or nonspeech session which were considered to reflect inattentiveness were excluded from the estimation of mean FøDLs. The exclusion criteria
were either the containment of 3 or more reversals of 76.8Hz in the session, or 1 or more
reversals of 76.8Hz in those selected for estimating FøDL, which could not be avoided with the
above adjustment. In total, 3 speech sessions, 1 from each of the 4-, 6-, and 10-year-old groups;
and 4 nonspeech sessions, 1 from each of the 3-, 4-, 5-, and 10-year-old groups were excluded.
This involved 1 subject from each of the 3-, 4-, 5-, and 6-year-old groups; and 2 subjects from
the 10-year-old group.

Mean FøDLs for each age group were calculated from all the remaining data after the
exclusion of 3 speech sessions and 4 nonspeech sessions. Table 2 shows the mean FøDLs for
speech and nonspeech stimuli and all age groups. There was a trend of gradual decrease in
FøDLs from the 3-year-old group to the adult group. More precisely, the data showed a decrease
in FøDLs from 3- to 4-, 5-, and 6-year-old groups. The 4-, 5-, and 6-year-old groups attained
similar FøDLs. Also, FøDLs further decreased from 4-, 5-, and 6-year-old groups to 10-year-old
group, and slightly decreased from 10-year-old group to adult group. However, the data of 3-, 4-,
and 6-year-old groups should be interpreted with caution. The data were obtained from a portion
of the recruited subjects in these age groups; those who failed the training were not included in
the test.

The results were analyzed using factorial analysis of variance (ANOVA) to compare the
performance between age groups, and between the natures of stimuli. The interaction
between age and type of stimuli was not significant, F (5, 83) = 0.3, p > .05. The main effect
for type of stimuli was not significant, F (1, 83) = 0.1, p > .05. Figure 1 shows the mean
FøDLs (Hz) for speech and nonspeech stimuli for all age group after excluding 6 subjects.
Table 2

*Mean FøDLs (Hz) and standard deviation for all age groups (after the exclusion of data from 3 speech and 4 nonspeech sessions)*

<table>
<thead>
<tr>
<th>Age (yrs)</th>
<th>Speech Mean (S.D.)</th>
<th>Nonspeech Mean (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>19.0 (9.7)</td>
<td>16.5 (10.1)</td>
</tr>
<tr>
<td>4</td>
<td>11.3 (7.7)</td>
<td>12.1 (8.0)</td>
</tr>
<tr>
<td>5</td>
<td>9.9 (8.6)</td>
<td>10.0 (9.1)</td>
</tr>
<tr>
<td>6</td>
<td>10.8 (7.4)</td>
<td>12.2 (9.5)</td>
</tr>
<tr>
<td>10</td>
<td>5.1 (6.3)</td>
<td>7.3 (8.2)</td>
</tr>
<tr>
<td>Adult</td>
<td>2.3 (2.5)</td>
<td>2.8 (3.3)</td>
</tr>
</tbody>
</table>

There was significant main effect of age, $F (5, 83) = 9.8, p < .05$. Results of the Post-hoc Tukey HSD showed support to the descriptive statistics. It was found that the FDLs of 3-year-old group were significantly larger than those of 5-year-old group and adult group, $p < .05$. Also, FDLs of all age groups except the 10-year-old group were significantly larger than those of adult group, $p < .05$. There was no significantly difference in other combinations of age.

To summarize, the results showed that the FøDLs decreased with age. The ability of Fø discrimination improved significantly from age 3 to 5, and from age 5 and 6 to adult. The adult’s level of performance was achieved at the age of 10. There was no significant difference between
the performance in discrimination of speech and nonspeech stimuli in all age groups, which meant linguistic and nonlinguistic processing of Fø level of complex tones were done equally well in the present study.

<table>
<thead>
<tr>
<th>age</th>
<th>Speech</th>
<th>Nonspeech</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>18.97</td>
<td>16.46</td>
</tr>
<tr>
<td>4</td>
<td>11.25</td>
<td>12.12</td>
</tr>
<tr>
<td>5</td>
<td>9.871</td>
<td>9.95</td>
</tr>
<tr>
<td>6</td>
<td>10.79</td>
<td>11.78</td>
</tr>
<tr>
<td>10</td>
<td>5.14</td>
<td>7.345</td>
</tr>
<tr>
<td>a</td>
<td>2.259</td>
<td>2.82</td>
</tr>
</tbody>
</table>

Figure 1. Mean FøDLs (Hz) for speech and nonspeech stimuli (after the exclusion of data from six subjects)

Discussion

Effect of age on Fø discrimination and the age achieving adult performance

In general agreement with a previous study on frequency discrimination in children (Maxon & Hochberg, 1982), the current study showed that FøDLs decreased with age. The results confirmed Kent’s suggestion (1997) of continual improvement in the ability to discriminate
auditory stimuli beyond the infancy. Precisely, F\o DLs estimated in the current study for the 10-year-old group and adult group were in close agreement with FDL obtained for 9-year-old group and adult group in Thompson et al.’s study (1999). However, F\o DLs estimated for 5-year-old group were smaller than FDL obtained in Thompson et al.’s study. Also, F\o DLs for 6- and 10-year-old group were similar with FDLs of 7- and 10-year-old groups in Maxon and Hochberg’s study (1982), while F\o DLs for 4-year-old group were smaller than FDLs.

The smaller F\o DLs obtained for the young age groups reflected the difference in the test paradigms used. The 3-interval, 2-alternative forced choice procedure used in the current study did not require verbal response, or the judgment of same or different when compared to Maxon and Hochberg’s study (1982). Also, in the current study, the participants did not have to memorize the long stimuli (each with 4 tones) as in Thompson et al.’s study (1999). This procedure eliminated the need of judging same or different and reduced the cognitive requirement in memorization, making this procedure more able to obtain children’s optimal performance, especially for young children. Similar F\o DLs at 104Hz and FDLs at 500Hz and 1000Hz was found for older children and adults. This finding could be explained in terms of the similarity between the underlying mechanisms in the pitch perception of pure tone and complex tone with resolved harmonics (Grimault, Micheyl, Carlyon, Bacon, & Collet, 2003). Pitch perception of complex tone involves first analyzing the complex tone into its frequency components, and then the pitch is determined by recognizing the pattern of the frequencies of the resolved components. In the current study, the complex tone at 104Hz was resynthezied from natural speech, it contained energy at 520Hz (the 5th harmonics) and 1040Hz (the 10th harmonic).
Therefore, FøDLs at 104Hz could be similar with FDLs of its component frequencies, at around 500Hz and 1000Hz.

However, the current study found adults’ FDL for speech stimuli to be 2.3Hz and 2.8Hz, which was larger than the value of 0.3Hz and 0.5Hz estimated at 120Hz Klatt’s study (1973). This could be explained by the difference in research methodologies in terms of procedure and the definition of FøDLs. In Klatt’s study (1973), 100 trials were obtained from each subject at each Fø difference of 0.15Hz, 0.3Hz, 0.6Hz, and 1.2Hz. The subjects responded by stating whether the second tone was higher or lower in pitch than the first one, and then the percentage accuracy was plotted. The Fø difference at which 75% accuracy was achieved was defined as the FøDL. In the current study, each adult subject participated in a 2-down 1-up procedure, and 8 reversals were obtained. FøDLs were calculated from the average of 4 smallest consecutive reversals within 3 or 4 steps. Therefore, in the current study, relatively fewer trials were obtained for each Fø difference. FøDLs were determined by their performance on few exposures with the stimuli. While in Klatt’s study (1973), listeners had a larger number of exposures to each Fø difference, they had more chance to practice responding to the Fø difference, which facilitated their optimal response. Besides, in the current study, any wrong response lead to one step increase in Fø difference, the Fø difference of the wrong trial became one of the reversals. In this procedure, respond of each trial directly affected the FøDL. Since in Klatt’s study (1973), there was larger number of trials, occasional mistakes had a smaller count towards the FøDL than in the current study. Besides, larger FøDLs in current study could be partially explained in terms of difference in stimuli. Klatt’s study (1973) utilized stimuli with V [e] and CV [ya] structure while
the current study used real word with CVV [lou] structure and a hum. The CV syllable yielded a larger FðøDL than V syllable in Klatt’s study (1973), which suggested that unstable formant frequencies may affect Fðø discrimination. It was possible that CVV syllable used in the current study yielded even larger FðøDL than CV syllable since it had more unstable formant frequencies.

The age achieving adult performance was found to be 10 in the current study. It was in general agreement with Thompson et al. (1999), who showed that the adult performance was achieved at the age of 9. However, the finding was not in accord with the suggestion of Maxon and Hochberg (1982), which stated that children did not achieve adult performance until the age of 12. Maxon and Hochbery (1982) reached their conclusion by comparing the performance of their child subjects with that of adult subjects in Harris’ study (as cited in Maxon & Hochberg, 1982). However, they did not mention whether the methodologies of the two studies were the same. It would be more reliable to compare children and adults performance with the same procedure. Therefore, adult’s level of performance in frequency discrimination was considered to be achieved at the age of 10.

*Comparing results in frequency discrimination with frequency identification*

The purpose of studying Fðø discrimination of complex tone at Fðø of human speech was to revealed more information on the development of tone perception ability. The current study was consistent with Ciocca & Lui’s study (2003) that there was room for improvement in tone perception ability after the age of 6 before children achieved adult’s level of performance. Besides, the current study showed that initiation maturation of tone perception ability on tone discrimination task was found between the age of 3 and 5. This trend was reflected in significant
improvement in identification of high level-mid level contrast between the age of 2;09-3;03 and 4;00-4;11 showed in Sze’s study (2004). The current study found that children at age 3 obtained mean FøDL of 19.0Hz for speech stimuli. With this ability, they would be able to identify high level-mid level contrast fairly well. The results in Sze’s study (2004) supported this finding; 3-year-old children achieved about 70% accuracy in identification of high level-mid level contrast. When the children reached the age of 4, their FøDL reduced to 11.3Hz for speech stimuli. Then they would be able to identify the mid level-low level in almost every encounter with it. Sze’s study (2004) provided evidence to this finding, since the 4-year-old subjects identified high level-mid level contrast with about 90% accuracy.

Nevertheless, there was discrepancy in the ages at which rapid improvement in tone perception ability occurred. In the current study, 4- and 6-year-old children performed equally well in tone discrimination, but the 6-year-old children perform better than the 4-year-old children did in the tone identification of mid level-low level contrast (Ciocca & Lui, 2003; Sze, 2004). Since the mid level-low level contrast represented an average Fø difference of 7Hz in Ciocca and Lui’s study (2003), the 4-year-old children, who got an average FøDL of 11.3Hz for speech stimuli, would not be able to identify the mid level-low level contrast consistently. This was supported by Ciocca and Lui’s study (2003) which found that 4-year-old subjects identify this contrast with an accuracy rate just above chance level (50%). In contrast, the six-year-old children in Ciocca and Lui’s study (2003) were able to identify this contrast with about 80% accuracy. However, in the current study, 6-year-old children could only discriminate Fø difference of 10.8Hz in average. Also, the current study showed that children were not able to
discriminate F0 difference of 7Hz until after the age of 6. This discrepancy raised question about
the reliability of the result for the 6-year-old children in the current study. More precisely, results
from identification study (Ciocca & Lui, 2004) suggested that 6-year-old children could
discriminate F0 better than the level estimated in the current study. Therefore, the finding for
6-year-old children should be interpreted with caution.

Effect of type of stimuli

Another objective of the current study was to compare the performance of speech and
nonspeech F0 discrimination. The results showed that Cantonese speakers from all age groups
performed equally well in both speech and nonspeech F0 discrimination. Since the F0 level of
speech stimuli was perceived as linguistic material, while F0 level of nonspeech stimuli was
perceived nonlinguistically (Francis & Ciocca, 2003), the current finding suggested that the
linguistic processing of F0 could be done as well as the nonlinguistic processing. This result was
supported by Semal, Demany, Ueda, and Halle’s suggestion (1996) of a common store for the
pitch of both speech and nonspeech stimuli. The objective of Semal et al.’s study was to
investigate whether the pitch of speech sound was retained in the same way as the pitch of a
nonspeech sound, using frequency discrimination task. The tone pair to be discriminated
contained 4 interfering tones in between. The frequency difference between the tone pair to be
discriminated and the interference tones could be large, medium, or small. Listeners were
required to tell whether the first and the last tones were the same or different in pitch. The
authors proposed a counter hypothesis which stated that there was a specific store for retaining
the pitch of speech stimuli, and that this memory store could not be accessed by nonspeech
stimuli. If this hypothesis was true, F₀ discrimination of speech stimuli would be affected more when the interference tones were speech stimuli than when they were nonspeech complex tones. However, the results showed that for F₀ discrimination of speech stimuli, whether the interference tones were speech or nonspeech stimuli did not make a difference in the performance, no matter if the interference tones had large, medium, or small F₀ difference with the discrimination pair. Therefore, the counter hypothesis was proved to be incorrect, the pitch of speech stimuli was not retained in a specific store. Instead, speech and nonspeech stimuli shared the same store for pitch analysis.

The finding of similar performance in speech and nonspeech F₀ was not contradicted to Francis and Ciocca’s finding (2003). Although they found that Cantonese-speaking subjects were able to perform F₀ discrimination better for nonspeech than speech stimuli, they suggested the possibility that the difference could be attributed to differences in the complexity of the stimuli. As both the types of stimuli and the complexity of stimuli varied, it could be not concluded whether the difference in type of stimuli or in complexity of stimuli accounted for the difference in performance of F₀ discrimination of speech and nonspeech stimuli.

That listeners were able to discriminate F₀ of speech stimuli as well as nonspeech stimuli was not surprising since the perception of Cantonese level tones is noncategorical (Francis & Ciocca, 2003; Ng, 2000). For categorical perception, listeners discriminate nearly as many categories as they can perceive contrastively. They are able to tell the difference between stimuli when it lays across stimuli categories rather than within the same category (Kent, 1997). The contrast of categorical perception is continuous perception. In continuous perception, listeners
discriminate continuous variation in the parameter; they can discrimination a lot more stimuli than they can identify. In Francis and Ciocca’s study (2003), Cantonese-speaking subjects showed similar performance in F0 discrimination across the high level-mid level and mid level-low level boundaries than within the categories of high, mid, and low level tones. Besides, in the current study, the 10-year-old children were able to detect frequency difference of 5.2Hz for speech stimuli on average. That meant they were able to discriminate between speech stimuli with 104Hz and 109.2Hz. However, words with F0 of 104Hz and 109.2Hz would be likely to be identified as the low level tone (Francis & Ciocca, 2003). This again suggested that Cantonese-speaking listeners were able to discriminate tokens within the same category in Cantonese tone system. Taken these, it was possible that F0 discrimination of speech stimuli was done as well as F0 discrimination of nonspeech stimuli.

Conclusion

In the current study, F0DLs were estimated for 3-, 4-, 5-, 6-, 10-year-old Cantonese-speaking children and adults for speech and nonspeech complex tones. F0DLs decreased with age and adult’s level of performance was achieved at age 10. This finding, being in accord with previous studies on frequency discrimination, proved that children’s ability of discriminating F0 improved during the childhood. The current study showed that significant improvement in F0 discrimination occurred between the 3- and 5-year-old groups, between the 5-, 6-year-old groups and adult. Decrease in F0DLs from 3 to 5 explained the improvement of identification of the high level-mid level tone contrast between age 3 and 4. However, the 4-, 5-, and 6-year-old children performed similarly in F0 discrimination, while previous studies
suggested tone identification of mid level-low level contrast improved significantly between the age of 4 and 6. This discrepancy suggested that the 6-year-old children would perform better in Fø discrimination. Moreover, the nature of stimuli did not make a difference in Fø discrimination, suggesting that linguistic and nonlinguistic processing of Fø of complex tone could be done equally well. This finding could be explained by the concept of common store for pitch storage for nontone language speakers. Also, this finding was consistent with the noncategorical perception of Cantonese level tones.

Limitation and further study

In the current study, significant difference in FøDLs were found between 3- and 5-year-old groups; 3- and 10-year-old groups; and all groups except 10-year-old group and the adult group. It is suggested that using larger subject groups would increase the chance of getting more significant improvements in FøDLs between these age groups.

The current study provided developmental normative data on FøDLs of speech and nonspeech sounds in Cantonese-speaking children. It served as the basis for comparison between normal children and other population of interest. Similar study on children with specific language impairment can reveal information about how these children perceive Fø difference in speech and nonspeech stimuli, and how their performance different from that of normal children.
Acknowledgements

I am grateful to Dr. Valter Ciocca from the Department of Speech and Hearing Sciences of the University of Hong Kong for the advice and supervision on the dissertation. I would like to thank the staff and pupils of SKH The Church of our Saviour Kindergarten, and of the Tai Po Public School for their participation in this study. I thank Yiu Sham for designing the animation. Also, I wish to thank my family and friends for the endless concern and support.

References


Hong Kong.
