The perception of lexical tone contrasts in Cantonese children with and without Specific

Language Impairment (SLI)

Anita M.-Y. Wong

University of Hong Kong

Valter Ciocca

University of Hong Kong*

Sun Yung

University of Hong Kong

* currently affiliated with the University of British Columbia

Running Head: Tone perception in Cantonese SLI

Abstract

Purpose: This study examined the perception of fundamental frequency (f0) patterns by Cantonese children with and without Specific Language Impairment (SLI).

Method: Participants were 14 five-year-old children with SLI, 14 age-matched (AM) and 13 four-year-old vocabulary-matched (VM) controls. The children identified a word from familiar word pairs that illustrated the eight minimally contrastive pairs of the six lexical tones. They discriminated the f0 patterns within contrastive tonal pairs in speech and non-speech stimuli.

Results: In tone identification, the SLI group performed worse than the AM but not the VM group. In tone discrimination, the SLI group did worse than the AM group on two contrasts, and showed a non-significant trend of poorer performance on all contrasts combined. The VM group generally did worse than the AM group. There were no group differences in discrimination performance between speech and non-speech stimuli. No correlation was found between identification and discrimination performance. Only the normal controls showed a moderate correlation between vocabulary scores and performance in the two perception tasks.

Conclusion: The SLI group's poor tone identification cannot be accounted for by vocabulary knowledge alone. The group's tone discrimination performance suggests that some children with SLI have a deficit in f0 processing.

Specific Language Impairment (SLI) is a developmental disorder in which children show language skills that are significantly below their age peers, in the absence of neurological, cognitive, psychosocial, or hearing impairments (Leonard, 1998). The nature of SLI has been documented in children learning a range of different languages (Leonard, 2009), including Cantonese Chinese (e.g., Fletcher, Leonard, Stokes, & Wong, 2009; Fletcher, Stokes, Leonard, & Wong, 2005; Fletcher, Stokes, & Wong, 2006; Leonard, Wong, Deevy, Stokes, & Fletcher, 2006; Stokes & Fletcher, 2003; Wong, Leonard, Fletcher & Stokes, 2004;). There have been several explanatory accounts of SLI (e.g., Johnston, 2004; Leonard, 1992; Rice, Wexler, & Cleave, 1995; van der Lely, 1998), and one of the them involves auditory processing deficits.

Auditory processing deficits in children with SLI

Earlier studies have reported that children with SLI have difficulties processing rapidly presented auditory information (Tallal, 2000; Wright et al., 1997). More recent evidence, however, argue that, at least for some children, SLI is associated with a frequency discrimination deficit. In McArthur and Bishop (2004a), 16 individuals with SLI and 16 people with no known spoken language difficulties between the age of 10 and 20 years took part in four auditory backward recognition masking (ABRM), and a frequency discrimination (FD) tasks. The ABRM tasks were designed to test rapid auditory processing abilities. Group differences were not statistically significant on any of the ABRM thresholds. The SLI group, however, showed a significantly higher threshold on the FD task, indicating poorer performance. Of the 15 individuals with SLI who successfully completed the task, five (33%) had much higher FD thresholds than any individual in the control group. In a follow-up study, McArthur and Bishop (2004b)

examined FD performance in a group of thirty-two 12 to 21 year old individuals. Eleven of the participants with SLI and 13 of those in the age-control group had participated in the previous study by the same authors (McArthur & Bishop, 2004a), which had been conducted 18 months earlier. The later study included three FD tasks, in which children were asked to identify the pure tone with a higher frequency within a sequence of two (2I-2AFC), or three sinusoidal tones (3I-2AFC). The SLI group obtained higher FD thresholds than their typically-developing age control group across all three tasks, although none of these differences reached statistical significance after Bonferroni correction for multiple comparisons. When the thresholds for the three tasks were combined, the difference between the mean threshold of SLI group ($1.40 \log_{10}$ Hz) and of the control group ($1.07 \log_{10}$ Hz) just failed to reach statistical significance. There was a substantial individual variability within the SLI group, and some individuals did perform at similar levels as those in the control group. However, of the 16 individuals with SLI, five (31%) received a composite mean threshold that was above the 95% confidence intervals of the control group, clearly showing a deficit in FD. A comparison between the performance of the individuals with SLI on the 2I-2AFC FD task in the two studies by McArthur and Bishop showed that the performance by SLI group had been consistent over time. Further evidence that individuals with SLI have a deficit in the processing of frequency information was provided in subsequent studies which showed that the SLI group performed poorly in FD (Mengler, Hogben, Michie, & Bishop, 2005) but not in backward masking tasks (Bishop, Carlyon, Deeks, Bishop, 1999; Hill, Hogben, & Bishop, 2005).

Within the context of research on specific language impairment, investigations on auditory processing have primarily focused on (low-level) perceptual abilities to detect, discriminate and categorize pure tones on the basis of some acoustic features. How deficits in auditory processing relate to higher-level language learning and use, which children with SLI have persistent problems with, is still a matter of contention. McArthur and Bishop (2004b) suggested that FD deficits "affect(ed) the development of neural representations for the different phonemes in the language" (p. 529). Phonemes are the smallest meaningful units that make up the words in a language, and they are characterized by complex spectral properties which may also include f0 patterns (for voiced sounds). It is however unclear by which mechanism FD deficits might affect phoneme discrimination. It is also unclear whether the poorer FD threshold reported for English-speaking children with SLI would still be adequate for the accurate discrimination of the spectral properties of phonemes, and hence their eventual development of neural representations. Last but not least, there have been no precise proposals about the relationship between FD deficits and the specific patterns of language difficulties documented in English-speaking children with SLI (Leonard, 1998 for a review). For example, it is unknown how poor FD abilities can account for these children's difficulties in grammatical morphemes that mark tense and agreement.

Perception of lexical tones in Cantonese-speaking children

For speakers of tonal languages such as Chinese, the ability to discriminate fundamental frequency (f0) patterns is critical. In Chinese, each syllable contains voiced portions (vowels, diphthongs and voiced consonants) that are characterized by a specific f0 pattern that corresponds to the different lexical tones. Tone identity affects the

meaning of a word such that syllables that are composed of the same sequence of vowels and consonants may have very different meanings when produced with different tones. To illustrate, the Cantonese syllable /ji/ with a high level (55) tone refers to clothes, but when it carries a mid-level (33) tone, it means 'idea'. In addition to these two, there are four other contrastive tones in the Cantonese spoken in Hong Kong: low-level (22), high-rising (25), low-rising (23) and low-falling (21) (Bauer & Benedict, 1997). The numbers in parentheses represent, in relative terms, the beginning and the ending f0 values, of each tone. In Cantonese these f0 values generally fall into five categories, also called registers: High (5), mid-high (4), mid (3), mid-low (2) and low (1) (Chao, 1947). When the beginning and the ending f0 values do not change substantially, the tone has a flat contour (e.g., level tones 55, 33 and 22). When f0 falls from a high to a low register, it has a falling contour (e.g., low-falling tone, 21). A tone has a rising contour when f0 rises from a lower to a higher register (e.g., tones 25 and 23). While for some tonal languages such as Mandarin, acoustic features such as amplitude and duration can be used to identify lexical tones (Whalen & Xu, 1992), in Cantonese there is evidence that suggests that such cues are not reliable ones, and that f0 patterns are the only necessary and sufficient cue (Fok-Chan, 1974; Ching, 1984). Difficulties with f0 discrimination will affect the ability to correctly identify these tonal patterns, which provide critical information in a Cantonese-speaker's phonological representation of a word. Such a deficit, if confirmed, would have a direct impact on a Cantonese-speaking child's acquisition of language, and vocabulary acquisition in particular.

Most developmental studies on Cantonese tone perception reported on children's abilities to identify different lexical tones. In the first experimental study, Ching (1984)

asked children between four to 10 years of age to select one of the six pictures that matched with the monosyllabic word just presented (the syllable /ji/ produced with one of the six Cantonese lexical tones). Cantonese-speaking children were able to identify above chance level five out of the six lexical tones at 5 years of age; however, their performance did not reach the adult level until the age of 10. This finding was generally replicated by Ciocca & Lui (2003), who presented the target tones in the medial position within a semantically neutral carrier phrase. The six stimuli were the same syllable /ji/ produced with one of the six lexical tones. They used a two-alternative forced choice procedure in which the participants were asked to choose the picture that matched with the stimulus syllable they just heard. The two pictures illustrated one of the eight minimal pair tonal contrasts constructed from combinations of the six lexical tones. The tones within the High Level-Mid Level (HL-ML), High Level-Low Level (HL-LL), Mid Level-Low Level (ML-LL) contrasts differed in both starting and ending f0 values. Members of the Low Rising-Low Level (LR-LL), Low Falling-Low Level (LF-LL) and Low Falling-Low Rising (LF-LR) contrasts differed in ending f0, and in f0 contour. Tones within the High Level-High Rising (HL-HR) pair differed in starting f0 and contour. Finally, tones within the High Rising-Low Rising (HR-LR) contrast differed in ending f0 only. An overall improvement in identification accuracy was reported between the 4- and 6-year-old groups, and between the 6- to 10-year-old groups. As in Ching (1984), the 10-year-old group showed comparable performance as the adult group. When the children's performance for each of the eight tonal contrasts was examined, the ML-LL and HR-LR contrasts were found to be the most difficult for all age groups. Ciocca and Lui suggested that the similarity in f0 contour and the small f0 difference between the members of the

contrasts (over the whole duration of the tone for the ML-LL, and over the first portion of the tones for the HR-LR contrasts) were the likely reasons for the difficulty with these two contrasts.

Using the same research paradigm with two sets of stimulus words (familiar and unfamiliar words), A. W.-L. Sze (2004) replicated the findings reported in Ciocca and Lui (2003). In addition to the two contrastive pairs reported in Ciocca and Lui (2003), the LR-LL contrast was also found to be difficult for the younger children between 2;09 and 3;03. The HL-LL and HL-HR pairs were reported to be the easiest contrasts since even children between the age of 2;09 and 3;03 were able to achieve an identification accuracy between 80% and 90%. Comparisons of performance between the two sets of words revealed no word familiarity effect on tone perception, except for the tone contrasts HL-LL and HR-LR for the youngest age group. In agreement with previous findings (Ciocca & Lui, 2003; A. W.-L. Sze, 2004), Ip (2006) also reported that the ML-LL, HR-LR and LR-LL were the most difficult and the HL-LL and HL-HR were the easiest contrasts for children between 2;00 to 5;11. None of these studies, however, have examined to what extent tone identification is related to the children's language.

Developmental studies on f0 discrimination

Previous studies on children's frequency discrimination used pure tones, rather than complex tones, as stimuli (e.g., Maxon & Hochberg, 1982; McArthur & Bishop, 2004a, 2004b). However, most natural sounds, including speech sounds, are complex. Moreover, f0 discrimination of complex tones is important for speakers learning a tonal language like Cantonese. For these reasons, A. P.-Y. Sze (2006) examined whether Cantonese-speaking children demonstrated developmental changes in the f0

discrimination of complex tones. The complex tones, which included 10 harmonics (from 2nd to 11th), were first filtered using a formant filter centred at 700 Hz and a bandwidth of 200 Hz; they had a duration of 100 ms. Complex tones with f0 of 120 Hz were used as the standard tones. Comparison tones were presented at an f0 that was between 0.075 Hz and 76.8 Hz higher than the standard. The f0 difference limen of stimuli presented was measured in children aged 4, 5, 6 and 10 years and in adults. Thresholds were estimated using a 2-down 1-up procedure (Levitt, 1971). The threshold generally reduced with age, showing increased sensitivity in the FD of complex tones; performance reached the adult level at age 10. Because of the relatively large variability in the data, however, only the 4-year group had a significantly higher threshold than the 10-year old group. These results are consistent with studies examining English-speaking children's frequency discrimination of pure tones (Maxon & Hochberg, 1982; Thompson et al., 1999), suggesting that FD continues to develop into the early school years. We have, however, as yet no information on whether Cantonese-speaking children with SLI had more difficulties than their age peers on the FD of complex tones, as their English-speaking counterparts did with pure tones. What also remains unclear is how the f0 discrimination of complex tones develops with the mental representation of the different lexical tones in the language. These are some of the issues we addressed in this study.

Research questions

Unlike previous studies on FD by children with SLI, complex tones were employed as stimuli in this study. The f0 patterns of complex tones distinguish word meanings in Cantonese. Therefore, sensitivity in the identification or discrimination of the f0 patterns of complex tones can be related to young Cantonese-speaking children's

vocabulary development. Children with a higher sensitivity in tone perception are more likely to establish two words that differ in f0 as separate tokens in the mental lexicon with greater ease and in a shorter time than children with lower sensitivity. In this study, we compared the perception of lexical tone contrasts among Cantonese-speaking children with SLI, an age-matched (AM) and a younger vocabulary-matched (VM) group of typically developing children. The VM group was included to control for the plausible effect of vocabulary knowledge on the SLI group's performance. Children were asked to perform tone identification, speech and non-speech tone discrimination. The speech stimuli were minimal pair tone contrasts that had been used in previous developmental Cantonese tone perception studies (Ip, 2006; A. W.-L. Sze, 2004). The amplitude and duration characteristics of the tone stimuli within each tone contrast were identical; therefore, children only needed to attend to differences in the f0 patterns of the stimuli. The tone identification task required the children to perceive the pitch pattern in the syllable presented, and to associate this particular syllable with a lexical tone category by choosing one of two word labels. To succeed in this task, children needed to have a mental representation of the six tonal categories in Cantonese. These mental representations develop with the learning of new words and with the discrimination of their f0 patterns using general pitch perception processes. The tone discrimination tasks required the children to perceive the difference in the two f0 patterns presented with the same syllable, and to report whether they were the same or different. One of the discrimination tasks involved speech stimuli, which were the same real words used in the identification task. The other discrimination task employed non-speech stimuli that were digitally processed such that they had the same duration, amplitude envelope and f0

characteristics as the voiced portion of their speech counterparts. To perform accurately in these discrimination tasks, children could in principle perform the tasks by solely relying on general pitch perception processes (most likely to occur with non-speech stimuli) but could also activate their mental representation of the tonal categories (likely to occur in the speech task). These tasks had the purpose to investigate the development of tone identification and discrimination in typically-developing children and children with SLI. Finally, this study compared the pattern of performance across the eight tonal contrasts in the three groups of children to gain an insight on the nature of FD deficits in children with SLI. In summary, this study addressed these questions:

1. Do Cantonese-speaking children with SLI identify and discriminate lexical tones at the same overall level, and in a similar way across the eight tonal contrasts, as typically-developing children?

2. Do Cantonese-speaking children with SLI show a similar pattern of performance as typically-developing children in the discrimination of tones for speech and non-speech stimuli?

3. Is performance in tone identification and tone discrimination related to vocabulary skills in children with SLI and in the typically developing controls?

<div align="center">Method</div>

Participants

Forty-one children between 4 and 6 years of age participated in this study. They were all attending preschool and among them, 14 were children with SLI. These 14 children were referred to us as language delayed by their speech-language pathologists, and they were subsequently confirmed to meet the conventional criteria for SLI. Each of

these children scored more than 1 SD below the mean for their age on the Receptive

(comprehension) Scale of the Cantonese version of the Reynell Developmental Language

Scales (RDLS, Hong Kong Society for Child Health and Development, 1987). In the

Columbia Mental Maturity Scale (CMMS; Burgemeister, Blum, & Lorge, 1972), a test of

nonverbal intelligence, these children scored no lower than 1 SD of their age mean,

indicating normal cognitive development. They all passed an oral motor screening

adapted from Robbins and Klee (1987) and showed no signs of psychosocial or

neurological problems. They also passed a pure tone hearing screening conducted at 30

dB HL (if seen in our research laboratory), or 35 dB HL (if seen in their child care centres)

at the frequencies of .25, .5, 1, 2, 4, 6 KHz. The passing criterion was a contingent

response to each tone on two out of the three times it was presented to each ear. Scores on

the Expressive Scale of the RDLS and the mean length of their utterances in a

conversational sample were also obtained from each participant. The former was not used

a criterion for inclusion in the group, since some children with SLI scored within the

normal range despite their clinical status. The latter, calculated using the same procedures

reported in Klee, Stokes, Wong, Fletcher and Gaven (2004) provided a complementary

measure of their expressive language. For later matching with a younger group of

typically-developing children, each child also completed the Cantonese Receptive

Vocabulary Test (CRVT; C. Lee, L. Lee, & Cheung, 1996). A summary of the children's

ages, RDLS, CMMS, CRVT scores, and MLUs is presented in Table 1.

Insert Table 1 about here

Of the 27 participants with typical language skills, 14 children had ages that were within 3 months of that of a child in the SLI group (typically-developing age-matched, AM group). All children in this group scored no less than .67 SD below the mean on the RDLS Receptive and Expressive Scales, and received significantly higher scores than the SLI group (t (26) = 6.27 p < .000) on the Receptive Scale. The AM group performed significantly better than the SLI group on the Expressive Scale (t (26) = 3.82, p < .001), MLU (t (26) = 3.11, p = .004), as well as on the CRVT (t (26) = 4.56, p = .000). All AM children scored no less than 1 SD below, **and no** more than 1.25 above the mean on the CMMS, and their scores were significantly higher than those for the SLI group (t (26) = 3.26, p < .003). All children in the AM group passed the hearing and oral motor screenings.

The remaining 13 typically developing children, as a group, were significantly younger than both the children in the SLI group (t (25) = 4.22, p < .000), and the children in the AM group (t (25) = 4.64, p < .000). Each of these 13 children was matched within three points on the CRVT with a child in the SLI group. Hereafter, these children will be referred to as the typically developing vocabulary-matched (VM) group. In order to ensure that the children in this group showed only average vocabulary knowledge, their scores on the CRVT were no less than 1.0 SD below, and no more than 1.0 SD above, their age mean. All VM children scored no less than .67 SD below the mean on the RDLS-Receptive and Expressive Scales. The group scored significantly better than the SLI group on the Receptive Scale (t (25) = 3.99, p < .000), but showed comparable performance as the SLI group on the Expressive Scale (t (25) = 1.98, p > .05). The MLU in words for children in the VM group were not collected. Everyone scored no less than

1.0 SD below their age mean on the CMMS, and all passed the hearing and oral motor screenings.

A group of 10 adults with no known hearing, speech or language disorders also participated in the two discrimination tasks to provide reference performance data from mature Cantonese-speakers. The five men and five women were between 19 and 50 years of age, and none of them had received professional or specialist training on these tasks.

Stimuli

The speech stimuli for the identification task were the monosyllabic words that were used by A. W.-L. Sze (2004). They were colloquial Cantonese terms (see Appendix A) that could be presented in picture format, and should be present in the vocabulary of normally developing children at age three (Fletcher, Leung, Stokes & Weizman, 2000). The speech stimuli were recorded in a sound attenuated room and stored in the memory of a Macintosh Power Mac 7100 computer. A male native Cantonese speaker produced the stimuli five times in random order (five blocks of stimuli produced in random order) within the carrier phrase: /ŋɔ23 wui33 tuk2 X pei35 nei23 tʰɛŋ55/ (I will read X for you to listen). The distance between the speaker's mouth and the microphone was kept constant at 10 cm so as to ensure the recording level was similar for all stimuli. Three native Cantonese listeners with four years of phonetic training were employed to listen to all the sentence stimuli. The best of the five utterances of the carrier with the /bei33/ target was selected by the three expert listeners, on a consensus basis, as the carrier sentence that was used for presenting all target words in the identification task. For each of the eight tonal contrasts, one of the five utterances of one of the two contrast members was selected as the "base" stimulus for that contrast. For example, one of the five

utterances of the word /lou22/ was selected as the base stimulus for the LL-LR contrast (/lou22/ vs. /lou23/). This selection was made such that the "base" i) had f0 values that were close to the average of the five utterances for that word, ii) was judged as an accurate production by the listeners, and iii) could be easily spliced out of its original carrier sentence for further processing. The amplitude of the selected base stimulus within its original carrier was normalized (to obtain a similar level across target words), before the base was extracted from its original carrier and spliced into the common carrier. The two target words for each contrast were obtained by re-synthesizing the base such that their f0 values corresponded to the average f0 patterns for the specific tones. By using the same stimulus to re-synthesize both members of each tonal contrast, we ensured that the target words differed only in their f0 patterns while duration and intensity characteristics were identical. All the sound editing and digital signal processing was carried out using the Praat software 4.3.1 (Boersma & Weenink, 2005); and the PSOLA procedure (Moulines & Charpentier, 1990), implemented in the Praat software, was employed for the re-synthesis of the target words with modified f0 values. After all the stimuli were thus re-synthesized, three new native Cantonese listeners with four years of phonetic training were asked to listen to the stimuli to ensure that they were good utterances of each word. The duration of the target words varied from 300 ms (/pou/) to 470 ms (/hai/). Table 2 gives the mean f0 values of the tone pairs for the eight tone contrasts at the beginning, middle and end-point of each word. Figure 1 shows the details of the f0 patterns for the members of each tonal contrast. These patterns are typical for tones produced by male speakers.

Insert Figure 1 about here

For the discrimination tasks, the speech stimuli were the same target words used in the identification task. The words were manually edited from the carrier phrases, their amplitudes were normalized, and they were presented in pairs with a 500-ms inter-stimulus interval (ISI). The non-speech stimuli were re-synthesized from the speech stimuli by extracting the f0 values and re-synthesizing the stimuli as a hum with the same f0 characteristics as the original stimuli -using the 'pulses-pitch' (hum) option in the Praat software. This procedure synthesizes signals on the basis of the pitch pulse information extracted by the software, using a schwa-like formant structure. The humming sounds thus resynthesized from the original words were low-pass filtered at 1900 Hz, and then a pre-emphasis filter was applied twice, in order to give a non-speech-like quality to the stimuli. The amplitude contour of the original speech stimuli was extracted and then multiplied with the re-synthesized and filtered sound, so that non-speech analogs of each word had similar amplitude and duration characteristics as the voiced portions of the corresponding speech stimuli. Finally, the amplitude of all stimuli was normalized in order to equate the presentation level of the stimuli. This processing was carried out for each of the target words using the Praat software. Figure 2 shows the amplitude waveform and f0 tracks of the speech and non-speech analogs for the LF-LR contrast. This figure shows that the duration and the amplitude envelope were identical for members of the tonal contrasts.

Insert Figure 2 about here

Procedures

The identification task tested children's ability to identify a Cantonese word from familiar word pairs that were minimally contrastive in lexical tones (e.g., /pou55/, "pot" vs /pou33/, "cloth"). The same eight tone contrasts used in previous developmental studies (Ciocca & Lui, 2003; A. W.-L. Sze, 2004) were used as stimuli. For each contrast, children were tested using two pairs of word stimuli originally developed by A. W.-L. Sze (2004) (Appendix A). The procedure for this task was the same as that used in earlier studies on the perception of lexical tones in young Cantonese-speaking children conducted in our lab (Ciocca & Lui, 2003; A. W.-L. Sze, 2004; Ip, 2006). The child first listened to a word (one of the two members of a tonal contrast) that was embedded in the middle of a carrier phrase (i.e., "ŋɔ23 wui23 tuk2 ____ pei25 nei23 tʰɛŋ55") and then

was asked to point to one of two pictures that was displayed on a computer screen. The two pictures represented the objects corresponding to the words for the relevant lexical contrast. Each member of a tonal contrast was presented eight times (four blocks of trials). Each picture was presented once on the left, and once on the right on the computer monitor in randomized order within each block. There were a total of 128 trials (four blocks of 32 trials each), using custom software written by the second author; the software controlled stimulus presentation and data collection on a Macintosh G4 iBook. All stimuli were presented bilaterally through Sennheiser HD280 Professional headphones at a comfortable listening level.

The discrimination tasks (speech and non-speech stimuli) tested children's ability to discriminate pairs of familiar words (or their non-speech analogs) that were minimally contrastive in their f0 patterns. The words were manually extracted from the sentences used in the identification task. The non-speech stimuli were complex sounds that did not contain any Cantonese phonemes, and they were identical to the words used in the speech discrimination tasks in their f0 characteristics. Within each trial, pairs of stimuli were presented with an inter-stimulus interval of 500 ms. After listening to a stimulus pair, the child was asked to indicate whether the two stimuli were the same or different. Responses were given either verbally, or by putting a check for "same" and a cross for "different" on a response grid. The response grid helped to keep the child on task as it gave visual feedback to the child on the number of trials completed or remaining. All but the first two older children gave their responses using the response grid. There were a total of 160 trials; half of the trials consisted of non-identical word pairs. The 16 non-identical word pairs were exactly the same stimuli used for the eight contrasts in the identification task; for example, the different pairs /pou55-pou33/ and /pou33-pou55/, and the corresponding "same" pairs were used to test the discrimination of stimuli with the same f0 values as the stimuli of the HL-ML contrast. The remaining 80 trials consisted of identical word pairs, which were constructed from words used in the eight non-identical word pairs. Examples for these identical pairs are /pou55-pou55/ and /pou33-pou33/. There were five blocks of trials; word pairs were presented in randomized order within each block.

Each child received 6-8 training trials before each of the experimental tasks. In the training trials for the identification task, the child listened to the stimuli and then was told which picture it corresponded to. In the training for the discrimination tasks, children

were told whether the two stimuli just presented were the same or different, and they were shown the corresponding correct response on the response grid. Most children volunteered a response during training, and were given corrective feedback if necessary. The tasks began when the children showed comprehension of the tasks. Given that there were predictable effects of one task on the others, the order of the tasks was counterbalanced by assigning each child to one of the six possible orders of the three tasks. Given a nearly equal number of children in each group, there was basically no difference in the number of times a particular order was presented across the groups. Each task took between 20 to 30 minutes to complete. Each child was seen individually and typically completed all three tasks on two to three separate days.

<div align="center">Results</div>

Tone identification

Children's accuracy in tone identification was measured as proportions of correct responses. Figure 3 gives the mean identification scores for each of the three groups across the eight contrasts. To compare performance across the groups and contrasts, we ran a two-way ANOVA and to follow up on main and interaction effects, we conducted post-hoc Tukey tests for unequal sample size.

<div align="center">------------------------------------</div>

<div align="center">Insert Figure 3 about here</div>

<div align="center">------------------------------------</div>

The three groups of children showed significant differences in their identification accuracy ($F_{(2,38)} = 6.21$, $p < .005$). The SLI group (mean = 67.6%, SD = 6.9%) was significantly less accurate ($p < .05$) than the AM group (mean = 79%, SD = 11.6%, $d =$

1.16). Cohen's *d* is used here as a measure of effect size, with values of .20, .50 and .80

suggesting small, medium and large effects (Cohen, 1988). Further evidence in support

for the group difference is the finding that among the 14 children with SLI, 11 scored

below 73% (the 95% confidence limit of the AM group mean). The SLI group's

performance, however, was not significantly different ($p > .05$) from that of the VM

group (mean = 66.3%, SD = 12%). Like the SLI group, the VM group was less accurate

than the AM group ($p < .01$, $d = 1.08$).

All groups combined, the children were more successful in tone identification for

some contrasts than others, F (7, 266) = 32.90, $p < .0001$. They received a significantly

higher score for the HL-LL and HL-HR contrasts than the other six contrasts, and for the

HL-ML contrast than the ML-LL and HR-LR, LR-LL and LF-LL contrasts ($p < .05$).

The groups did not, however, show the same pattern of performance across the eight

contrasts, as revealed by a significant "group by contrast" interaction effect, F (14, 266) =

1.97, $p < .05$. The SLI group, but not the AM group, performed better for the HL-LL and

HL-HR than for the HL-ML, LF-LR and LR-LL contrasts ($p < .05$). The VM group' s

performance pattern was similar, albeit at a lower overall level, than that of the AM group;

however, unlike the AM group, the VM group performed better for the LR-LF than the

ML-LL, and for the HL-LL than the LF-LL and LR-LL contrasts ($p < .05$). When

performance was compared directly among groups, the SLI group was significantly less

accurate than the AM group in the identification of the HL-ML ($d = 1.13$), LR-LL ($d =$

1.12), and LF-LL ($d = 1.2$) contrasts. No other group differences were statistically

significant.

To examine whether there was a relationship between vocabulary and tone identification, Pearson correlation coefficients were calculated on the children's CRVT and identification scores. When scores from all children were included, r was equal to .49, and when only children in the AM and the VM groups were considered, r was equal to .52. These r values were statistically significant ($p < .05$, one tail). No statistically significant relationship between receptive vocabulary and tone identification was reported for the SLI group (r = .24, $p > .05$ one tail).

Tone discrimination

The children's ability to discriminate tones were examined using the measure *d'*. This measure reports an individual's sensitivity to discriminate differences between two stimuli, and it is calculated on the basis of hit and false alarm rates (see, for example, MacMillan, & Creelman, 2005). Children received a *d'* above zero when their hit rates were higher than their false alarm rates, and a zero when the former was equal to the latter, and a negative *d'* score when their false alarm rates were higher than their hit rates. A *d'* score of 1 denotes a moderate sensitivity to differences between stimuli. In the current study, *d'* was calculated by using the tables for same-different designs (differencing model) provided by MacMillan and Creelman (2005). The maximum value for *d'* was 6.93, when the hit rate was 1 and the false alarm rate was 0.

The means and standard deviations of *d'* scores across groups and contrasts for speech and non-speech stimuli are shown in Figure 4a-b. The SLI group received a *d'* score that was smaller than 1 for three speech contrasts, HR-LR, LR-LL and LF-LL, and for one non-speech contrast, HR-LR. The VM group also scored less than 1 for the ML-LL contrast with speech stimuli, and for the same set of contrasts as reported for the SLI

group (HR-LR, LR-LL, LF-LL) and for the LF-LR contrast for both speech and non-speech stimuli. In fact, the VM group's *d'* score for the speech HR-LR and LF-LR contrasts fell slightly below zero. An examination of individual scores in the VM group revealed that all but two of the 13 children received between two to five negative *d'* scores for the speech stimuli, and all but three of the children received between one to five negative d' scores for the non-speech stimuli. The number of children who received negative *d'* scores was only between 3 to 4 in the AM group, and 7 in the SLI group. Negative scores indicated that the child gave more "different" responses when the stimuli were actually identical than when the stimuli were different. These results suggest that a majority of the children in the VM group failed to make a distinction between the stimuli within many of the minimal contrasts. To compare the children's sensitivity in tone discrimination across groups (3), type of stimuli (2, speech vs. non-speech) and tonal contrasts (8), we ran a three-way ANOVA with group as a between-subject factor, and the type of stimuli and tonal contrasts as within-subjects factors; post-hoc Tukey tests for unequal sample size were conducted to further investigate main and interaction effects.

Insert Figure 4a-b about here

The three groups showed significant differences in their sensitivity in tone discrimination, $F(2, 38) = 6.84$, $p < .005$. The VM group (mean = .82, SD = 1.81) performed significantly worse than the AM group (mean = 3.22, SD = 1.59) ($p < .005$, *d* = 1.41). The SLI group (mean = 1.82, SD = 1.70) received a higher *d'* score than the VM

group, but this difference was not statistically significant ($p > .05$). Among the 14 children with SLI, eight scored outside the 95% confidence interval of the AM group mean both with speech (2.41), and non-speech (2.29) stimuli; one additional child with SLI scored outside the 95% confidence limit of the AM group with speech stimuli only. As a group, the children with SLI received a lower d' score than the AM group, but given the large within group variability, the difference was not statistically significant ($p > .05$).

As was the case in the identification task, performance varied across tonal contrasts, and the main effect of contrast was statistically significant, $F (7, 266) = 18.61$, $p < .000$. Similar to the identification task, the HL-ML, HL-LL, and HL-HR contrasts were the easiest to discriminate overall. The "group by contrast" interaction was significant, $F (14, 266) = 1.96$, $p < .01$. For both the AM and the SLI groups, the HL-ML and HL-LL contrasts were significantly better discriminated than the HR-LR contrast ($p < .05$); children from these two groups were also more sensitive to the f0 differences for the HL-HR than for the ML-LL, HR-LR, LF-LL and LF-LR contrasts ($p < .05$). Children in the AM group, but not children with SLI, had also lower sensitivity for the HR-LR than the LR-LL and LF-LR contrasts. Only the SLI group performed worse for the LR-LL than for the HL-HR contrast. Overall, however, the pattern of d' across contrasts of the SLI group was similar to that of the AM group. By contrast, the VM group's pattern of sensitivity was virtually flat across the tonal contrasts and no pair-wise comparison between contrasts was statistically significant for the VM group.

When performance of the three groups was directly compared for each of the eight contrastive pairs, the SLI group was found to be less sensitive than the AM group for two contrasts ($p < .05$), including the LR-LL ($d = 1.40$), and the LF-LR ($d = .77$)

contrasts. Compared to the VM group, the SLI group performed significantly better for one of the contrasts (HL-HR, $p < .05$; $d = .73$). The AM group was more sensitive than the VM group for all but one contrast (HR-LR), with effect size ranging from .90 to 1.96.

With the second research question we wanted to determine whether Cantonese-speaking children with and without SLI perform in a similar way in the discrimination of f0 differences for speech and non-speech stimuli. The three groups did not differ in discrimination sensitivity with speech and non-speech stimuli, as indicated by the absence of a significant "group by type of stimuli" interaction effect, $F(2, 38) = .14$, $p > .05$. In fact, when the children's scores were combined across groups and contrasts, their performance with speech stimuli (mean = 1.98, SD = 2.01) was similar to that with non-speech stimuli (mean = 1.99, SD = 1.97), resulting in a non-significant main effect on nature of stimuli, $F(1, 38) = .01$, $p > .05$. There was a significant "contrast by type of stimuli" effect ($F(7, 266) = 3.03$, $p = .004$), but post-hoc analysis indicated no significant differences between the discrimination of speech and non-speech stimuli for any one of the eight contrastive pairs ($p > .05$). The three-way interaction effect was not statistically significant, $F(14, 266) = 1.00$, $p > .05$.

While the three groups of children received a mean d' scores between .82 and 3.22, the adult group was almost at the ceiling level of performance ($d' = 6.93$) for both speech (mean = 6.40, SD = 0.52) and non-speech stimuli (mean = 6.27, SD = 0.71). Of the eight contrasts, the HR-LR contrast was the most difficult; it was the only one for which a significant difference was found between non-speech (mean = 4.59, SD = 1.47) and speech stimuli (mean = 5.54, SD = 1.53), $t(9) = 2.58$, $p < .05$, $d = 0.63$.

To examine whether there was a relationship between vocabulary and tone discrimination, Pearson correlation coefficients were calculated on the children's CRVT and d' scores. With speech stimuli, when scores from all children were included, the r was equal to .44, and when only children in the AM and the VM groups were considered, the r was equal to .58. With non-speech stimuli, when scores from all children were included, the r was equal to .46, and when only children in the AM and the VM groups were considered, the r was equal to .60. These r values indicated a statistically significant moderate relationship ($p < .05$, one tail). No statistically significant relationship between receptive vocabulary and tone discrimination for either speech (r = .15) or non-speech stimuli (r = .21) ($p > .05$, one tail) was reported for the SLI group.

To summarize, the SLI group performed significantly worse than the AM group in tone identification, but performed at a similar level as the VM group. In tone discrimination, the SLI group performed significantly worse than the AM group for two of the eight contrasts, and showed a non-significant trend of poorer performance for all contrasts combined. A majority of the children in the VM group failed to make a distinction of the f0 differences between many of the contrasts. Both the SLI and the AM group showed a stronger discrimination sensitivity than the VM group, although a significant difference was only reported between the AM and the VM groups. The SLI group however did score significantly better than the VM group on the easiest HL-HR contrast. In both tasks, the HL-LL and HL-HR contrasts were perceived most accurately while the HR-LR and ML-LL were the most difficult to perceive. While the overall pattern of results for children with SLI was similar to that of their typically-developing AM peers, differences in both overall performance level and in specific tonal contrasts

were also found between the two groups. There was a moderate but significant relationship between vocabulary scores and performance on the identification and discrimination tasks for the typically-developing children.

Discussion

Development of tone perception in typically-developing children

When averaged across the eight contrasts, the overall identification accuracy of the 5-year-old AM group (79%) in this study was similar to that of the 5;00 to 5;11 group in Ip (2006), who scored 82%. By contrast, the performance the 4-year-old VM group in this study was between 9 to 11% lower than the means reported for children at the same age. In fact, the VM group's mean identification accuracy was more similar to that of the 3;00 to 3;03 group in A. W.-L. Sze's study (2004), who scored 68%, and that of the 3;0 to 3;11 group in Ip (2006), who scored 67%. There are several reasons for differences in these findings. In this and other earlier studies, the sample size was small with only between 14 to 16 children in each age group. Sampling error might have an effect particularly on findings for children younger than four-years of age, because of large individual differences. Moreover, both A. W.-L. Sze (2004) and Ip (2006) used shorter testing periods, since they only included the identification task. In Ip's study children also received a smaller number of trials for each contrast than children in this study. For example, only two of the 15 4-year-old children completed all 128 trials in Ip's (2006) identification task; the remaining 13 children completed only 64 trials. In this study, every participating child completed all 128 trials. Structured tasks that require the child to stay attentive for an extended period of time might be a challenge for some children, especially young preschoolers.

26

The children as a whole, including those with SLI, did better in tone identification for some tone contrasts than others. This finding replicates earlier reports (Ciocca & Lui, 2003; A. W.-L. Sze, 2004; Ip 2006) that Cantonese-speaking children did the best for the HL-LL and HL-HR contrasts, and the worse with the ML-LL, HR-LR and LR-LL contrasts. A clear developmental difference in mean identification accuracy was observed between the two groups of typically-developing children. The 5-year-old AM group was generally more accurate in tone identification than the 4-year-old VM group.

This study reports for the first time the discrimination of contrastive pairs of complex tones in Cantonese speakers. The adults in this study were very accurate in this task and performed almost at ceiling for all contrasts, except for the most difficult HR-LR contrast (See Figure 4). For this particular contrast, they performed better with speech than non-speech stimuli, suggesting that adults may have been able to use tone categorization processes to improve their discrimination ability for this tonal contrast. While the children identified the ML-LL contrast at a significantly lower accuracy than the HL-ML and HL-LL contrasts, they showed no significant difference in their discrimination sensitivity among these contrasts. The f0 difference between members of the ML-LL pairs (about 15-17 Hz on average) is larger than the f0 discrimination threshold obtained with similar stimuli (9.5 Hz for 4-year old children; A. P.-Y. Sze, 2006). Therefore, members of this contrast should be relatively easy to discriminate by most children as young as four, and this was generally true for the children in this study. However, the identification of these tones requires children to have a mental representation of separate tonal categories that differ by relatively small f0 differences, and to match the f0 pattern of the speech input with these representations. The difference

in performance between identification and discrimination for level-tone contrasts

suggests that the discrimination task was carried out mainly by relying on (f0-based)

pitch perception processes.

Clear developmental differences were observed in the AM and the VM group's

overall sensitivity in tone discrimination. The AM group received a higher mean *d'* score

than the VM group. The 5-year-old AM group also showed a clearly different pattern of

performance across the tonal contrasts, showing early development of discrimination

sensitivity for some contrasts before the others. A majority of the children in the 4-year-

old VM group failed to make a frequency distinction between many of the contrasts.

However, as a group their *d'* score was larger than one on the three contrasts that were

found to be the easiest in the AM group (HL-ML, HL-LL and HL-HR) for both speech

and non-speech stimuli. These findings suggest that the same pattern of discrimination

sensitivity to the different contrasts emerged at 4 years of age if not earlier. Among the

four or five contrasts in which the VM group failed to make a distinction, the group's

performance in the identification task was at chance for only two (ML-LL and HR-LR)

(Binomial test; $n = 208$, $p = 1/2$, alpha = .05). This suggests that identification

performance was not totally dependent on discrimination ability. In fact, if it were, then

one would expect that identification and discrimination scores were highly correlated.

However, Figure 5 shows that there is no clear relationship between identification and

discrimination performance for children in the AM and VM groups. In the AM group, all

but one child showed some discrimination sensitivity but their scores were spread over a

wide range. For all but one of these children, their identification scores were however,

uniformly and more narrowly spread between .7 and 1. For the VM group, most children

showed poor sensitivity, but variable identification scores, except for three children

whose *d'* scores were above 1. For these children, there seems to be a positive correlation

between identification and discrimination, although it is difficult to draw firm

conclusions given the small sample size. These findings support the conclusion that the

identification of lexical tones requires the ability to match speech percepts with mental

representations of tonal categories in children's lexicon, and that these representations are

learned over time during childhood. Once children are able to discriminate the f0 patterns

corresponding to the various tonal categories, their identification performance is

generally accurate.

Insert Figure 5 about here

Deficits in tone perception in children with SLI

The SLI group was less accurate than the AM group in tone identification. The

SLI group's performance in the tone discrimination tasks suggests that they had a deficit

in the processing of f0 patterns. When their performance was compared with that of

children in the AM group, the children with SLI showed discrimination deficits for two of

the eight tonal contrasts, indicating a deficit in f0 processing. Evidence from the

comparison of the SLI group and the younger VM group shows that, despite their age

advantage, the SLI group performed better than the VM group on only one of the

developmentally easiest contrasts (HL-HR).

Specific patterns of performance of the SLI group relative to the AM group might

further characterize their deficit in f0 processing. Evidence from earlier tone

identification studies suggests that contrasts whose members have similar contour but differ on the basis of small f0 differences (ML-LL and HR-LR) are difficult for typically-developing children; in fact, contrasts marked by small f0 differences are difficult to perceive even though tones differ in contour (LR-LL). The SLI group performed more poorly than the AM group on the LR-LL contrast in both the identification and the discrimination tasks. However, the SLI group's deficit in f0 processing did not seem to be restricted to contrasts with small f0 differences, as they performed more poorly than the AM group on the LF-LR contrast as well. Such group differences cannot be easily explained by differential familiarity with certain tones, as an examination of words used in daily conversations did not reveal that some lexical tones or tone pairs are used more frequently than others (Fok Chan, 1974). Nor could it be explained by differential knowledge in the stimuli since only familiar words were used in the present study. A more likely explanation for the performance of children with SLI in the present study is that they have deficits in the processing of f0 patterns. These deficits result in imprecise mental representations of tonal categories, which in turn impair their abilities to specify tonal information for words in their mental lexicon. Such an impairment affects vocabulary, as well as grammar, since both lexical and grammatical forms are marked by tones in Cantonese.

Given that lexical tones in Cantonese words help mark lexical meanings, it is possible to hypothesize that poor f0 processing abilities leads to the inadequate development of mental representations of tonal categories by children with SLI, resulting in poor tone identification performance and in poor vocabulary knowledge. In fact, the SLI group had significantly poorer receptive vocabulary scores than the AM group. The

fact that there was no significant difference between the SLI group and the VM group of younger typically-developing children with comparable vocabulary knowledge also seems to support this argument. Poor vocabulary knowledge alone, cannot explain the SLI group's poor tone perception performance. In fact, from the zero-order correlations reported earlier, it could be estimated that vocabulary knowledge accounted for only about 24% of the variance in the typically developing children's performance in tone identification, and for about 19% and 21% of the variance in their performance in the tone discrimination involving speech and non-speech stimuli respectively. The percentages were even lower for the SLI group given the lack of significant correlations. These observations provide further evidence for the role of f0 processing deficits on the impaired identification and discrimination of lexical tones by children with SLI. Given that lexical tones distinguish word meanings in Cantonese, such deficits will have a direct and plausibly a cascade effect on language development in children with SLI.

Deficits in fundamental frequency processing in children with SLI

As reported in frequency discrimination studies with English-speaking children with SLI (Hill, et al., 2005; McArthur & Bishop, 2004a; 2004b; Mengler et al., 2005), there were individual differences in the perception of complex lexical tones in our group of Cantonese-speaking children with SLI. Some of the children in the SLI group actually scored within the 95% confidence level of the mean of the AM group. Three of the fourteen children with SLI were as accurate as their age peers in the identification of tonal contrasts. Six children with SLI were as accurate as their age peers in the discrimination of both speech and non-speech stimuli, and one child with SLI was as accurate their age peers only in the discrimination of non-speech stimuli. Having said that,

there was not a single child with SLI who did not have problems with either identification or discrimination of lexical tones: seven children had deficits in both tasks, five were impaired in identification and the remaining two in discrimination. These results show the presence of subgroups in Cantonese-speaking children with SLI. Some of the children with SLI did not perform at an age-appropriate level in the processing of the f0 patterns of complex tones, and some had difficulties with the labeling and categorization of the tonal patterns in the language. Another subgroup of children with SLI seems to have difficulties in both F0 processing and in the categorization of f0 patterns (see Figure 5).

The patterns of performance in the two tasks shows that although the ability to discriminate f0 patterns might facilitate the accurate categorization of lexical tones, good sensitivity does not necessarily lead to accurate identification for these children. For example, five SLI children who had a *d'* score of 3 or higher still performed at a relatively low identification accuracy (70% or lower; see Figure 5). This finding suggests that some children with SLI have mainly categorization deficits. The other subgroup of children with SLI who had lower than age-appropriate sensitivity did not show a clear relationship between discrimination and identification scores, suggesting that f0 processing deficits are often, but not always, associated with categorization deficits.

Future directions

The current findings complement previous reports of a pure tone frequency discrimination (FD) deficit in children with SLI who are learning English (Hill, et al., 2005; McArthur & Bishop, 2004a; 2004b, Mengler et al., 2005). The perceptual deficits reported in the current study resulted in impairments in the identification and the discrimination of f0 patterns associated with Cantonese lexical tones. Therefore, our

findings suggest that the perceptual deficits in children with SLI are not limited to deficits in frequency discrimination of pure tones. It is not yet known to what extent deficits in the processing of f0 patterns can be accounted for by FD deficits. Future research on Cantonese-speaking children with SLI should examine their FD performance in order to further understand the nature of their perceptual deficits.

It is possible that deficits in the processing of the frequency of both pure and complex tones result from a general impairment in the processing of spectral information (McArthur & Bishop, 2005). If present, such an impairment might affect the processing of other spectrally-based attributes of sounds, such as timbre. Therefore, it will also be important to investigate the processing of other perceptual attributes that rely on the processing of spectral features of sounds -such as timbre- by children with SLI, relative to typically-developing age peers.

Future research on pitch perception in children with SLI should also collect detailed information on the children's general language skills, such as receptive and expressive vocabulary (McArthur & Bishop, 2001), as well as general cognitive processing abilities, and examine how these abilities might contribute to the interpretation of the children's performance. Successful completion of perception tasks also requires adequate comprehension, multiple cognitive abilities, including attention, motivation and memory. Among these skills, children with SLI were reported to have problems with working memory (Ellis Weismer, Evans, & Hesketh, 1999; Montogmery, 2000), attention (Lum, Conti-Ramsden, & Lindell, 2007), and a domain-general information processing deficit (Im-Bolter, Johnson, & Pascual-Leone, 2006). It is plausible that some of these cognitive processing abilities account for the SLI group's performance on tone perception,

in addition to their deficits in vocabulary, categorization and f0 processing. While it is not clear how FD deficits relate to the language difficulties observed in English-speaking children with SLI, the relationship between tone perception and word learning is more direct for Cantonese-speaking children, given that tonal information is critical in the phonological representation of all Cantonese words. Future research should examine whether SLI is associated with defective cognitive and auditory processing abilities that cause a failure to meet the processing demands of Cantonese.

References

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter.

Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research, 42*, 1295-1310.

Boersma, P., & Weenink, D. (2005). Praat: A system for doing phonetics by computer. Retrieved on February 15, 2005, from University of Amsterdam, Institute of Phonetics Sciences web site: *http://www.praat.org*.

Burgemeister, B., Blum, L., & Lorge, I. (1972). *The Columbia Mental Maturity Scale*. New York: Harcourt Brace Jovanovich.

Ching, Y. C. (1984). Lexical tone pattern learning in Cantonese children. *Language Learning and Communication, 3*, 317-334.

Chao, Y.-R. (1947). *A Cantonese Primer*. Cambridge, MA: Harvard University Press.

Ciocca, V., & Lui, J. Y.-K. (2003). The development of the perception of Cantonese lexical tones. *Journal of Multilingual Communication Disorders, 1*, 141-147.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences (2nd ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ellis Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1249-1260.

Fletcher, P., Leonard, L. B., Stokes, S. F., & Wong, A. M.-Y. (2009). Morphosyntactic deficits in Cantonese-speaking children with Specific Language Impairment. In S.-P. Law, B. Weekes & A. M.-Y. Wong (Eds.), *Language disorders in Chinese-speaking children and adults* (pp. 75-88). Bristol: Multilingual Matters.

Fletcher, P., Leung, S. C.-S., Stokes, S., & Weizman, Z. (2000). *Cantonese pre-school language development: A guide*. Hong Kong: The University of Hong Kong.

Fletcher, P., Stokes, S. F., Leonard, L. B., & Wong, A. M.-Y. (2005). The expression of aspect in Cantonese-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*, 621-634.

Fletcher, P., Stokes, S. F., & Wong, A. M.-Y. (2006). Specific language impairment in Chinese. In P. Li, L. H. Tan, E. Bates & O. J. L. Tzeng (Eds.), *Handbook of Asian Psycholinguistics, Vol. 1*. (pp. 296-307). Cambridge: Cambridge University Press.

Fok Chan, Y. Y. (1974). *A perceptual study of tones in Cantonese*. Hong Kong: Hong Kong University Press.

Hill, P. R., Hogben, J. H., & Bishop, D. V. M. (2005). Auditory frequency discrimination in children with specific language impairment: A longitudinal study. *Journal of Speech, Language, and Hearing Research, 48*, 1136-1146.

Hong Kong Society for Child Health and Development (1987). *Manual of the Reynell Developmental Language Scales, Cantonese (Hong Kong) version*. Hong Kong: Author.

Im-Bolter, N., Johnson, J. R., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*, 1822-1841.

Ip, V. W.-K. (2006). *Perception and production of lexical tones by young children.* University of Hong Kong, Hong Kong.

Johnston, J. R. (2004). Fearless and bold in following ideas wherever they lead, *The Elizabeth Bates Memorial Lecture at the 25th Symposium on Research in Child Language Disorders*. Madison, WI.

Klee, T., Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Gavin, W. (2004). Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 47*, 1396-1410.

Lee, K. Y.-S., Lee, L. W.-T., & Cheung, P. S.-P. (1996). *Hong Kong Cantonese Receptive Vocabulary Test*. Hong Kong: Hong Kong Society for Child Health and Development.

Leonard, L. B. (1992). The use of morphology by children with specific language impairment: Evidence from three languages. In R. Chapman (Ed.), *Processes in language acquisition and disorders* (pp. 186-201). St Louis: Mosby-Yearbook.

Leonard, L. B. (1998). *Children with Specific Language impairment*. Cambridge, MA.: The MIT Press.

Leonard, L. B. (2009). Cross-linguistic studies of child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 308-324). New York: Psychology Press.

Leonard, L. B., Wong, A. M.-Y., Deevy, P., Stokes, S. F., & Fletcher, P. (2006). The Production of Passives by Children with Specific Language Impairment Acquiring English or Cantonese. *Applied Psycholinguistics, 27*, 267-299.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 73*, 663-667.

Lum, J. A. G., Conti-Ramsden, G., & Lindell, A. K. (2007). The attentional blink reveals sluggish attentional shifting in adolescents with specific language impairment. *Brain and Cognition, 63*, 287-295.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum Associates. Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behaviour: Sensitivity and discrimination. *Ear and Hearing, 3*, 301-308.

McArthur, G. M., & Bishop, D. V. M. (2001). Auditory perceptual processing in people with reading and oral language impairments: Current issues and recommendations. *Dyslexia, 7*, 150-170.

McArthur, G. M., & Bishop, D. V. M. (2004a). Which people with specific language impairment have auditory processing deficits? *Cognitive Neuropsychology, 21*, 79-94.

McArthur, G. M., & Bishop, D. V. M. (2004b). Frequency discrimination deficits in people with specific language impairment: Reliability, validity, and linguistic correlates. *Journal of Speech, Language, and Hearing Research, 47*, 527-541.

McArthur, G. M., & Bishop, D. V. M. (2005). Speech and non-speech processing in people with specific language impairment: A behavioural and electrophysiological study. *Brain and Language, 94*, 260-273.

Mengler, E. D., Hogben, J. H., Michie, P. T., & Bishop, D. V. M. (2005). Poor frequency

discrimination is related to oral language disorder in children: A psychoacoustic

study. *Dyslexia, 11*, 155-173.

Montogmery, J. W. (2000). Verbal working memory and sentence comprehension in

children with specific language impairment. *Journal of Speech, Language and

Hearing Research, 43*, 293-308.

Moulines, E. & Charpentier, F. (1990). Pitch-synchronous waveform processing

techniques for text-to-speech synthesis using diphones. *Speech Communication, 9*,

453-467.

Reynell, J. and Huntley, M. (1985) *Reynell Developmental Language Scales: Revised*.

Windsor: NFER-Nelson.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific Language Impairment as a

period of extended optional infinitive. *Journal of Speech and Hearing Research,

38*, 850-853.

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development

in young children. *Journal of Speech and Hearing Disorders, 52*, 271-277.

Stokes, S. F., & Fletcher, P. (2003). Aspectual forms in Cantonese children with Specific

Language Impairment. *Linguistics, 41*, 381-406.

Sze, A. P.-Y. (2006). *Discrimination of fundamental frequency by children: A

comparison of two alternative procedures*. Unpublished B.Sc. Honours Thesis,

Division of Speech and Hearing Sciences, University of Hong Kong, Hong Kong.

Sze, A. W.-L. (2004). *The development of tone perception in Cantonese-speaking children.* Unpublished B.Sc. Honours Thesis, Division of Speech and Hearing Sciences, University of Hong Kong, Hong Kong-SAR-China.

Tallal, P. (2000). Experimental studies of language learning impairments: From research to remediation. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairment in children: Causes, characteristics, intervention and outcome* (pp. 131-155). Hove, U.K. : Psychology Press.

Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *Journal of Speech, Language, and Hearing Research, 42*, 1061-1068.

van der Lely, H. (1998). SLI in children: Movement, economy, and deficits in the computational-syntactic system. *Language Acquisition, 7*, 161-192.

Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments, *Phonetica, 49*, 25–47.

Wong, A. M.-Y., Leonard, L. B., Fletcher, P., & Stokes, S. F. (2004). Questions without movement: A study of Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language and Hearing Research, 47*, 1440-1453.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387*, 176-178.

Acknowledgements

Table 1. The age, RDLS-R, RDLS-E, CMMS, CRVT scores and MLU (words) for the SLI, age-matched (AM) and vocabulary-matched (VM) groups.

		age	RDLS-R	RDLS-E	CMMS	CRVT	MLU
AM	mean	64.86	59.21	60.93	111.29	60.79	4.02
	SD	(7.82)	(3.51)	(3.36)	(6.35)	(3.89)	(0.43)
	range	55-77	50-64	56-69	101-120	52-65	3.39 - 4.80
SLI	mean	64.00	47.14	52.64	103.86	52.86	3.35
	SD	(8.30)	(6.29)	(7.38)	(5.68)	(5.22)	(0.68)
	range	54-76	35-54	40-66	94-114	44-63	2.23 - 4.53
VM	mean	51.00	55.38	57.15	112.92	54.46	---
	SD	(7.67)	(4.15)	(3.72)	(5.88)	(4.18)	
	range	43-68	49-65	52-65	99-120	48-61	

RDLS-R: the Cantonese version of the Reynell Developmental Language Scale (RDLS, Hong Kong Society for Child Health and Development, 1987)-Receptive subtest

RDLS-E: the Cantonese version of the Reynell Developmental Language Scale (RDLS, Hong Kong Society for Child Health and Development, 1987)-Expressive subtest

CMMS: Columbia Mental Maturity Scale (Burgemeister, Blum, & Lorge, 1972)

CRVT: Cantonese Receptive Vocabulary Test (K. Lee, L. Lee, & Cheung, 1996)

MLU: mean length of utterance in words

Table 2. Fundamental frequency values for members of each pair of tonal contrasts, measured at the beginning, middle, and end-point of the voiced portion of each word (duration shown in parentheses). The end-point value corresponds to the last voice cycle (level and falling tones) or the peak value towards the end of the utterance (rising tones).

Tonal contrasts (duration in ms)	Tones	Starting f0 (Hz)	Middle f0 (Hz)	Ending-peak f0 (Hz)
HL-ML (270 ms)	HL tone	139.5	139.6	133.1
	ML tone	114.5	115.4	107.1
HL-LL (250 ms)	HL tone	148.3	137.8	132.3
	LL tone	110.6	104	101.1
ML-LL (264 ms)	ML tone	116.9	116.2	114.5
	LL tone	104.4	101.9	97.0
HR-LR (282 ms)	HR tone	104.1	102.7	134.2
	LR tone	103.7	98.6	112.8
LR-LL (333 ms)	LR tone	106.3	102.3	113.2
	LL tone	107.6	104.7	98
HL-HR (330 ms)	HL tone	140.2	137.1	128.2
	HR tone	111.5	101	138.1
LF-LL (343 ms)	LF tone	108.7	91.2	86.1
	LL tone	100.9	100.9	95.7
LF-LR (290 ms)	LF tone	106.6	94.9	88.6
	LR tone	107.3	98.3	112.3

Figure 1. Fundamental frequency (f0) patterns of the voiced portions of the speech stimuli used for each of the eight tonal contrasts employed in this study. Stimulus duration differed among contrasts; the duration of the time window is indicated (in seconds) in the lower left portion of each panel for each contrast.
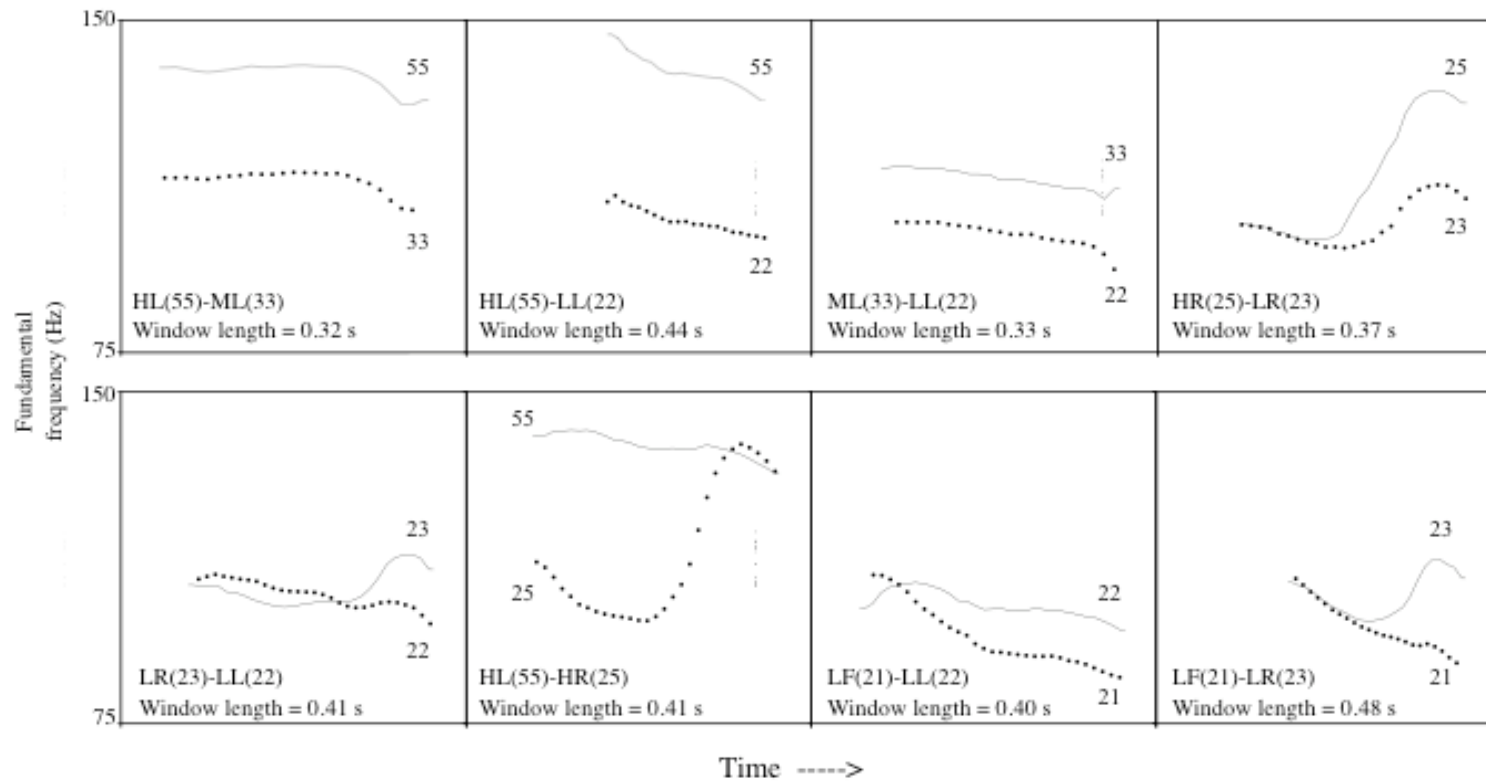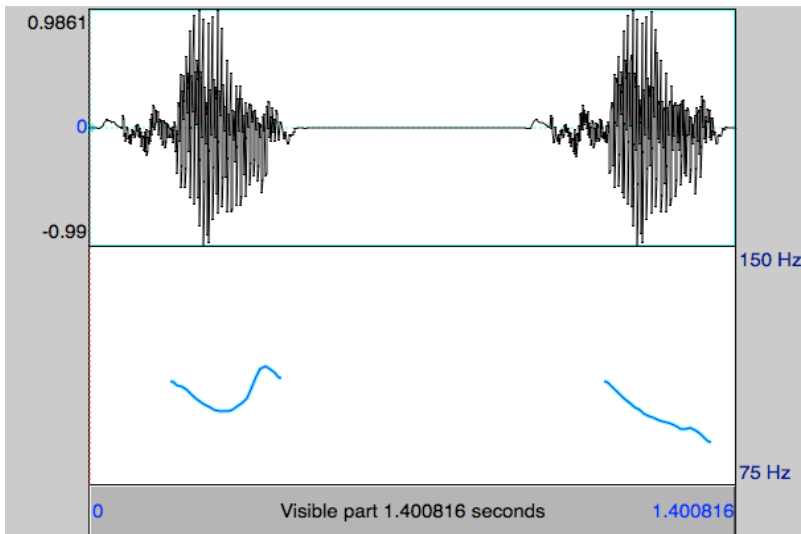
Figure 2. Amplitude waveform and f0 tracks of the speech stimuli (/hai23/-/hai21/; part a)
and of the corresponding nonspeech stimuli (part b) for the LF-LR contrast. Amplitude
waveforms are displayed in the top and the f0 track at the bottom of each panel.

(a)

(b)
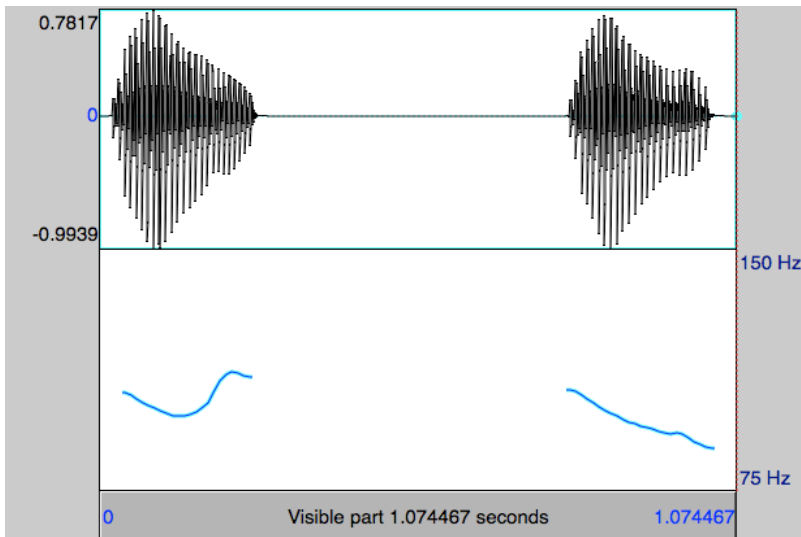
Figure 3. Mean identification scores (% correct) by age-matched (AM), SLI and

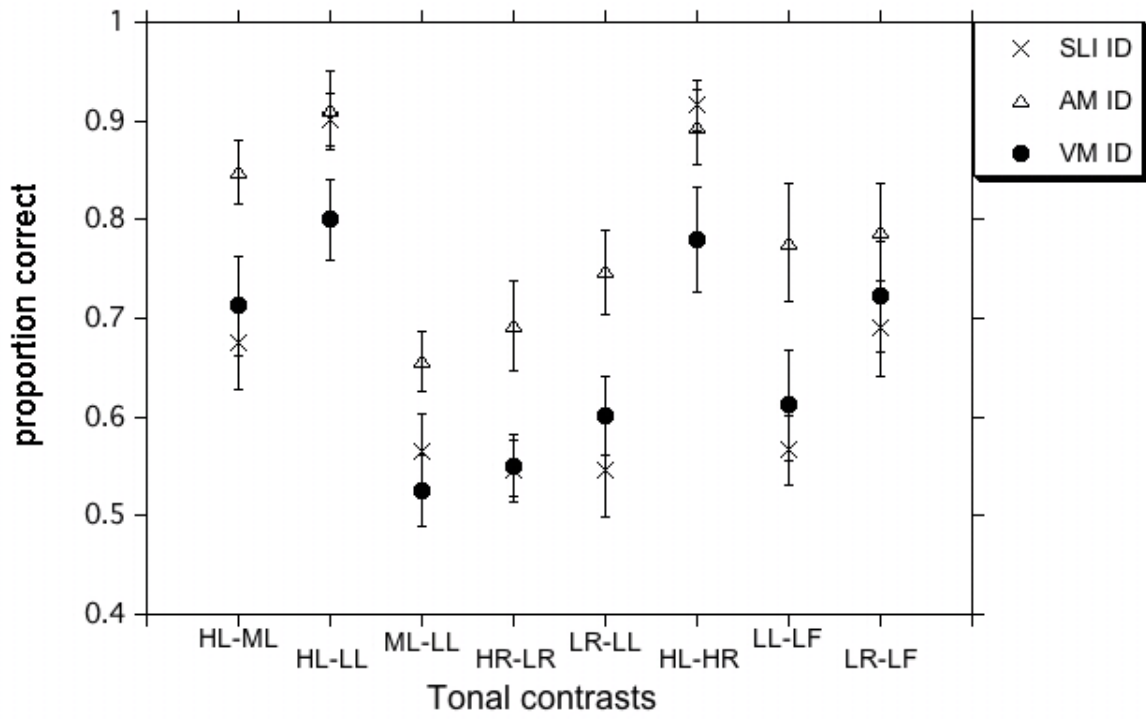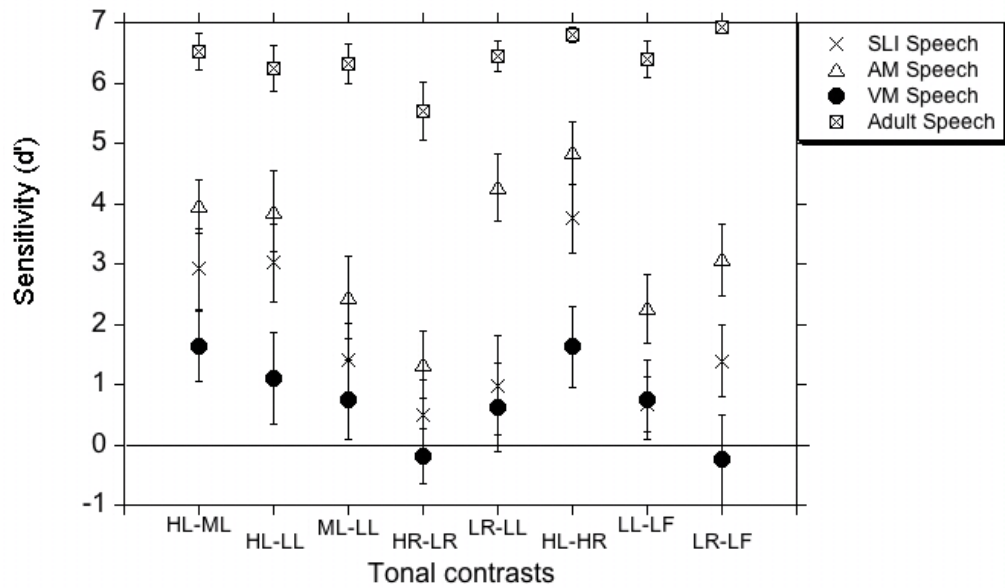vocabulary-matched (VM) children for each tonal contrast.

Figure 4. Mean sensitivity (*d'*) scores for the discrimination task with speech stimuli (a) and for nonspeech stimuli (b), with standard error bars. Adult data for the same contrasts are also displayed for comparison purposes.
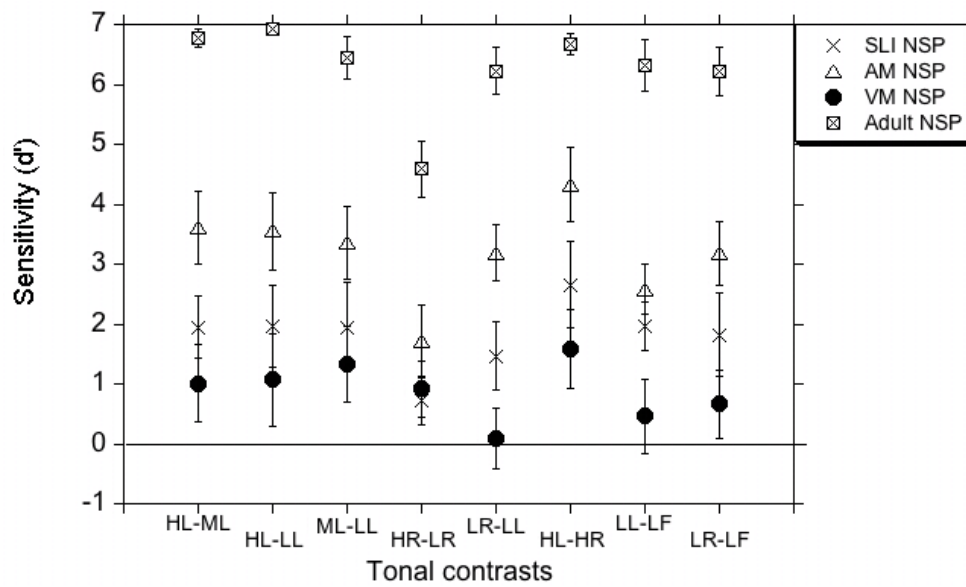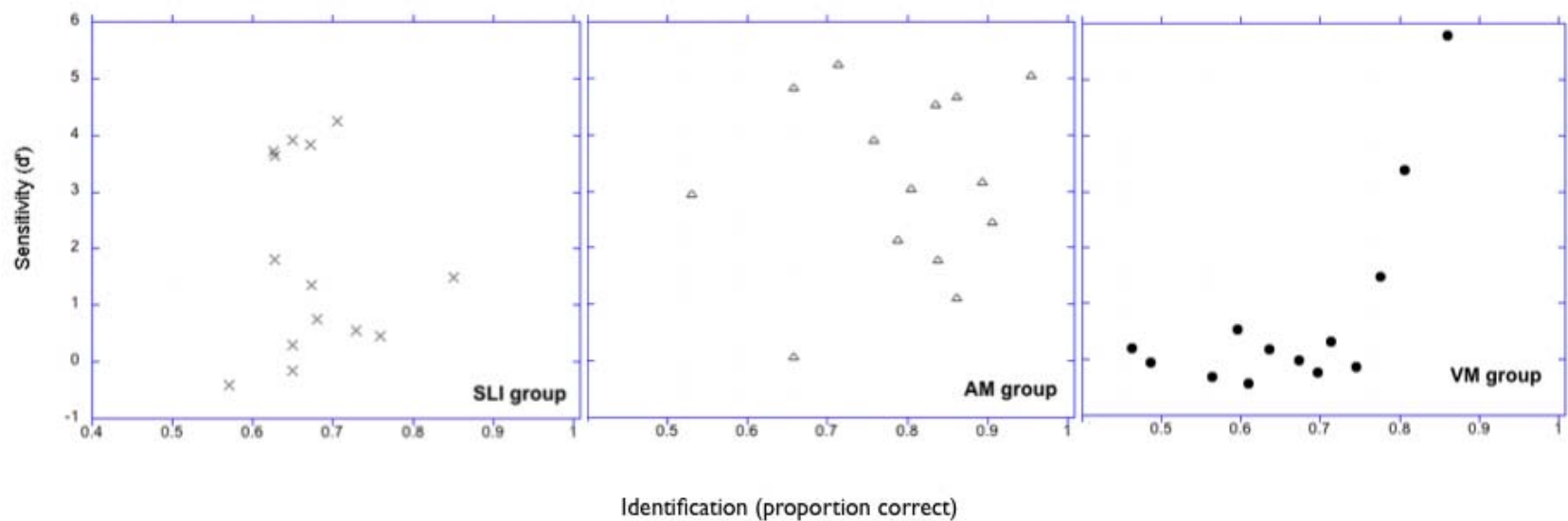
(a)

(b)

Figure 5. Scatterplots showing sensitivity scores (calculated as means of speech and non-speech discrimination scores, averaged across all contrasts; x-axis) plotted against identification scores (y-axis), for each participant from the SLI (a), AM (b) and VM (c) groups.

Appendix A. A sample of the contrastive pairs used in the tone identification and

tone discrimination (speech stimuli) tasks.

/pou55/, "Pot" - / pou33/, "Cloth" (HL–ML contrast);

/sy55/, "Book" - /sy22/, "Tree" (HL-LL contrast);

/pei33/, "Arm" - /pei22/, "Nose" (ML-LL contrast);

/jy25/, "Fish" - /jy23/, "Rain" (HR-LR contrast);

/hai23/, "Crab" - /hai21/, "Shoes" (LR-LF contrast);

/tʰɔŋ55/, "Soup" - /tʰɔŋ25/, "Sweet (candy)" (HL-HR contrast);

/lou22/, "Road" - /lou23/, "Old (man)" (LL-LR contrast);

/min21/, "Cotton" - /min22/, "Noodle" (LF–LL contrast).

For /lou22/ and /tʰɔŋ25/, the words within quotes indicate the literal meaning of the corresponding Chinese characters; the words in parentheses describe the corresponding picture for those characters whose meaning was not associated with a single specific image.