CAN CANTONESE CHILDREN WITH COCHLEAR IMPLANTS PERCEIVE LEXICAL TONES?

V. Ciocca, R. Aisha, A. Francis, and L. Wong

Department of Speech & Hearing Sciences, The University of Hong Kong, Hong Kong, Hong Kong SAR, China

ABSTRACT
The purpose of this study was to determine whether prelingually deaf children fitted with a cochlear implant are able to identify Cantonese tones in monosyllabic words. Eight children participated in this study. The age of the children ranged between 4:06 and 8:11. The duration of the post-operative period varied between 11 and 41 months. All children were fitted with the Nucleus 24-channel cochlear implant system. The stimuli were the six contrastive Cantonese tones produced with the monosyllabic target word /ji/. Each stimulus represented a concrete object: for example, the target word uttered with the high-level tone (/ji55/) means "clothing", while the same word produced with the low-level tone (/ji22/) means "two". The six target words were produced by a male native speaker of Cantonese within a carrier phrase. Listeners had to identify the target words, presented within the carrier phrase, by selecting one of two pictures. These pictures represented one of eight minimal pair contrasts: high- vs mid-level; high- vs low-level; mid- vs low-level; high-level vs high-rising; high-rising vs low-rising; low-rising vs low-level; low-falling vs low-rising; low-falling vs low-level. Within each trial, one picture represented the object corresponding to the word/tone that had been presented; the other picture represented the other word of a minimal pair. Subjects were tested individually in a soundproof room. Each contrast was presented sixteen times; each member of the minimal pairs was presented eight times for each contrast. The order of the stimuli was randomised. No feedback was given. The results showed that performance was at chance level (50%) for all contrasts: the percentage of correct responses for each contrast varied between 47 and 58%. An analysis of the data grouped by tones, rather than tone contrasts, revealed that none of the tones was identified above chance. These findings suggest that this group of cochlear implant users could not extract sufficiently accurate pitch information to identify Cantonese tones. The implications of this finding for the perception of pitch information through cochlear implants will be discussed.

1. INTRODUCTION
Profoundly deaf children with little residual hearing who gain no or little benefit from the use of conventional acoustic hearing aid have been able in recent years to use cochlear implants. Cochlear implantation may provide access to audition by electrically stimulating the functioning fibers of the auditory nerve in the cochlea to induce a sensation of hearing [1].

Even though profoundly deaf individuals can make effective use of lipreading for speech comprehension, the effective use of auditory information is necessary for the perception of the lexical tones of tonal languages such as Cantonese and Mandarin. Since lexical tones are produced by physiological changes of the larynx—which correspond to the acoustic features of fundamental frequency (F0), amplitude and duration- it is unlikely that reliable visual cues exist for the perception of lexical tones. Unlike Mandarin tones, for which not only F0 changes but also amplitude patterns and duration have been found to be reliable acoustic cues to tone perception (see, e.g., [2], [3], [4]), Cantonese tones have been found to be cued almost exclusively by F0 pattern [5], [6]. Therefore, Cantonese tones are suitable stimuli for testing the capacity of cochlear implant listeners to estimate the fundamental frequency of phonation for the purpose of perceiving the pitch patterns of speech sounds. The current study investigate the perception of Cantonese lexical tones by prelingually deaf children using cochlear implants.

2. METHOD

2.1. Subjects
Eight subjects aged between four to nine year-old were employed in this study. Only children older than the age of four years were included in this study as the study done by [7] showed that normal children are unable to reliably recognize isolated lexical tones until age four. All of these subjects were prelingually deaf children and native speaker of Cantonese. All subjects used the Nucleus 24-Channel cochlear implant system and the SPEAK processing strategy. There were four females and four males. Their postoperative period ranged from eleven to 41 months. According to a study done by [8], prelingually deaf children using cochlear implant require at least 12 months experience to have performance level of above 50% accuracy in stress recognition task and close-set word recognition test. All subjects were fitted with cochlear implants either at the Queen Elizabeth, the Prince of Wales or the Queen Mary Hospital in Hong Kong, where they also received auditory and speech training.

2.2. Stimuli
Natural stimuli were used because children respond best to natural speech tokens in lexical tone identification tasks [7]. The word /ji/ is chosen as the target word as it can be represented by simple and concrete lexical labels: Tone 55 /ji/, (clothing), Tone 25 /ji/, (chair), Tone 33 /ji/, (spaghetti), Tone 21 /ji/, (child), Tone 23 /ji/, (ear) and Tone 22 /ji/, (two). The stimuli were produced by a native Cantonese male speaker aged 21. The six words were presented in random order to the speaker, who read the words aloud in a sound proof room. Each word was produced ten times within the carrier phrase: /ŋɔ wai ʨɛk ɛi liɛ tɛŋ/ (I will read ___ for you to hear). The carrier phrase continued the target words in medial position to prevent the rise or drop of intonation that might affect the perceived tone values [9]. The minimum and maximum excursion of fundamental frequency (F0) for each word were measured; these values were used to calculate the mean and range of variation for each tone. The words with the smallest differences from their respective computed F0 average were used to make a ‘most average’ sentence. This sentence was designated as the context sentence for presenting the target words to the listeners. The productions of each lexical tone that represented the most extreme F0 difference, in the appropriate direction, from the mean F0 values of the corresponding tone were used as the stimuli for the identification task. For example, the ‘best’ /ji55/ (high level) syllable was the syllable that had the greatest positive F0 difference in F0 from the mean /ji55/ values. The best six productions for each tone were clipped out of their respective sentence, normalized in amplitude and inserted into the context sentence to make the stimuli sentences.

The six stimuli were grouped into the following eight tonal contrasts:

i) High Level vs Mid Level (‘HL_ML‘; Tone 55 vs 33)
ii) High Level vs Low Level (‘HL_LL‘; Tone 55 vs 22)
iii) Mid Level vs Low Level (‘ML_LL‘; Tone 33 vs 22)
iv) High Rising vs Low Rising (‘HR_LR‘; Tone 25 vs 23)
v) Low Rising vs Low Level (‘LR_LL‘; Tone 23 vs 22)
vii) Low Falling vs Low Rising (‘LF_LR‘; Tone 21 vs 23)
vii) Low Falling vs Low Level (‘LF_LL‘; Tone 21 vs 22)

Contrasts HL_ML, HL_LL and ML_LL were used to investigate the effect of separation between the three pitch levels (high, mid and low) on tone perception. If listeners were sensitive to F0 separation between level tones, then performance should be more accurate for pair HL_LL which contains larger F0 separation than contrasts HL_ML, and ML_LL. Pairs HR_LR, LR_LL, LF_LR and LF_LL were used to test listeners’ sensitivity to F0 differences in the end-point of tones, since the tones for these pairs start at similar frequencies but end at different frequencies. Tones in pair HR_LR have the same (rising) contour while tones for pairs LR_LL, LF_LR and LF_LL have different contours. Finally, pair HL_NR contains tones which have a similar F0 end-point but different initial F0.

Within a block of trials, each contrast was presented four times. For each tonal contrast, each member of the contrast was presented twice as the auditory stimulus (target word). Six visual stimuli corresponding the objects for each of the six tones were used for performing the identification task. The visual stimuli were color pictures of the objects corresponding the members of each tonal contrast. The place of presentation (left-right) of the visual stimuli for each contrast was counterbalanced.

2.3. Procedure

Subjects were tested individually in a double-walled IAC soundproof room. One experimenter sat inside the soundproof room behind the listener, while the other experimenter and the caregiver sat outside the soundproof room. A computer (Pentium by Macintosh 7100/80AV) was placed outside the soundproof room. A Hypercard 2.4 program was used for the presentation of the visual and auditory stimuli. Visual stimuli were projected from the computer onto a screen placed in the soundproof room using a CTX EzPro 500 projector. The order of presentation of the stimuli was randomized by the computer. The auditory stimuli were output through an Audiomedia II D/A board into a Madsen 0B822 audiometer, and then were presented through a Westra LAB-501 loudspeaker. Subjects sat one meter away from the speaker in the soundproof room.

Before testing began, all the subjects were familiarized with all the lexical items and the corresponding visual stimuli (pictures) used for labeling, to make sure that the children knew all the six words. Each trial began with the presentation of the two pictures corresponding to a tonal contrast. After presentation of the pictures, the target word was presented within the carrier phrase. The subjects were asked to choose the picture that matched the target word. They were given the following instructions: ‘You will hear each word once, then you should point to one of the two pictures to tell me which word you have heard’. The subjects responded by pointing to one of the two pictures on the screen; the experimenter who was sitting outside the sound proof room stored the selected response into the computer. The subjects were encouraged to guess if they were not sure about the correct response. Before the experimental session, all the subjects were given 10-15 practice trials. These trials were identical to the experimental trials except that the listeners were given feedback on their responses. No feedback was given to the subject on their performance during the experimental session. Subjects were allowed to take a short break whenever they requested it.

The level of the stimuli inside the soundproof room was 60dB SPL as measured with a sound-level meter (Brul & Kjær Type 1625). Each subject had to complete four blocks of trials; each block consisted of 32 tonal contrasts.

This study was carried out in conjunction with another study that involved tone discrimination and tone identification tasks. Subjects were tested in the order of (a) I-D-I or (b) I-D-I, where I was the tone identification task of this study while D and I were the tone discrimination task and the tone identification task of the other study. Four subjects were tested for each of the two orders of the tasks.
3. RESULTS

3.1. Performance for Tonal contrasts

The average correct scores for each tone contrast for the cochlear implant subjects range from 48% to 58% (see Table 1).

<table>
<thead>
<tr>
<th>Tonal Contrasts</th>
<th>% correct cochlear implant Ss</th>
<th>% correct hearing impaired Ss</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL-ML</td>
<td>58</td>
<td>92</td>
</tr>
<tr>
<td>HL-LL</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>ML-LL</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>HR-LR</td>
<td>47</td>
<td>58</td>
</tr>
<tr>
<td>LL-LR</td>
<td>51</td>
<td>92</td>
</tr>
<tr>
<td>HL-HR</td>
<td>55</td>
<td>92</td>
</tr>
<tr>
<td>LL-LF</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>LR-LF</td>
<td>52</td>
<td>93</td>
</tr>
</tbody>
</table>

Table 1: This table shows the percentage of correct responses for each tonal contrast by the cochlear implant listeners and by the hearing impaired subject.

The cochlear implant subjects did not perform above chance for any of the tonal contrasts by a binomial test (N= 128, p = 1/2, α = 0.05). Only one cochlear implant subject (S7) performed above chance (binomial test, N = 16, p = 1/2, α = 0.05), but only in one condition (contrast HL-ML). A one-way ANOVA with repeated measures was carried out on the mean percent correct identification for each subject and each contrast. The results of ANOVA showed that the null hypothesis that the means of tonal contrasts were equal could not be rejected. F (7,49) = 1.21, p > 0.05. Table 1 shows that the moderately hearing impaired subject performed much better than the cochlear implant subjects for all contrasts except for the contrast between High Rising and Low Rising tones. His performance on this contrast (58% correct) was at chance (binomial test, N = 16, p = 1/2, α = 0.05).

3.1. Performance for tones

The performance for each of the six tones was calculated by dividing the number of times each tone was identified correctly by the total number of times that each tone appeared as the target word in tonal contrasts. These scores were used to calculate the percentage correct responses for each tone. The average correct scores for each tone pair range from 42% to 57% for the cochlear implant subjects (see Table 2).

<table>
<thead>
<tr>
<th>Tonal Contrasts</th>
<th>% correct cochlear implant Ss</th>
<th>% correct hearing impaired Ss</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Level</td>
<td>54</td>
<td>89</td>
</tr>
<tr>
<td>Mid Level</td>
<td>57</td>
<td>96</td>
</tr>
<tr>
<td>Low level</td>
<td>42</td>
<td>100</td>
</tr>
<tr>
<td>High Rising</td>
<td>52</td>
<td>100</td>
</tr>
<tr>
<td>Low Rising</td>
<td>56</td>
<td>72</td>
</tr>
<tr>
<td>Low Falling</td>
<td>56</td>
<td>92</td>
</tr>
</tbody>
</table>

Table 1: This table shows the percentage of correct responses for each tone by the cochlear implant listeners and by the hearing impaired subject.

Performance by the cochlear implant subjects did not exceed chance for any of the tones (binomial test, N= 128-192, p = 1/2, α = 0.05). One cochlear implant subject (S7) scored well for the Mid Level and the Low Falling tones (88% and 75% correct, respectively). However, none of the other subjects exceeded 70% correct performance for any of the tones. A one-way ANOVA with repeated measures was carried out on the mean percent correct identification for each subject and each tone. The results of the ANOVA showed that the main effect was not significant, F (7,49) = 1.21, p > 0.05. The performance of the moderately hearing impaired child who participated in the pilot study was well above the performance of the cochlear implant subjects. His worse score was obtained for the Low Rising tone (72%).

4. DISCUSSION

The results for the tonal contrasts and the individual tones show that cochlear implant patients could not perform above chance in a tone identification task in which they had to choose between two minimal pair alternatives. The current results have been obtained from a limited sample of children, and are limited to one cochlear implant device (Nucleus 24-channel) and one processing strategy (SPEAK). Nonetheless, these preliminary findings suggest that cochlear implant children who are prelingually deaf have great difficulty in extracting F0 information on the basis of the input provided by cochlear implants. This difficulty affects the subjects’ capacity of perceiving Cantonese lexical tones accurately.
There are several factors that could affect performance in the present study, such as the etiology of deafness, the age of the child, the age of implant fitting, the postoperative period, the type of cochlear implant, the frequency and quality of speech and auditory training, motivation and learning ability of the child, and the amount of practice at home. For this study, all of the children were prelingually deaf and sixteen of the subjects were using the Nucleus 24 cochlear implant system. All the subjects have received or were receiving at least one hour of speech and auditory training by speech therapists in their first two years of implantation in the hospital where they had the implantation. Preliminary analyses suggest that neither the age of the subjects, the age of implant fitting, nor the duration of the postoperative period are correlated with the performance of the subjects.

It is interesting to compare the results on a similar task obtained by subjects who were fitted with a multi-channel device similar to the one used by our subjects. Huang et al. performed a study which employed a four-alternatives, forced choice task [10]. Their subjects were Mandarin-speaking adult implantees who were fitted with the Nucleus 22-channel cochlear mini system. The subjects were able to perceive Mandarin tones with about 68% accuracy compared to a preoperative performance of 34.5% correct. [10] suggested that the acoustic cues to fundamental frequency of the four Mandarin tones can be extracted by the speech-coding strategy of the 22-channel cochlear implant system and stimulate the auditory nerve where they are perceived as pitch. There are two differences between the two studies could account for the seemingly contrasting findings by [10] and by the present study. First, the subjects in that study were adults who had auditory and phonological knowledge before the implantation, and therefore might have been able to benefit more from the auditory capacity provided by the implant than prelingually deaf children [11]. Adults may also perform better than children in making decision in a tone identification task because of more advanced cognitive skills in comparison with the young children. Second, Cantonese speaker do not rely upon temporal cues for tone perception [5] while Mandarin speakers can use temporal cues effectively [2], [3], [4]. Therefore, it is possible that the better performance of listeners with cochlear implants in perceiving Mandarin tones may have to do with the perception of temporal cues rather then pitch perception. The present data suggest that although the Nucleus 24 cochlear implant system might be able to extract fundamental frequency information from speech sounds, the effectiveness of the implant in transmitting this information to the auditory system is still far from equivalent to that of a normal hearing system.

5. REFERENCES