Tone production and perception and intelligibility of produced speech in Mandarin-speaking cochlear implanted children

Yi-Lu Li, Yi-Hui Lin, Hui-Mei Yang, Yeou-Jiunn Chen & Jiunn-Liang Wu

To cite this article: Yi-Lu Li, Yi-Hui Lin, Hui-Mei Yang, Yeou-Jiunn Chen & Jiunn-Liang Wu (2017): Tone production and perception and intelligibility of produced speech in Mandarin-speaking cochlear implanted children, International Journal of Audiology, DOI: 10.1080/14992027.2017.1374566

To link to this article: https://doi.org/10.1080/14992027.2017.1374566

Published online: 14 Sep 2017.

Article views: 55
Tone production and perception and intelligibility of produced speech in Mandarin-speaking cochlear implanted children

Yi-Lu Li1, Yi-Hui Lin2, Hui-Mei Yang1, Yeou-Jiunn Chen3, and Jiunn-Liang Wu1

1Department of Otolaryngology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, 2Department of Otolaryngology, Tainan Municipal Hospital, Tainan, Taiwan, and 3Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan

Abstract

Objectives: This study explored tone production, tone perception and intelligibility of produced speech in Mandarin-speaking prelingually deaf children with at least 5 years of cochlear implant (CI) experience. Another focus was on the predictive value of tone perception and tone production as they relate to speech intelligibility. Design: Cross-sectional research. Study sample: Thirty-three prelingually deafened children aged over eight years with over five years of experience with CI underwent tests for tone perception, tone production, and the Speech Intelligibility Rating (SIR). A Pearson correlation and a stepwise regression analysis were used to estimate the correlations among tone perception, tone production, and SIR scores.

Results: The mean scores for tone perception, tone production, and SIR were 76.88%, 90.08%, and 4.08, respectively. Moderately positive Pearson correlations were found between tone perception and production, tone production and SIR, and tone perception and SIR (p < 0.01, p < 0.01 and p < 0.01, respectively). In the stepwise regression analysis, tone production, as the major predictor, accounted for 29% of the variations in the SIR (p < 0.01).

Conclusions: Mandarin-speaking cochlear-implanted children with sufficient duration of CI use produce intelligible speech. Speech intelligibility can be predicted by tone production performance.

Key Words: Cochlear implant, tone perception, tone production, speech intelligibility, Mandarin Chinese

Introduction

Cochlear implants (CIs) can provide significant improvements in the lives of people with profound deafness, aiding in the acquisition of auditory information and linguistic development (Tye-Murray, Spencer, and Woodworth 1995; Svirsky et al. 2000). Although the CI is effective for most recipients, the perceptual and productive performance of CI users does not reach the same level as that seen in the normal-hearing population (Peng, Tomblin, and Turner 2008). As a lexical tone language, Mandarin consists of suprasegmental parts based on variations in fundamental frequency (F0). The four major tones forming the F0 contour in a syllable can be characterised as follows: flat for tone 1 (T1), rising for tone 2 (T2), falling and rising for tone 3 (T3), and falling for tone 4 (T4). Unlike most Western syllabic languages, variations in the different tones in the F0 contour carry distinct meaning in such tonal languages, even though the phonemic components (consonants and vowels) are identical. Due to the limited number of channels used to transmit the wide range of speech frequencies, the low spectral resolution of the CI device restricts the coding of fine spectral detail necessary for resolving F0 harmonics as a cue to pitch (McKay, McDermott, and Clark 1994; Zeng 2002; Moore 2003; Liu, Chen, and Lin 2004). While spectral detail is restricted, F0 information is coded in the temporal envelope of the electrical stimulus signal, which is capable of eliciting cues to F0 pitch although they are far weaker than those for listeners with normal hearing. These technological restrictions may increase the differences in linguistic performance between children with CIs and those in populations with normal hearing.

An acoustic analysis indicated that CI recipients have obvious difficulties in producing all tone patterns with regard to the irregular and flat contours of F0, with large individual differences (Xu et al. 2004). In order to improve pitch perception, a number of novel speech processing strategies have been developed. Advanced combination encoder (ACE) and continuous interleaved sampling (CIS) approaches have been developed to enhance F0 coding, but...
only small and inconsistent benefits have been shown with these approaches (Wong et al. 2008; Han et al. 2009; Schatzker et al. 2010; Teng 2011; Hwang et al. 2012). Perception of lexical tones thus remains challenging for CI users who communicate using tonal languages. The feasibility of enhanced temporal cues using novel pitch coding strategies other than ACE or CIS for tone language perception has been studied and proven (Milczynski et al. 2012; Vandali, Dawson, and Arora 2016). New pitch coding strategies focus on temporal envelope modulations, which are important cues for lexical tone perception in the electric hearing population. The results from such studies show significantly better tone recognition, albeit no difference in sentence recognition (Luo et al. 2008; Milczynski et al. 2012).

Recent studies have demonstrated that prelingually deaf children with CIs face difficulties with perceiving lexical tone languages (Xu et al. 2009, 2011). An average tone recognition score of 73% was reported by Peng, Tomblin, and Cheung (2004) in a study involving 30 paediatric CI users, but the performances varied from chance (50% correct) to nearly perfect. The lack of adequate auditory input as well as feedback may give CI users problems with regard to learning tone production. As seen with the tone perception of CI users, individual variability in tone production performance in previous studies has been high, with scores ranging from chance to nearly perfect (Han et al. 2007; Zhou and Xu 2008; Xu et al. 2009). Age at implantation was found to be a significant contributor to this large variability in tone production (Han et al. 2007; Xu et al. 2009).

A recent study of Cantonese tone production in congenitally deaf children with CIs also indicated that early implantation (before the age of 4 years) is crucial for good tone production (Lee, van Hasselt, and Tong 2010).

In lexical tone identification, secondary acoustic cues, such as amplitude contour and tonal duration, also provide essential information (e.g. Liang 1963; Liu and Samuel 2004; Kuoa, Rosen, and Faulkner 2008), especially when F0 information is degraded. In a Mandarin identification test, Whalen and Xu (1992) found that the population with normal hearing got 80% of the questions correct without the F0 information. Their study indicated that the amplitude contour contributed to this result more than the tone duration among the secondary cues. For CI users, this is particularly important given the poor coding and perception of F0 information in electrical stimulation.

Speech intelligibility is a measure of the extent to which listeners receive the verbal information that speakers intend to present (Whitehill and Ciocca 2000). For congenitally deaf children receiving a CI, speech intelligibility has been recognised as a key factor in measuring the development of speech production. Among the various measurements used to evaluate the linguistic outcomes of CI children, the Speech Intelligibility Rating (SIR), a hierarchical scale ranking speech into one of five categories, has been shown to be a reliable and simple tool (Allen, Nikolopoulos, and O’Donoghue 1998; Cox and McDaniel 1989).

In a previous study, although speech intelligibility was found to be poorer in CI children compared to the population with normal hearing, the communicative outcomes as measured with SIR remained satisfactory (Huang et al. 2005). The possible contributors to speech acquisition and intelligibility in CI users, such as residual hearing, duration of implant use, and age at implantation, have been examined in many studies. Speech intelligibility in CI children has been shown to be negatively correlated with age at implantation, but positively correlated with the length of time the child has been using a CI (Mondain et al. 1997; Nikolopoulos, O’Donoghue, and Archbold 1999; Wu and Yang 2003; Huang et al. 2005; Han et al. 2007).

Since Mandarin is a tonal language, the same phonemes spoken within different tones have different meanings. Therefore, just like the vowels and consonants that appear as phonemes in words, tones in Mandarin represent an irreplaceable component in words and sentences. As yet, however, the relationship between tone production and perception and speech intelligibility has not been studied. Consequently, the first aim of this study was to examine the predictive values of tone perception and tone production as they relate to speech intelligibility. As mentioned above, the age at implantation and the degree of experience with CIs are the two main factors that impact speech intelligibility. We thus also tried to evaluate the speech intelligibility and tone production and perception of Mandarin-speaking children who received a CI at a young age who had already experienced many years of using their CI.

Materials and methods

Subject selection

In this cross-sectional research, children with prelingually bilateral sensorineural hearing loss were recruited from the Otolaryngology Department of National Cheng Kung University Hospital (NCKUH), Tainan, Taiwan. Patients who had received a CI by the same operator were assembled for our research and extracted by convenience sampling from a pool of subjects that met the following selection criterion: to attain more reliable results, all the participants in this study were required to be able to read a sufficient number of vocabulary items and follow instructions. As a consequence, prelingually deaf children aged at least 8 years old who had received a unilateral CI were eligible for this study. Speech intelligibility tends to develop steadily over time after implantation (Allen, Nikolopoulos, and O’Donoghue 1998; Calmels et al. 2004; Bakhshaee et al. 2007), especially after a period of three to five years, and thus the participants in this study all were required to have used CIs for more than five years. The subjects also had to be able to communicate with others using auditory-verbal communication in Mandarin Chinese. Children who had other physical disabilities, such as blindness, neuromuscular disease, craniofacial deformities, upper limb defects, and mental retardation were excluded, as their other impairments would potentially cause problems when they were completing the required tasks. The participants’ background data, such as gender, aetiology, age at hearing loss, age at implantation, device manufacturer and type, and age at the time of testing were collected. All subjects’ guardians signed an informed consent as permission for enrolment. The use of human subjects authorised by guardians was reviewed and approved by Institutional Review Board of National Cheng Kung University Hospital (IRB No: B-BR-101-104).

A total of 33 children with unilateral CIs were enrolled in this study. Seven of these prelingually deaf children had a family history of hearing impairments; one was diagnosed with Mondini syndrome; one was diagnosed with large vestibular aqueduct syndrome (LVAS), and the aetiology of the remaining 24 children remained unknown. Twenty-seven subjects used Nucleus’s CI devices (Cochlear, Sydney, NSW, Australia), and six used Clarion’s devices (Advanced Bionics, Valencia, CA). The chronological ages of the CI subjects at the time of testing ranged from 8.5 to 17.7 years old (mean = 13.4 years). The age at the diagnosis of hearing loss ranged from 0.1 to 3.7 years old (mean = 1.1 years). Figure 1 reveals the age at implantation (ranging from 0.9 to 4.9 years,
mean $= 2.5$ years) and the duration of CI use (range 6.0 to 14.8 years, mean $= 10.8$ years). The mean pure tone average (PTA) of the unaided CI subjects was over $100$ dB HL, and the PTA for the better ear was about $60$ dB HL after fitting with hearing aids. After implantation, the PTA was $30$ dB HL, on average.

Environment and equipment of the experiment

The tests were conducted using a repeated-measures counterbalanced design across the subjects to minimise the effects of order on outcomes. Every three subjects took turns undergoing the different tests at different places at the same time, with a ten-minute break between tests. Each participant performed the tone perception test in an anechoic room, and quiet rooms (ambient noise level $= 40$ dB SPL; long-term averaged level, A-weighting) were provided for the tone production and sentence imitation tests. Testing materials were displayed on the screen of a laptop with or without the audio being broadcast through an Elephant Aoide PC speaker (signal-to-noise ratio $\geq 85$ dB), and a microphone (Samson C01U) was used to record the speech produced by the subjects. Each participant was asked to sit in a comfortable position in front of the laptop, and he/she could control a mouse to complete the tasks. All testing materials, procedures, and assessments were integrated into a programme to enable this process to be carried out on a computer.

Tone perception test

The tone perception test used materials based on the Mandarin Lexical Tone Recognition Test (MLTR), as edited by Teng (2011). The materials were retrieved from 637 monosyllables composed of vowels and consonants in different tones of Mandarin Chinese, and 2,548 speech files were recorded by four speakers to build the material pool. In proportion to the tones used for each phoneme, 132 test items were selected randomly from the material pool to be used by each participant.

In the testing, a tonal monosyllabic stimulus was presented at the most comfortable listening (MCL) level via a loudspeaker positioned in front of the subject; the MCL varied across children. The order of presentation was fully randomised. Following the twelve pre-tests, the formal questionnaire including 132 questions was administered. One tonal monosyllable was broadcasted at a time, using a 4-alternative forced choice (4-AFC) method, the subjects were instructed to choose one of the four tones shown on the computer screen that corresponded to what they heard. Subjects did not get any feedback after each response, thus they did not know if their answer was right or wrong.

Once the subject made his/her choice, both the answer and response time were recorded by the computer, and the percentage of responses that were correct was calculated, with a range from 0 to 100%.

Tone production test

Although the F0 contour is rather clear in terms of a monosyllabic word, disyllabic words represent a more natural tone interaction in real life. In the tone production test, each participant was thus examined individually and asked to pronounce a sampling of disyllabic utterances. The material was edited by the researchers (Supplementary Appendix 1) and consisted of 15 pairs of two-word phrases comprised of combinations of the four tones. To avoid the confusing situation in which the participants might encounter tonal co-articulation in Mandarin, we excluded the disyllable in a T3 structure followed by T3 (T3-T3). In speaking the T3-T3 disyllable, the former T3 term should be produced as T2. It is classified as the “anticipatory co-articulation.”

Most of our subjects were students in elementary school who had well-developed quite ability to recognise vocabulary items. To reduce the negative impact of unfamiliar words, the materials in this test were common, concrete nouns retrieved from textbooks used in the junior grades of elementary school.

After the four training pre-tests, each participant was provided with and pronounced 30 trials labelled with Traditional Chinese and Mandarin Phonetic Symbols that were displayed randomly on the computer screen, and the speech produced by the child was recorded at a sampling rate of 44,100 Hz with a 16-bit resolution.
Two speech therapists with normal hearing specialising in working with hearing-impaired children judged the recorded utterances subjectively. As with the tone perception task, the participants’ tone production scores were also converted into percentages. The results were counted in phonemes. Thus, the total number of points for each subject was divided by 60 and multiplied by 100% to obtain an accurate score, with a range from 0 to 100%.

Speech intelligibility rating (sentence imitation test) We used the SIR as a measure of results for the sentence imitation test, which classifies spontaneous speech into a five hierarchic categories as follows: connected speech is intelligible to all listeners, and the child is understood easily in everyday contexts (category 5); connected speech is intelligible to a listener who has little experience with the speech of a deaf person (category 4); connected speech is intelligible to a listener who concentrates and lip-reads (category 3); connected speech is unintelligible, and intelligible speech is developing in single words when context and lip-reading clues are available (category 2); connected speech is unintelligible, and pre-recognizable words appear in spoken language (category 1). The sentence imitation test involved identification and repetition of 20 sample sentences of five different lengths (four, six, eight, ten, and twelve words), and each sentence ended in a different tone (T1, T2, T3, and T4) (Supplementary Appendix 2). The material was pre-recorded using the speech of a 12-year-old girl with normal hearing in a quiet room.

Speech was presented randomly at the MCL level, which was specific to each subject, via a loudspeaker placed in front of the subject, and subtitles in Traditional Chinese with Mandarin Phonetic Symbols were also shown on the computer screen. Five sentences were provided for practice before the formal test. Each subject repeated the sentences he/she heard, and the speech was recorded at a sampling rate of 44,100 Hz with a 16-bit resolution.

The speech recorded from all the participants was pooled together and renumbered, and each sentence was retrieved in a random order to be scored subjectively by two speech therapists, and the child is understood easily in everyday contexts (category 5); connected speech is intelligible to all listeners, and the child is understood easily in everyday contexts (category 5); connected speech is intelligible to a listener who has little experience with the speech of a deaf person (category 4); connected speech is intelligible to a listener who concentrates and lip-reads (category 3); connected speech is unintelligible, and intelligible speech is developing in single words when context and lip-reading clues are available (category 2); connected speech is unintelligible, and pre-recognizable words appear in spoken language (category 1). The sentence imitation test involved identification and repetition of 20 sample sentences of five different lengths (four, six, eight, ten, and twelve words), and each sentence ended in a different tone (T1, T2, T3, and T4) (Supplementary Appendix 2). The material was pre-recorded using the speech of a 12-year-old girl with normal hearing in a quiet room.

Speech was presented randomly at the MCL level, which was specific to each subject, via a loudspeaker placed in front of the subject, and subtitles in Traditional Chinese with Mandarin Phonetic Symbols were also shown on the computer screen. Five sentences were provided for practice before the formal test. Each subject repeated the sentences he/she heard, and the speech was recorded at a sampling rate of 44,100 Hz with a 16-bit resolution.

The speech recorded from all the participants was pooled together and renumbered, and each sentence was retrieved in a random order to be scored subjectively by two speech therapists, without providing the identities of the speakers. Listeners ranked each sentence into one of the five categories, with higher scores for more intelligible speech. The mean scores of all utterances from the same speaker and all the sentences spoken by participants’ tone production scores were also converted into percentages. The results were counted in phonemes. Thus, the total number of points for each subject was divided by 60 and multiplied by 100% to obtain an accurate score, with a range from 0 to 100%.

Speech was presented randomly at the MCL level, which was specific to each subject, via a loudspeaker placed in front of the subject, and subtitles in Traditional Chinese with Mandarin Phonetic Symbols were also shown on the computer screen. Five sentences were provided for practice before the formal test. Each subject repeated the sentences he/she heard, and the speech was recorded at a sampling rate of 44,100 Hz with a 16-bit resolution.

The speech recorded from all the participants was pooled together and renumbered, and each sentence was retrieved in a random order to be scored subjectively by two speech therapists, without providing the identities of the speakers. Listeners ranked each sentence into one of the five categories, with higher scores for more intelligible speech. The mean scores of all utterances from the same speaker and all the sentences spoken by all the subjects were summarised and then analysed.

Due to the subjective scoring method used in the tone production test and SIR, Kendall’s coefficient of concordance was used to estimate the consistency between the two speech therapists, and the W = 0.88 indicated that the two sets of results were in good conformity with each other, significant at p < 0.01.

Statistical analysis To assess the mismatch between the most difficult tone for the subjects in regard to identification and production, confusion matrices were produced using the results of the tone perception and production tests. To estimate the relationships among tone perception, tone production, and SIR, a Pearson correlation was applied to evaluate the three co-variants. If any significant correlation existed, then a stepwise regression analysis was performed to clarify the possible causality. The statistical analysis was processed using SPSS version 22.0. Statistical significance was defined as p values less than 0.05.

Results The performance of the 33 subjects in tone perception is presented as a confusion matrix in Table 1. The mean score of overall four-tone perception was 76.88% (SD = 10.26%). The mean scores of tone 2 (mean ± SD = 58.83% ± 18.92%) and tone 3 (mean ± SD = 72.38% ± 15.08%) were significantly lower than those of tone 1 (mean ± SD = 84.05% ± 16.96%) and tone 4 (mean ± SD = 92.25% ± 10.26%) (p < 0.01). Tone 2 and tone 3 were more often confused in terms of identification. Among the 41.17% of incorrect results for tone 2, 28.84% were misidentified as tone 3, and 12.23% as tone 1. Among the 27.61% of incorrect results for tone 3, most were misidentified as tone 2 (25.97%).

Table 1. Confusion matrix for tone perception (%).

<table>
<thead>
<tr>
<th>Recognized tone</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target tone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>84.05</td>
<td>9.27</td>
<td>5.44</td>
<td>1.25</td>
</tr>
<tr>
<td>T2</td>
<td>12.23</td>
<td>58.83</td>
<td>28.84</td>
<td>0.10</td>
</tr>
<tr>
<td>T3</td>
<td>1.47</td>
<td>25.97</td>
<td>72.38</td>
<td>0.17</td>
</tr>
<tr>
<td>T4</td>
<td>1.34</td>
<td>1.87</td>
<td>4.55</td>
<td>92.25</td>
</tr>
</tbody>
</table>

Table 2. Confusion matrix for tone production (%).

<table>
<thead>
<tr>
<th>Recognized tone</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Produced tone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>98.58</td>
<td>0.85</td>
<td>0.38</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>T2</td>
<td>4.55</td>
<td>81.54</td>
<td>12.69</td>
<td>0.76</td>
<td>0.19</td>
</tr>
<tr>
<td>T3</td>
<td>5.11</td>
<td>10.98</td>
<td>83.62</td>
<td>0.28</td>
<td>0.00</td>
</tr>
<tr>
<td>T4</td>
<td>1.80</td>
<td>0.00</td>
<td>0.85</td>
<td>96.59</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Unknown: the pronunciation of the target tone cannot be correctly recognised.
SIR was the dependent variable in a multiple-factor regression analysis. In general, a stepwise regression was applied to choose the predictors for fitting regression models, and the $p$-value of F-test was set at 0.1 for the purpose of removing a predictor. The results indicated that tone perception was removed. Therefore, tone perception had the least statistically deleterious effect on the model fit. In contrast, tone production was the major predictor of the SIR results ($R^2 = 0.312$, $R^2_{adj} = 0.290$, $p < 0.01$). This meant

Figure 2. Results of SIR (speech intelligibility rating) for individual subjects. The mean score of SIR is 4.08, presented as the horizontal bar (SD = 0.50).

Figure 3. (a) Scatter diagram with a regression line for tone perception and tone production ($r = 0.672$, $p < 0.01$). (b) Scatter diagram with a regression line for tone perception and SIR (speech intelligibility rating) ($r = 0.513$, $p < 0.01$). (c) Scatter diagram with a regression line for tone production and SIR (speech intelligibility rating) ($r = 0.559$, $p < 0.01$).
that 29% of the variation in SIR data was correlated with that of tone production.

**Discussion**

Good speech perception is a prerequisite for children to produce intelligible speech (Laches, Pisoni, and Kirk 2001; Casserly and Pisoni 2010). As discussed earlier, because Mandarin is a tonal language, a word’s meaning changes when it is produced with a different tone contour even though the phonemic components are the same. For Mandarin speakers, tone performance plays an important part in oral communication as well as other phonemes such as vowels and consonants. We assume intelligible Mandarin speech also depends on good perception of tonal information. The results of our study support our assumption. First, a positive correlation between tone perception and tone production was found. A few studies in the past have also discussed the relationship between tone recognition and production in children with CIs. Consistent with the hypothesis that perception precedes production in children that are developing normally (Edwards 1974), CI children with better tone perceptions tend to have higher accuracy in tone production tasks (Peng, Tomblin, and Turner 2008; Xu et al. 2009; Zhou et al. 2013). Xu et al. (2009) reported a significantly strong correlation between tone perception and production performance, where the Pearson correlation coefficient \( r \) was 0.805. Secondly, SIR was positively correlated to tone perception and production. Tone production accounted for 29% of the variation in SIR. In Mandarin speakers with a CI, intelligibility of produced speech can mainly be predicted by their tone production performance. This result, which has not been stated in previous literature on this topic, is reasonable with the regard to SIR criteria depending on connected speech. The more clearly the four tones are pronounced, the more accurate the meaning is.

In the present study, all the subjects had received implantation before 5 years of age and had used their CIs for an average of a decade. Presumably, this experience had helped them to develop audio-oral communication. In spite of the limited perceptive ability with regard to lexical tone languages that can be obtained using CI technologically, most of the speech produced by these early-implanted children was intelligible to listeners. Secondary acoustic cues that are transmitted by the CI devices in these children may be helpful in tone identification and thus influence their production. Similar results were found in past studies, although small differences existed in the experimental results among these studies. Fang et al. (2014) examined 84 Mandarin-speaking children receiving CIs before five years old who were not aware of speech sounds and had an unintelligible speech before implantation. The results showed that after five years of implant use, conversations of most of the subjects could be understood without lip-reading, and more than half were fully intelligible to all listeners. Likewise, all our subjects had received aural rehabilitation and speech training for ten years on average before joining the study. Experienced speech therapists combined both primary and secondary cues into the training material and helped these implanted children to recognize the characteristics of each tone and to pronounce them correctly. We speculate that after a sufficient period of CI use and integrated speech training, prelingually deafened children will be able to speak understandable Mandarin.

The mean scores of our CI subjects in the tone production and perception tasks were higher than those reported in the previous studies (Peng, Tomblin, and Cheung 2004; Han et al. 2007; Zhou and Xu 2008; Xu et al. 2009, 2011), as was the SIR performance. Moreover, a smaller variance in tone perception was found in this study compared with Teng (2011). Teng’s trial involving 39 CI subjects shared similar demographic profiles with our participants in categories including mean chronological age and geographic distribution in Taiwan. The accuracy of tone perception in Teng’s research ranged from 31.06% to 93.18% correct \( (\text{mean} \pm \text{SD} = 65.86\% \pm 15.39\%) \), but better performance with fewer differences was revealed in our study \( (\text{mean} \pm \text{SD} = 76.88\% \pm 10.26\%) \). Our subjects had a lower mean age at implantation and a longer mean duration of CI use, respectively, which may have contributed to the difference in the results. In addition, good performance in tone production with reduced individual variability was also found in our subjects, in contrast to the results of previous studies (Peng, Tomblin, and Cheung 2004; Han et al. 2007; Zhou and Xu 2008; Xu et al. 2009), although differences in test methods and materials may explain this difference. Furthermore, age at implantation has been demonstrated to be a contributor to variability in tone production (Xu et al. 2009; Lee, van Hasselt, and Tong 2010; Zhou et al. 2013). We assume the effect of implantation age and the experience with CI use to be the main factors influencing the results of our study. Most of our subjects received a CI before 4 years of age, followed by an average 10-year period of auditory rehabilitation. With regard to maturity in language development and sufficient experience with CI use, the CI recipients who had been implanted at a young age had excellent tone production and intelligibility of produced speech performance in the current study.

Although tone production correlates to tone perception, a lower overall correct rate of tone perception was found in this study compared with that for tone production \( (\text{mean} \pm \text{SD} = 76.88\% \pm 10.26\% \text{ vs. } 90.08\% \pm 6.64\% \) respectively). The differences in materials, testing methods, and scoring systems may have contributed to this. In the tone perception test, the participants had to answer the most questions among the tests being administered and had only auditory stimuli. In contrast, fewer questions were asked in the tone production test and SIR, and visual clues were displayed on the screen at the same time. In contrast to the objective scoring method used with the tone perception test, two adults with normal hearing assessed the results of the tone production test and SIR subjectively although the inter-rater consistency was verified to be high.

Among the four different tones in Mandarin, the results of the tone perception test and tone production test indicated that the CI children faced more difficulties in identifying T2 and T3 as well as in producing them. A number of previous studies have also shown that after implantation, congenitally deaf children still perform unsatisfactorily in discriminating between T2 and T3 (Peng, Tomblin, and Cheung 2004; Lin and Peng 2009). The ambiguity in the F0 contour between these two tones, with T2 rising, and T3 falling and then rising, may contribute to this phenomenon. The characteristics of amplitude and tone duration among the four tones also vary a lot, leading to their influence on tone perception and production. In general, T4 has the largest amplitude and the shortest duration, while T3 has the least amplitude and the longest duration (Chuang et al. 1975; Tseng 1990). Among children with normal hearing, T1 and T4 have been demonstrated to be mastered earlier than T2 and T3 (Li and Thompson 1977; Tse 1978; Su 1985). A similar but slower developmental pattern of tone acquisition in hearing-impaired children compared to children with normal hearing was found by Chen (1986). Some studies have also
proposed that the Mandarin tone system is acquired early (by age 6) with respect to its production in children with normal hearing. Consequently, congenitally deaf children with early implantation may exhibit similar tone production and perception to that of the population with normal hearing. Furthermore, as mentioned above, the current perceptual limitations in lexical tone languages remain to be addressed by further developments in current CI technology. Although the findings investigating the correlation between SIR and tone production, perception in this study have not been discussed in previous studies, some limitations exist in the present work that should be noted. First of all, based on the convenience sampling method and selection of subjects used in this work, the small sample size and the limited region from which the subjects were obtained mean that the results may not represent the entire population of Taiwan or other Mandarin-speaking countries. All subjects in the present study received CIs at under five years of age, and their performance cannot thus represent the level of intelligible speech among deaf people who received implants when they were older. Secondly, due to the lack of a standard test for tone production, the materials in this study were self-produced without an analysis of their validity. Finally, although the consistency between listeners was verified using Kendall’s coefficient test, the scoring by adults with normal hearing in the tone production test and SIR may have been too subjective to reveal the actual performance of the individuals involved in this work. An ongoing task is to obtain a much larger sample size so that the results of any future study will have greater statistical power. Also, the effect of the F0 contour and secondary acoustic cues on perceptual and produced speech for implanted children should be investigated in further more objective experiments.

**Conclusions**

The results of the present study implied that prelingually deaf children who received implants when they were older. Secondly, due to the lack of a standard test for tone production, the materials in this study were self-produced without an analysis of their validity. Finally, although the consistency between listeners was verified using Kendall’s coefficient test, the scoring by adults with normal hearing in the tone production test and SIR may have been too subjective to reveal the actual performance of the individuals involved in this work. An ongoing task is to obtain a much larger sample size so that the results of any future study will have greater statistical power. Also, the effect of the F0 contour and secondary acoustic cues on perceptual and produced speech for implanted children should be investigated in further more objective experiments.

**Acknowledgements**

This study was supported by a grant from the Ministry of Science and Technology, Taiwan (NSC101-2221-E-006-019-MY3). The research was also supported by National Cheng Kung University Hospital, Tainan, Taiwan.

**Declaration of interest:** The authors have neither financial disclosures nor conflicts of interest relevant to this article.

**References**


**Intelligibility of produced speech in mandarin CI children**


Supplementary material available online