Clinical Linguistics & Phonetics, September 2011; 25(9): 784–803

Acoustic characteristics of vowels and plosives/affricates of Mandarin-speaking hearing-impaired children

SHU-CHUAN TSENG, KO KUEI, & PEI-CHEN TSOU

Institute of Linguistics, Academia Sinica, Taipei, Taiwan

(*Received 15 April 2010; Accepted 21 February 2011*)

Abstract
This article presents the results of an acoustic analysis of vowels and plosives/affricates produced by 45 Mandarin-speaking children with hearing impairment. Vowel production is represented and categorized into three groups by vowel space size calculated with normalized F1 and F2 values of corner vowels. The correlation between speech intelligibility and language abilities assessed by the level of word comprehension and the complexity of sentence structure is statistically significant. Vowel space grouping is correlated with speech intelligibility and spike percentage of plosives/affricates production. The generalized linear model analysis also shows that the level of word comprehension and the degree of hearing loss are the two most significant factors in predicting speech intelligibility. The statistical results suggest that the interplay of acoustic characteristic and speech ability is complex.

Keywords: *Mandarin, hearing impairment, acoustics, vowels, plosives, affricates*

Introduction

When using speech for communication, the phonetic structure including segmental and prosodic contrast is articulated and decoded. During normal speech development, children have many opportunities to hear others speak. For children with hearing impairment, much of the auditory input may not be accessible, which may hinder speech development (Serry and Blamey, 1999; Blamey, Barry, Bow, Sarant, Paatsch, and Wales, 2001a; Peng, Weiss, Cheung, and Lin, 2004c; Bouchard, Normand, and Cohen, 2007). For example, Serry and Blamey (1999) concluded that the development of speech in hearing-impaired children followed a process similar to that of normally hearing children, although at a slower rate. For example, diphthongs are acquired later than monophthongs, and fricatives and affricates are more difficult than nasals and plosives. Bouchard et al. (2007) stated that stops and labials are the most frequently produced consonants soon after cochlear implantation, consistent with a schedule of typical speech development. From the methodological point of view, it is not realistic to collect direct data on how hearing-impaired children hear. But we can study the production data and assess their perception ability of discriminating linguistic units and comprehending the meaning of a spoken discourse.

Correspondence: Shu-Chuan Tseng, Institute of Linguistics, Academia Sinica, Nankang, 115 Taipei, Taiwan. Tel: 886-2-26525014. Fax: 886-2-27856622. E-mail: tsengsc@gate.sinica.edu.tw

ISSN 0269-9206 print/ISSN 1464-5076 online
DOI: 10.3109/02699206.2011.565906

Speech production and perception

What one hears affects one's speech intelligibility. Thus, the speech of hearing-impaired children is often less intelligible than that of typically developing children. Using speech intelligibility scores, Peng et al. (2004c) reported that Mandarin-speaking hearing-impaired children showed lower intelligibility and delayed patterns in their development of consonant production compared with children with normal hearing. Based on results of a 6-year study of conversational speech in nine hearing-impaired children with cochlear implants (CIs), Blamey et al. (2001a) concluded that 90% of their syllables were intelligible and that production of speech sounds and words steadily improved. However, after 6 years of implantation, the speech acquisition process was still incomplete. Maassen and Povel (1985) manipulated the acoustic features of deaf children's speech by replacing segments with those spoken by hearing speakers and examined intelligibility by listeners. The segmental correction of vowels caused a dramatic increase in intelligibility from 24% to 72%, whereas corrections of temporal structure and intonation caused only a small improvement from 24% to 34%. Chuang (2006) used percentages of correctly produced vowels, consonants, segments, words, tones and the overall speech intelligibility of hearing-impaired children to study their relationships with vowel space (i.e. formant relationships). The vowel space size was correlated with accuracy rates of vowels, segments and words, but not with those of consonants, tones and speech intelligibility. Peng, Spencer, and Tomblin (2004a) reported that the percentage of correctly produced segments was highly correlated with rating-scale intelligibility of speech produced by Mandarin-speaking hearing-impaired children. In other words, impressionistic assessment obtained by listener ratings of speech is suitable for representing speech intelligibility and for conducting research on speech production of hearing-impaired children. Taken together, past results suggest that speech intelligibility of hearing-impaired subjects needs to be analysed together with different assessment methods such as acoustic properties and impressionistic evaluations.

It is reasonable to assume that speech perception problems lead to speech delays and reduced intelligibility in children with hearing impairment. Support for this assumption can be found in studies of children with CIs, which tend to find that speech perception improves more with earlier implantation (Blamey, Sarant, Paatsch, Barry, Bow, Wales, Wright, Psarros, Rattigan, and Tooher, 2001b; Wu and Yang, 2003; Mildner, Šindija, and Zerinski, 2006). In a study of Mandarin-speaking hearing-impaired children, Wu and Yang (2003) concluded that speech perception improved over a 2-year follow-up period, with improvements negatively correlated with the age at implantation and positively correlated with the duration of implant use. Despite these improvements, some speech problems apparently remain after implantation. For example, Ciocca, Francis, Aisha, and Wong (2002) stated that young children with CIs had great difficulty in extracting voice pitch information from natural speech sounds to accurately identify Cantonese lexical tones. Speech perception ability is also associated with the language level of hearing-impaired children. Blamey et al. (2001b) reported that most of their hearing-impaired children can achieve a level of approximately 90% sentence recognition in the audio–visual test condition when their language level is equivalent to that of a normally hearing child aged from 7 years or older.

Factors

Both non-linguistic and linguistic factors need to be considered when studying speech abilities of children with hearing impairment. The non-linguistic factors proposed by Musselman and Kircaal-Iftar (1996) included audiological factors, demographic factors and educational history. More were mentioned by Peng et al. (2004c): age at onset of deafness, age at implantation, duration of deafness before implantation, duration of CI use, physiological or device-related factors, and psychological, educational and social factors such as a recipient's motivation and level of intelligence. The linguistic factors include different levels of linguistic properties. Acoustic parameters such as duration, pitch, intensity and formant pattern are used to evaluate articulation manner, articulation place, voicing of vowels and consonants and prosodic qualities (Ling, 1976). Vowel centralization and reduction of segments and tones are often observed in terms of vowel space shrinkage, duration change and variation of F0 (Pickett, 1999; Kent and Read, 2002). Word tokens, part of speech and semantic association are studied to explore the structure of the mental lexicon, whereas sentential structure and types reflect the ability to produce sentences.

Acoustic properties of vowels and plosives/affricates

Acoustic features of speech are empirical cues to physical properties measured from digitized speech signals. Acoustic measurements help to specify patterns of articulation, as well as prosodic features such as intonation, tone, rhythm and stress (Lieberman and Blumstein, 1988; Ho, 1996; Pickett, 1999). The articulatory label of tongue height is represented by F1 in acoustic phonetics. High vowels have low F1 values, whereas low vowels have high F1 values. The tongue position of the vowel articulation performing the anterior–posterior (or front–back) contrast is represented by F2. The F2 of back vowels is low and the F2 of front vowels is high (Lieberman and Blumstein, 1988; Pickett, 1999; Kent and Read, 2002). Tye-Murray, Spencer, and Woodworth (1995) found that hearing-impaired children with CIs were able to enlarge their phonological system of vowels with increased access to formant information. Lin and Wang (1995) presented the typical formant patterns of Mandarin vowels. In addition to the explanatory power of F1 and F2 in terms of front/back and high/low contrasts, vowel space size can also be used to assess vowel production in terms of their acoustic quality. Liker, Mildner, and Šindija (2007) reported a smaller and fronted vowel space of children with CIs compared with that of normally hearing children. Ling (1976) provided a fine discussion on the interaction of acoustic properties and their sensory correlates of speech produced by hearing-impaired children. He stated that most vowels spoken in isolation can be identified if both the first and second formants are audible. Mildner and Liker (2008) recorded nonsense syllables to investigate the formant-defined vowel space and noise frequency of fricatives /s, ʃ/ produced by hearing-impaired and normally hearing children matched for age and gender. They found that hearing-impaired children produced vowel space similar to the hearing controls. However, their fricatives did not present clear distinctions in the noise spectrum.

Plosives should be audible to children who have only low-frequency residual hearing due to the fast F1 transition phase starting at lower frequencies (Ling, 1976). Typical production of plosives and affricates contains five articulation phases that can be clearly identified in the spectrogram: the occlusion, the transient, the frication, the aspiration and the transition (to the vocalic part) (Lieberman and Blumstein, 1988; Kent and Read, 2002). Consonant production is reported to be correlated with speech intelligibility by Shriberg, Austin, Lewis, McSweeny, and Wilson (1997). Lee (1999) studied the affricates produced by 16 hearing-impaired and 35 normally hearing children at primary school and found that both groups of subjects produced aspirated affricates longer than the unaspirated counterparts. Another way to analyse acoustic features of plosives and affricates is to examine the occurrence of spike. Spike is the brief pulse of acoustic energy produced by the release burst of a plosive/affricate, which can be clearly observed in the spectrogram (Kent and Read, 2002). The hearing-impaired group produced spike in affricates less frequently. In another study, Lee (2001) studied the relationship between the production and the intelligibility of plosives and affricates produced by 22 hearing-impaired children. She concluded that voice onset time cannot be used to distinguish whether a plosive or an affricate is aspirated in hearing-impaired children. However, the degree of hearing loss is correlated with speech intelligibility. For children with better hearing, their intelligibility is higher. Khouw and Ciocca (2007) studied the voice onset time of plosives of 10 Cantonese-speaking hearing-impaired adolescents and found no significant difference between the voice onset time of aspirated and unaspirated plosives. As temporal features cannot directly reflect the correctness of the plosives/affricates production, the occurrences of spike production will be examined instead in our studies later.

Research Question

This article presents an acoustic analysis of 45 Mandarin-speaking children with hearing impairment. The acoustic properties of vowels and plosives/affricates are quantified, normalized across subjects and then statistically analysed in comparison with the speech abilities of our subjects. Speech abilities of the subjects include speech intelligibility scored by impressionistic judgements, word comprehension ability scored by the maximum number of content words in sentences that can be understood and sentence production ability scored by the degree of complexity of the sentence structure.

Method

Subjects

Forty-five Mandarin-speaking children with hearing impairment participated in this study. The mean age of the children at the time of recording was 6;1, ranging from 3;3 to 12;5. The children were receiving regular auditory-verbal therapy (AVT) sessions in the Children's Hearing Foundation (CHF) (2010) in Taipei. The AVT emphasizes the importance of cultivating spoken language ability of hearing-impaired children by maximally utilizing their residual hearing (Dornan, Hickson, Murdoch, and Houston, 2007). The CHF (2010) uses the 'level of auditory memory' to evaluate the level of word comprehension of hearing-impaired children. The level of auditory memory is defined as the number of 'content words' a child can maximally recognize from designed sentences spoken to her/him following a testing procedure defined by the CHF (2010). We recruited children whose auditory memory was level 4 and above (level 5+ being the highest) to ensure that our subjects had a comparable level of language ability. We were aware of the fact that the adopted level of auditory memory is only partially representative of the overall language ability of our subjects. This assessment was used as a perceptual measure in the statistical analysis. The degree of hearing loss ranged from mild to profound. As a result, 15 CI users and 30 traditional hearing aids (HAs) users were recorded. Our method of selecting subjects led to a rather diverse group of subjects in respect of degree of hearing loss, age and CI/HA subgroups, as listed in Appendix 1.

Recording

Video and audio data were recorded in a soundproof room at the CHF (2010) during children's regular therapy sessions. Mandarin has 22 consonants including plosives, affricates, fricatives, nasals and approximants. Depending on the definition of phonemic contrast and use of symbols, the number of vowels in Taiwan Mandarin may differ, but three corner vowels /i, a, u/ are uncontroversial (Ho, 1996). To collect Mandarin consonants and three corner vowels, 18 sentences were designed with words that children of their age were familiar with. Word–picture tests indicated that all children understood the meaning of each target word. The complete sentence list with translation is given in Appendix 2. The sentences were read and recorded by an adult female native speaker of Mandarin with a clear pronunciation. Presented by PowerPoint, each sentence was played twice with a matched cartoon picture shown on the screen for the children to repeat. As a result, 741 sentences were used for analysis.

Data processing

We used the Adobe Audition 1.0 to extract audio data from the original video recording with a sampling rate of 44.1 kHz, 16-bit, mono-track audio format. The speech data were processed for phonetic analysis by using PRAAT v.4.5.12. Main segmentation criteria were transition demarcation and pause segmentation. Two authors of this article segmented and labelled the syllabic and segmental boundaries as well as the articulation phases of plosives/affricates. Each was responsible for one half of the data and double-checked the other remaining half. Inconsistency was discussed with the first author to achieve consensus on the labelling results. As a result, 4080 syllables equivalent to 9460 segments were annotated. Among them, 1701 plosives and affricates were annotated for articulation types.

Formant measurement and vowel space calculation

Formants were automatically extracted for all tokens of /i, a, u/ in open syllables at the point where the maximal intensity was measured. Because normally the most stable formant pattern is observed at the point of maximal intensity, the point may correspond to the most sonorous location from the phonological point of view. The means of F1 and F2 values of each vowel were taken as the representative formant values for respective vowels. Vowel space was then drawn with F1 on the y-axis and F2 inversely on the x-axis, making it spatially comparable to the conventional IPA vowel chart. Normalization was done using the z-score to eliminate individual difference across multiple speakers. The z-score measure ($z = (x–\mu)/\sigma$) was adopted to normalize the F1 and F2 values of the vowels across different speakers, where x is the measured value, μ the mean and σ the standard deviation. The final statistical analysis was done using the normalized z-scores of the measured values.

The shape and the size of vowel space, that is, the acoustic characteristic, are assumed to reflect the degree of the articulated contrast in the tongue height and position of vowels. Conventionally, the vowel space size is the area of the triangle determined by /i, a, u/. We adopted a slightly different calculation method to represent the size of vowel space. The difference of F2 values of /i/ and /u/ indicates the front–back contrast of the tongue position, represented by $|F2(i)–F2(u)|$, whereas the difference between the minimum F1 values of /i/ and /u/ and the F1 value of /a/ reflects the maximum high–low contrast of the tongue height, represented by $|\min(F1(i), F1(u))–F1(a)|$. Vowel space size is then defined as the area of the quadrangle determined by these two differences. In other words, we emphasize the importance of the maximal F1 and F2 differences in the corner vowels. In normal cases, the quadrangle area is proportional to the triangle area. However, due to hearing impairment, the produced vowels may not always be located in the relative position as described in a standard vowel chart. By using the quadrangle, abnormally positioned vowels can be more clearly illustrated.

Labelling plosives and affricates

Taiwan Mandarin has six plosives /p, t, k, p^{h}, t^{h}, k^{h}/ and six affricates /ts, tʂ, tɕ, ts^{h}, $tʂ^{h}$, $tɕ^{h}$/. Plosives and affricates that occur in the 18 sentences are explained in Appendix 2. We labelled fine articulation types of 1701 plosives and affricates. Formant patterns and intensity in spectrogram were used as primary cues to identify boundaries. The occlusion phase is labelled as occ. Spike is typical in the transient phase (tra). Noise in high-frequency areas is the main cue to identify the frication phase (fri). Aspiration is usually identified in lower frequency areas (asp). Transition is labelled only if the formant pattern of the following vowel can be identified (trans). When more than one phase is merged, _ is used to denote concatenated phases.

Results

Vowel space

Figure 1 summarizes the relative position of three corner vowels produced by all 45 hearing-impaired children. The space determined by the corner vowels varies to a great extent among the children. Some of the children produced clear contrasts in both F1 and F2 with a large vowel space. Theoretically, the larger the vowel space is, the clearer the vowel contrast will be. In other words, the acoustic properties of vowels will reflect their intelligibility in regard to articulation. By applying the maximal contrast approach (i.e. drawing a quadrangle connecting the points that form the maximal differences in F1 and F2 values), we observed that 21 out of the 45 hearing-impaired produced /a/ outside of the quadrangle. This deviation of F2 values of /a/ from the normal range means the articulation of /a/ is either in a further front

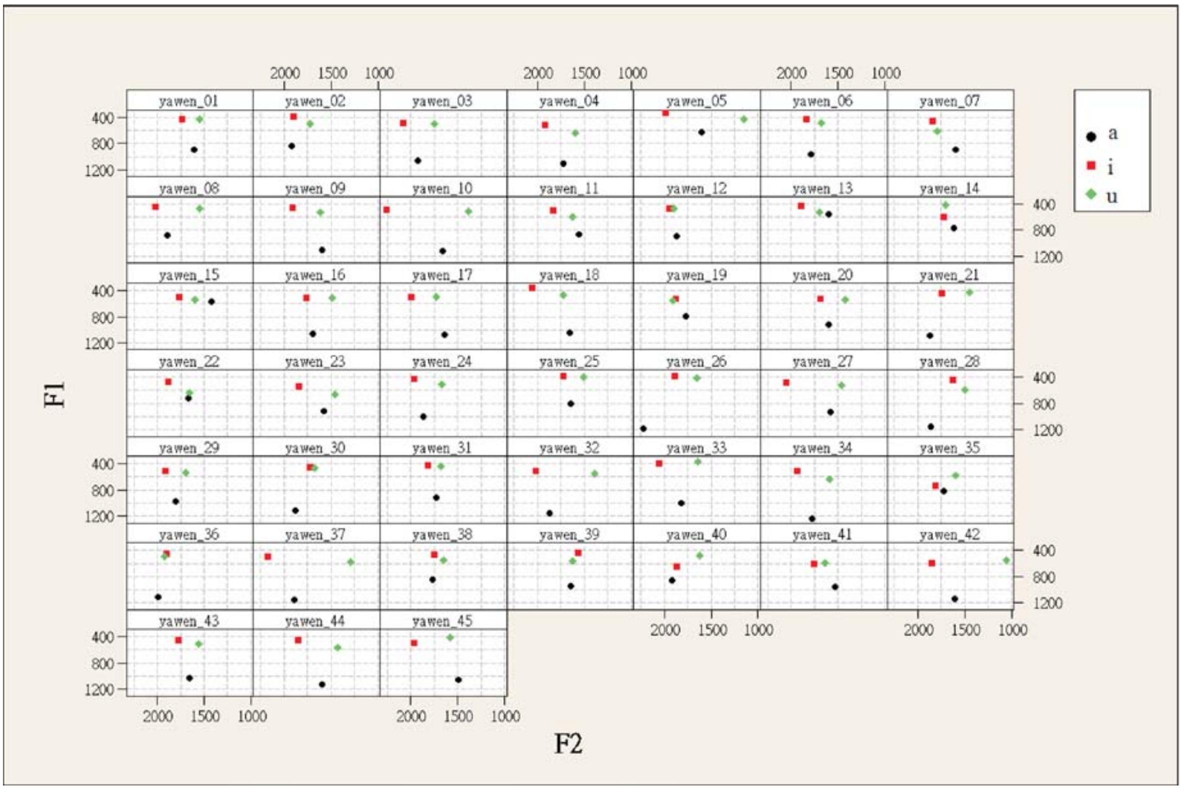

Figure 1. Vowel space produced by hearing-impaired children.
Note: *Yawen* is the abbreviation for the Children's Hearing Foundation.

tongue position than /i/ or in a further back tongue position than /u/. Eight children had more severe problems in producing /i/ and /u/; their F2 difference is small with minimal contrast. Four of them even produced /i/ and /u/ in a reversed order. And vowel space shrinkage can be clearly seen. In particular, the F2 values show that the articulation related to the tongue position results in a more severe problem than that to the tongue height. Therefore, its vowel space shows a small long shape. In Taiwan Mandarin, there is no low back vowel. Although 21 children had their /a/ outside the quadrangle, the deviation of /a/ should not hinder its discrimination in speech communication. However, if the F2 difference is not clear enough, it will be difficult to differentiate /i/ from /u/.

Speech intelligibility

The speech intelligibility of the children was assessed by adopting a 5-scale system: 1–5 to represent the poorest, poor, ordinary, good and clearest. Three trained linguists scored from 1 to 5 for each child, making the minimum score 3 and the maximum score 15. The Pearson correlation coefficients among the three linguists were all above 0.8, $p < 0.01$.

Vowel space grouping

One of our main attempts was to find appropriate threshold levels to categorize the production skills of vowels based on their acoustic properties. We found that children whose vowel space size was below 1 had a relatively poor pronunciation. In contrast, the pronunciation of those whose vowel space size was above 3 was very clear. Thus, we divided the children into three subgroups according to the threshold levels: above 3, between 1 and 3 and below 1. Also taking speech intelligibility into account, we calculated the mean value of speech intelligibility scores of each subgroup. By dividing vowel space size (separated by 1 and 3 on the y-axis) and their respective mean of speech intelligibility scores (separated by 8 and 11 on the x-axis) into subgroups, we observed a positive correlation trend between vowel space size and speech intelligibility, as shown in Figure 2.

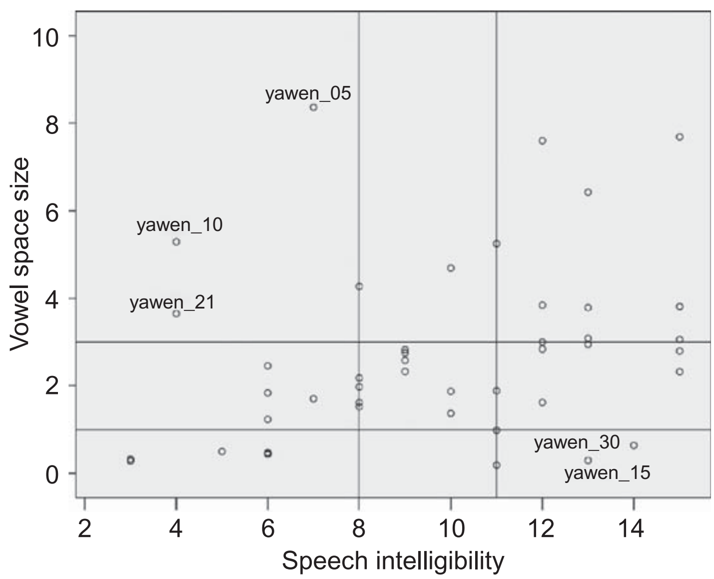

Figure 2. Scatterplot of vowel space size and speech intelligibility.

Five particular cases

Five particular cases, yawen_05, yawen_10, yawen_21, yawen_15 and yawen_30, are located in the upper-left and in the bottom-right blocks in Figure 2. The first three cases had a large vowel space with a low score in speech intelligibility; the last two cases were given high intelligibility scores with a small vowel space. Vowels produced by yawen_15 and yawen_30 were not clear; however, it did not affect the clarity of the sentence meaning, as the prosody was quite fluent. In contrast, vowels produced by yawen_05, yawen_10 and yawen_21 were clear with a nice contrast. But the tone and intonation were monotonous and flat and it was hard to discriminate some of the consonants. These apparently influence speech clarity. These five cases suggest that in addition to vowel production, factors such as prosody and consonants also need to be considered for the relationship between acoustic properties and speech intelligibility.

Vowel space and formant pattern in each vowel group

All 45 children were categorized into three groups according to their vowel space size: (1) Group I ($N = 15$): above 3; (2) Group II ($N = 20$): between 1 and 3; and (3) Group III ($N = 10$): below 1. For comparison, we recorded a normally hearing 5-year-old girl's data in the same scenario set-up. Figure 3(a) illustrates her vowel space with the area calculated for the quadrangle defined above with the representative formants of /i, a, u/ in red lines. The blue triangle represents the conventional vowel space. The children whose vowel space size falls in Group I tended to speak quite clearly. Figure 3(b) shows that their vowel space is relatively large when compared with the other two groups of children. Children grouped into Group II vary to a great degree, making up half of the total number of children. Differently, the children in Group III have mostly compressed vowel space, resulting in a long shape, as shown in Figure 3(d). Also shown in Figure 3 are their formant patterns in red dots. The main difference with the formants lies in the values of F2. The formant patterns produced by the normal-hearing child are similar to those represented in Lin and Wang (1995) for adults: (1) F1 and F2 fall widely apart for /i/; (2) F1 and F2 fall together in the medium frequency area for /a/; and (3) F1 and F2 fall together in the low-frequency area. By comparing the formant patterns in Figure 3(b)–3(d), we found that the F2 values of /u/ deviate from those of the normal-hearing child in Figure 3(a) to different extents. In particular, the contrast of F1 and F2 in Group III is only minimally represented.

Types of plosives and affricates

The labelling results of plosives and affricates were summarized into three types: (1) Type I – no clear boundaries within the plosives/affricates except for the transition phase; (2) Type II – except for the boundary between the frication and the aspiration phases, the rest of the phases can be clearly distinguished; (3) other types – all five phases including the transition phase were merged into one phase and other types of occurrences. Please note that we did not find a single case in our data in which all five articulation phases were clearly separable. Further investigations are necessary to find out if this is typical of hearing-impaired children or if it is a common process of children's phonological development. Figure 4(a) illustrates an aspirated bilabial plosive labelled as Type I, in which a clear formant pattern on the end phase of the consonant is identified, and only one pattern is observed in the spectrogram before the transition part. Type II represents the clearest pronunciation of plosives and affricates, as shown in Figure 4(b). In some extreme cases, as shown in Figure 4(c), we cannot even identify the transition part. The entire consonant is acoustically presented in terms of one pattern.

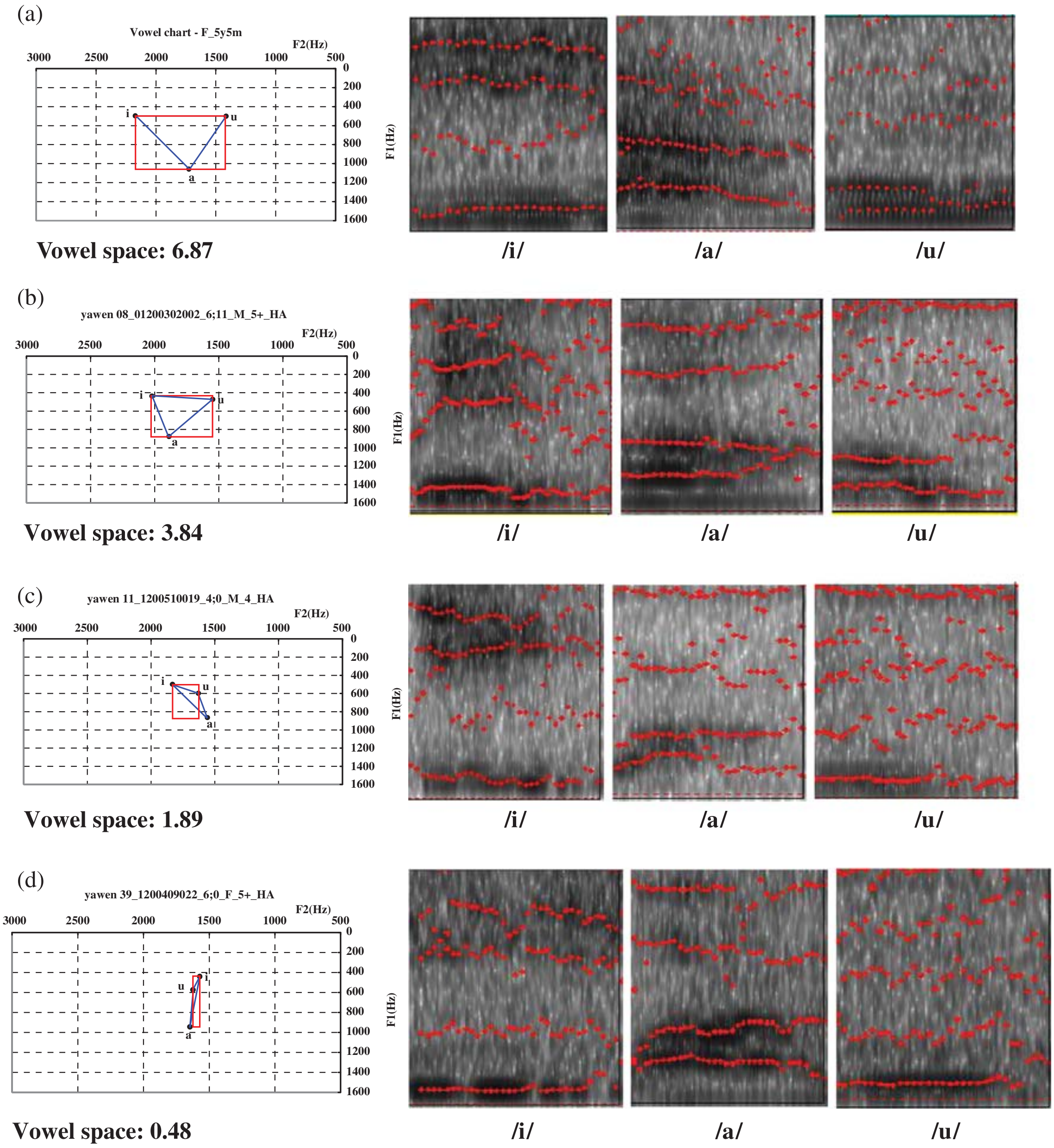

Figure 3. Vowel space and formant patterns. (a) From a normal-hearing child. (b) From a subject in Group I. (c) From a subject in Group II. (d) From a subject in Group III.

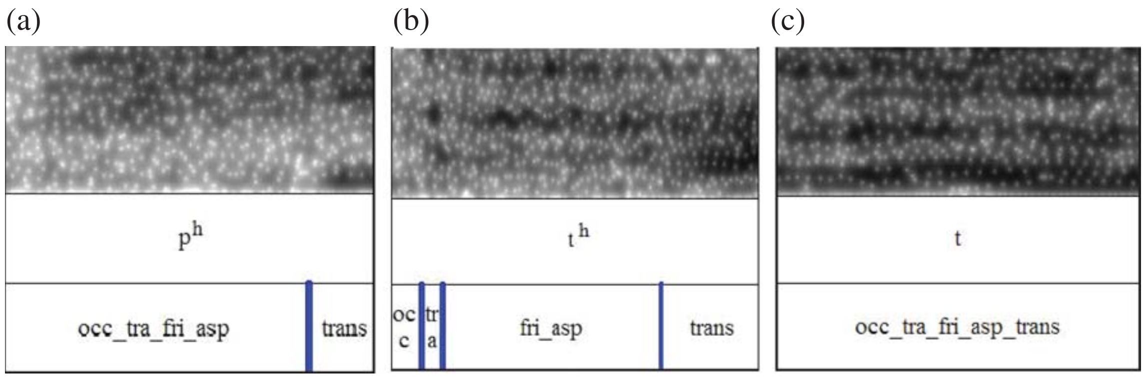

Figure 4. Labelled types of plosives and affricates. (a) Type I. (b) Type II. (c) Others.

Table I. Distribution of plosives and affricates types.

	Group I (%)	Group II (%)	Group III (%)
Type I	78.04	85.02	88.65
Type II (with spike)	12.16	7.15	4.60
Other types	9.80	7.82	6.75

As demonstrated in Table I, Group I children whose vowel space was the largest produced the clearest articulation type (Type II) the most frequently. And Group III with the poorest vowel contrast produced Type II the least frequently. That is, those whose acoustic contrast of vowels is better produce plosives/affricates in a more precise way in terms of the articulation types. Based on these results, we defined the spike percentage as the percentage of Type II produced by each child and used it as a factor in the statistical analysis.

Analysis

Linguistic and non-linguistic factors

For the linguistic factors, the level of auditory memory is the maximum number of content words in sentences a child can understand, adopted as a perceptual measure for representing the word comprehension ability of our subjects. As we only recruited children whose level of auditory memory was above 4, the level ranges from 4,5 to 5+. Also adopted from the CHF (2010) assessment method is sentence complexity. Sentence complexity is categorized into three levels to indicate how complex the sentences the children produce can be. Sentence complexity 1 means that the children used to produce simple and short sentences; 2 the children can produce long sentences without connectives, that is, no relative clauses are used; and 3 the children can produce long sentences with connectives. Similar to the level of auditory memory as a perceptual measure, sentence complexity is adopted as a measure for speech production. The non-linguistic factors include degree of hearing loss (1 stands for mild, 2 for moderate, 3 for moderate-severe, 4 for severe, 5 for severe-profound, 6 for profound), age of the children at test (in years), duration of receiving AVT (in years), age of receiving sensory aids (HA or CI; in years) and size of vowel space. The assessment of hearing-loss degrees was based on pure tone average thresholds, conducted by the CHF (2010).

Correlation

We examined the relationship of the above linguistic and non-linguistic factors with speech intelligibility, vowel space size and spike percentage by conducting the Pearson correlation test.

As shown in Table II, the degree of hearing loss is negatively correlated with speech intelligibility scores. The less severe the hearing loss of the children is, the clearer their speech will be. The size of vowel space is weakly correlated with speech intelligibility. Nevertheless, the relationship between vowel space grouping and speech intelligibility is clear, as illustrated in Figure 5. Group I performs the best in the impressionistic scores, with Group III performing the worst.

Table II also shows that the level of auditory memory is correlated with sentence complexity, speech intelligibility and age of receiving sensory aids. Sentence complexity is correlated with speech intelligibility and age at test. In contrast to speech intelligibility, vowel space size is only correlated with age at test at a statistically significant level. Figures 6 and 7 illustrate the

Table II. Correlation results.

	Vowel space size	Speech intelligibility	Level of auditory memory	Sentence complexity	Spike percentage	Degree of hearing loss	Age at test	Duration of receiving auditory-verbal therapy
Speech intelligibility	0.29, $p = 0.053$							
Level of auditory memory	0.095, $p = 0.535$	0.636, $p = 0$						
Sentence complexity	0.215, $p = 0.156$	0.396, $p = 0.007$	0.57, $p = 0$					
Spike percentage	0.065, $p = 0.67$	0.224, $p = 0.139$	0.177, $p = 0.245$	0.245, $p = 0.105$				
Degree of hearing loss	0.002, $p = 0.991$	−0.34, $p = 0.022$	−0.166, $p = 0.276$	−0.098, $p = 0.523$	−0.100, $p = 0.513$			
Age at test	0.315, $p = 0.035$	0.183, $p = 0.229$	0.263, $p = 0.081$	0.284, $p = 0.059$	0.131, $p = 0.392$	0.093, $p = 0.544$		
Duration of receiving auditory-verbal therapy	0.187, $p = 0.22$	−0.134, $p = 0.379$	−0.055, $p = 0.72$	0.004, $p = 0.977$	−0.042, $p = 0.782$	0.458, $p = 0.002$	0.652, $p = 0$	
Age of receiving sensory aids	0.173, $p = 0.255$	0.404, $p = 0.006$	0.32, $p = 0.032$	0.227, $p = 0.134$	0.263, $p = 0.08$	−0.434, $p = 0.003$	0.361, $p = 0.015$	−0.392 $p = 0.008$

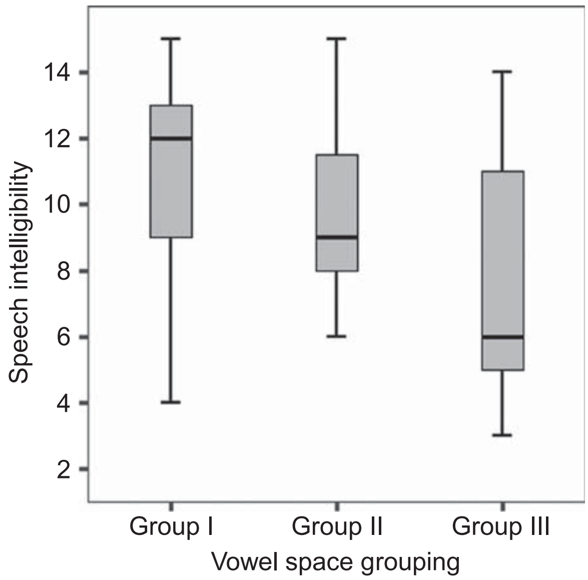

Figure 5. Speech intelligibility of vowel space groups.

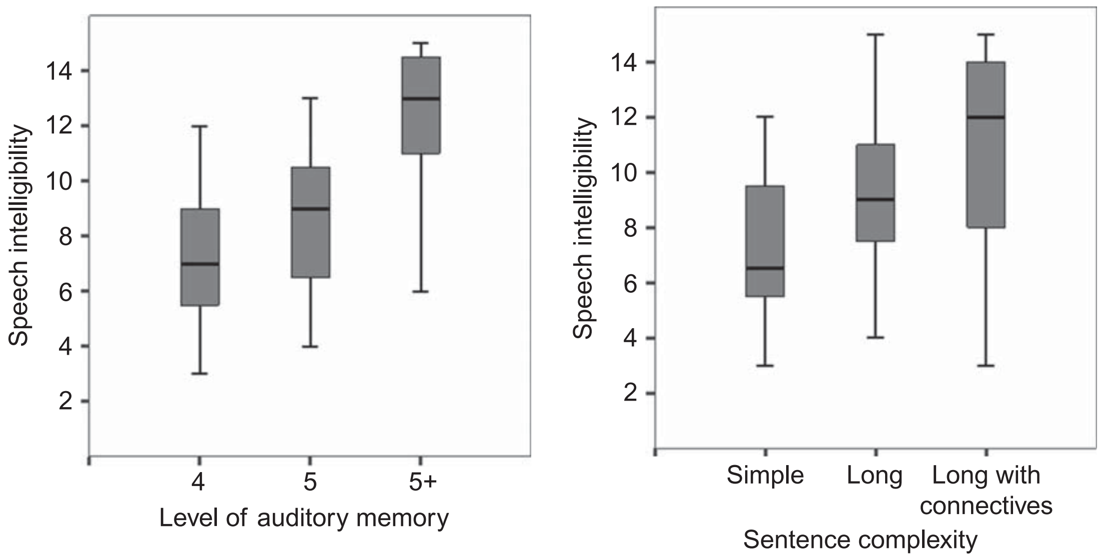

Figure 6. Correlation of speech intelligibility and speech abilities.

differences between speech intelligibility and vowel space size in the correlation with the perceptual and production measures: auditory memory level and sentence complexity, in box plots showing the median, the first and third quartiles and extreme cases with subject abbreviations in the statistics.

General linear model

Three generalized linear models (GLMs) were conducted to statistically verify the contribution of multiple factors on these three dependent variables: speech intelligibility, vowel space and spike percentage. Predictors are gender, hearing aids, level of auditory memory, degree of hearing loss, sentence complexity, duration of receiving AVT and age of receiving sensory aids. Coding of those categorical variables are the same as previously mentioned.

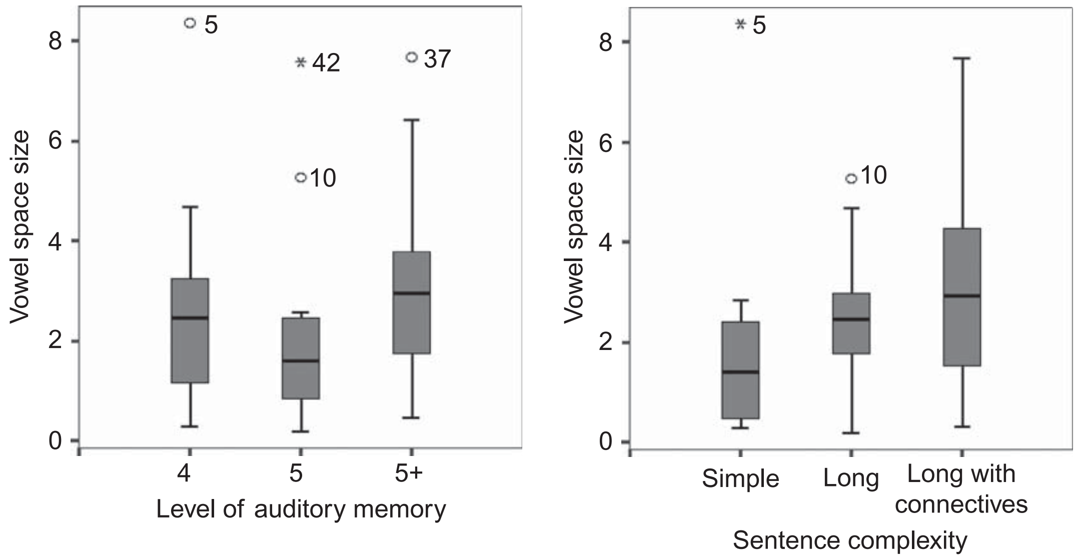

Figure 7. Correlation of vowel space size and speech abilities.

Table III. Results of multiple factor analysis.

Source	DF	Seq SS	Adjusted SS	Adjusted MS	*F*	*P*
Analysis of variance for vowel space size, using adjusted SS for tests						
Gender	1	4.026	7.542	7.542	1.87	0.18
Level of auditory memory	1	1.676	1.108	1.108	0.27	0.603
Hearing aids	1	0.647	2.209	2.209	0.55	0.464
Degree of hearing loss	1	0.843	0.713	0.713	0.18	0.676
Sentence complexity	1	6.276	4.13	4.13	1.02	0.318
Duration of receiving auditory-verbal therapy	1	7.102	12.336	12.336	3.06	0.089
Age of receiving sensory aids	1	12.385	12.385	12.385	3.07	0.088
Error	37	149.167	149.167	4.032		
Total	44	182.122				
Analysis of variance for speech intelligibility, using adjusted SS for tests						
Gender	1	1.381	1.892	1.892	0.27	0.609
Level of auditory memory	1	217.108	99.541	99.541	14.04	0.001
Hearing aids	1	0.062	16.573	16.573	2.34	0.135
Degree of hearing loss	1	48.427	31.462	31.462	4.44	0.042
Sentence complexity	1	1.259	0.777	0.777	0.11	0.742
Duration of receiving auditory-verbal therapy	1	0.024	0.657	0.657	0.09	0.762
Age of receiving sensory aids	1	6.049	6.049	6.049	0.85	0.362
Error	37	262.268	262.268	7.088		
Total	44	536.578				

Note: Seq SS, sequential sums of squares.

Although some of the factors were correlated with each other in Table II, they may not have a significant contribution to predicting the above variables when multiple factors are considered. The GLM results in Table III show that none of the predictors has a significant contribution to spike percentage. The duration of receiving AVT and age of receiving sensory aids are the two most significant predictors to vowel space size. For predicting speech intelligibility, the level of auditory memory and degree of hearing loss are the most significant contributing variables.

Table IV. Correlation results of vowel space grouping.

Vowel space grouping	
Spike percentage	$-0.362, p = 0.015$
Speech intelligibility	$-0.331, p = 0.026$
Level of auditory memory	$-0.194, p = 0.202$
Sentence complexity	$-0.327, p = 0.028$

Vowel space grouping

As this study also attempts to find simple thresholds to categorize the level of speech production abilities of hearing-impaired children, we tested whether the grouping of vowel space defined above is correlated with spike percentage, speech intelligibility, level of auditory memory and sentence complexity. Although vowel space size is correlated only with speech intelligibility in Table II, Table IV shows that there is a significant correlation between vowel space grouping and spike percentage, speech intelligibility and sentence complexity. These three factors are all related to speech production abilities, suggesting that vowel space grouping may be a preliminary, simple metric for diagnosing the speech production ability of hearing-impaired children. In contrast, the factor that is related to speech perception ability – the level of auditory memory – is not correlated with vowel space grouping.

Discussion

Acoustic properties of vowels are not equivalent to speech intelligibility

The empirical evidence provided in this article suggests that the acoustic properties of vowels are related to the speech intelligibility of hearing-impaired children. In spite of the heterogeneous subject group, vowel space size measured by the maximal contrast in F1 and F2 values normalized across subjects was correlated with speech intelligibility at a near significant level. This result implies that acoustic properties reflect upon the associated impressionistic perception. Furthermore, speech intelligibility was correlated with word comprehension level and sentence complexity, but vowel space size was not, as shown in Figures 6 and 7. In the analysis of GLM, the degree of hearing loss had a contributing role in predicting the speech intelligibility score, but not in the case of vowel space size. Apparently, the speech intelligibility score and vowel space size are two dependent, but different measures for describing language ability. In addition to vowels, acoustic properties of other linguistic factors should be taken into consideration to properly represent speech intelligibility as illustrated by the five particular cases in Figure 2. Nevertheless, the suggested vowel space thresholds for grouping acoustic properties of vowels have proven to be a useful assessment index, as this grouping was correlated with speech intelligibility, sentence complexity and spike production that reflect the speech production abilities of the subjects.

What is the role of consonant production in using speech?

We did not find any token of plosives or affricates, in which all five articulation phases can be clearly separated in our data when taking consonant production into account. The finest articulation type, whose boundary between the frication and the aspiration was merged, makes up around 10% of the labelled plosives/affricates. Statistically, vowel space grouping was correlated with spike production. The best vowel space group produced spike most frequently, the poorest group the least often. But spike production was not correlated with any of the linguistic or non-linguistic factors. This implies that consonant production is a measurement for assessing speech production performance, but the role it plays in assessing language ability may not be as important as vowels.

What about lexical tones?

Lexical tones are one of the most important prosodic features of Mandarin. They determine the form and meaning of the associated syllables. A number of studies examined how Mandarin tones are produced and perceived by hearing-impaired children (Barry, Blamey, and Martin, 2002; Lee, van Hasselt, Chiu, and Cheung, 2002; Wu and Yang, 2003; Peng, Tomblin, Cheung, Lin, and Wang, 2004b; Han, Zhou, Li, Chen, Zhao, and Xu, 2007). The pattern of tone perception after cochlear implantation may not fully follow that of normal children. Barry et al. (2002) concluded that average pitch height and pitch direction were more salient cues than pitch level while being perceived. Currently, we are using the same speech data to cluster the contour types by F0-stylization and model each tone shape by two optimized slopes. The clustering result will be used to investigate whether tone production is correlated with the speech abilities we have studied in this article.

Conclusion

This study has shown that there was a weak correlation between vowel space size and speech intelligibility. Vowel space size increased with age and duration of AVT training, and high speech intelligibility was associated with good perception and production abilities, as illustrated by the level of auditory memory and sentence complexity. These results support the notion that vowel space size and speech intelligibility are two different but dependent measures for speech production ability. Although consonant production (spike production) was correlated with vowel space grouping, it was not significantly associated with vowel space size or any other factor investigated in the GLM analysis. This evidence helps us to better understand how the speech information available to hearing-impaired children influences speech acquisition. Consequently, concrete treatment and linguistic training could be designed accordingly to strengthen the identified weaknesses. For further studies, linguistic factors should include prosody such as lexical tones in addition to vowels and consonants. Data types should include more authentic, spontaneous speech data such as story telling and conversation to analyse word use and communication skills.

Acknowledgements

We are grateful to the most useful comments including suggestions for statistical analysis provided by three anonymous reviewers of *Clinical Linguistics & Phonetics*. The study presented in this article is supported by the Children's Hearing Foundation and the National Science Council under Grant NSC 98-2410-H-001-066. We thank the teachers and assistants of the Children's Hearing Foundation for helping us record the speech data.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Barry, J. G., Blamey, P. J., & Martin, L. F. A. (2002). A multidimensional scaling analysis of tone discrimination ability in Cantonese-speaking children using a cochlear implant. *Clinical Linguistics & Phonetics*, *16*, 101–113.

Blamey, P., Barry, J., Bow, C., Sarant, J., Paatsch, L., & Wales, R. (2001a). The development of speech production following cochlear implantation. *Clinical Linguistics & Phonetics*, *15*, 363–382.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001b). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech and Hearing Research*, *44*, 264–285.

Bouchard, M. G., Normand, M., & Cohen, H. (2007). Production of consonants by prelinguistically deaf children with cochlear implants. *Clinical Linguistics & Phonetics*, *21*, 875–884.

Children's Hearing Foundation (2010). Accessed 13 March 2011 from: http://www.chfn.org.tw/en/index_e.htm

Chuang, C. (2006). Vowel formant characteristics, speaking rate, and speech intelligibility of Mandarin in prelingually deafened elementary students with cochlear implants (PhD Dissertation). National Taipei College of Nursing (in Chinese).

Ciocca, V., Francis, A. L., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *Journal of the Acoustical Society of America*, *111*, 2250–2256.

Dornan, D., Hickson, L., Murdoch, B., & Houston, T. (2007). Outcomes of an auditory-verbal program for children with hearing loss: A comparative study with a matched group of children with normal hearing. *The Volta Review*, *107*, 37–54.

Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., & Xu, L. (2007). Tone production of Mandarin Chinese speaking children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, *71*, 875–880.

Ho, D. (1996). *Concept and method in phonology*. Taipei: Daah Publisher (in Chinese).

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. San Diego, CA: Singular.

Khouw, E., & Ciocca, V. (2007). An acoustic and perceptual study of initial stops produced by profoundly hearing impaired adolescents. *Clinical Linguistics & Phonetics*, *21*, 13–27.

Lee, K. Y., van Hasselt, C. A., Chiu, S. N., & Cheung, D. M. C. (2002). Cantonese tone perception ability of cochlear implant children in comparison with normal-hearing children. *International Journal of Pediatric Otorhinolaryngology*, *63*, 137–147.

Lee, P. (1999). A study on acoustic characteristics of Mandarin affricates of hearing-impaired children. *Special Education and Rehabilitation*, *7*, 79–122.

Lee, P. (2001). Intelligibility of aspirated and unaspirated words by hearing-impaired students. *Special Education and Rehabilitation*, *9*, 133–158.

Liberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Liker, M., Mildner, V., & Šindija, B. (2007). Acoustic analysis of the speech of children with cochlear implants: A longitudinal study. *Clinical Linguistics & Phonetics*, *21*, 1–11.

Lin, T., & Wang, L. (1995). *Yuyinxue jiaocheng*. (Course in phonetics). Beijing: Beijing University Press (in Chinese).

Ling, D. (1976). *Speech and the hearing-impaired child: Theory and practice*. Washington, DC: Alexander Graham Bell Association for the Deaf and Hard of Hearing.

Maassen, B., & Povel, D. (1985). The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. *Journal of the Acoustical Society of America*, *78*, 877–886.

Mildner, V., & Liker, M. (2008). Fricatives, affricates, and vowels in Croatian children with cochlear implants. *Clinical Linguistics & Phonetics*, *22*, 845–856.

Mildner, V., Šindija, B., & Zerinski, K. V. (2006). Speech perception of children with cochlear implants and children with traditional hearing aids. *Clinical Linguistics & Phonetics*, *20*, 219–229.

Musselman, C., & Kircaali-Iftar, G. (1996). The development of spoken language in deaf children: Explaining the unexplained variance. *Journal of Deaf Studies and Deaf Education*, *1*(2), 108–121.

Peng, S., Spencer, L. J., & Tomblin, J. B. (2004a). Speech intelligibility of pediatric cochlear implant recipients with 7 years of device experience. *Journal of Speech, Language, and Hearing Research*, *47*, 1227–1236.

Peng, S., Tomblin, J. B., Cheung, H., Lin, Y., & Wang, L. (2004b). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing*, *25*, 251–264.

Peng, S., Weiss, A. L., Cheung, H., & Lin, Y. (2004c). Consonant production and language skills in Mandarin-speaking children with cochlear implants. *Archives of Otolaryngology Head & Neck Surgery*, *130*, 92–97.

Pickett, J. M. (1999). *The acoustics of speech communication. Fundamentals, speech perception theory, and technology*. Boston, MA: Allyn and Bacon.

Serry, T. A., & Blamey, P. J. (1999). A 4-year investigation into phonetic inventory development in young cochlear implant users. *Journal of Speech and Hearing Research*, *42*, 141–154.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*, 708–722.

Tye-Murray, N., Spencer, L., & Woodworth, G. G. (1995). Acquisition of speech by children who have prolonged cochlear implant experience. *Journal of Speech and Hearing Research*, *38*, 327–337.

Wu, J., & Yang, H. (2003). Speech perception of Mandarin Chinese speaking young children after cochlear implant use: Effect of age at implantation. *International Journal of Pediatric Otorhinolaryngology*, *67*, 247–253.

Appendix 1.

Subject	Sex	Age at test	Level of auditory memory	Sensory aid	Duration of receiving auditory-verbal therapy	Age of receiving sensory aids	Degree of hearing loss	Sentence complexity	Vowel space size	Vowel space grouping	Speech intelligibility	% of spike
yawen_01	F	5;5	5	HA	3;8	1;9	Severe-profound	Long	2.58	2	9	5.26
yawen_02	F	3;7	5+	CI	2;6	0;7	Profound	Long	1.87	2	10	19.23
yawen_03	F	4;11	5	HA	0;1	4;10	Moderate-severe	Long	2.33	2	9	4.76
yawen_04	F	5;11	5+	HA	1;9	4;0	Moderate	Long/c	3.08	1	13	40.48
yawen_05	M	12;5	4	CI	10;1	1;11	Profound	Simple	8.36	1	7	8.33
yawen_06	F	4;3	5	HA	3;8	0;8	Profound	Long	1.70	2	7	7.14
yawen_07	M	4;0	4	CI	1;0	0;9	Profound	Long/c	0.32	3	3	2.86
yawen_08	M	6;11	5+	HA	5;1	1;10	Severe-profound	Long/c	3.84	1	12	23.81
yawen_09	F	5;2	4	HA	0;2	4;11	Moderate-severe	Long	4.69	1	10	16.67
yawen_10	M	5;4	5	HA	4;5	0;11	Profound	Long	5.29	1	4	2.56
yawen_11	M	4;0	4	HA	2;2	1;7	Moderate-severe	Long	1.89	2	11	15.38
yawen_12	F	6;1	4	HA	5;9	0;3	Profound	Simple	0.45	3	6	0.00
yawen_13	F	5;0	4	HA	4;7	0;6	Profound	Simple	0.50	3	5	0.00
yawen_14	M	5;0	4	CI	4;7	0;6	Profound	Simple	0.28	3	3	2.86
yawen_15	F	4;11	5	HA	4;4	0;7	Moderate	Long	0.30	3	13	0.00
yawen_16	F	6;9	5+	CI	5;6	0;6	Profound	Long/c	3.81	1	15	7.32
yawen_17	M	4;11	5+	HA	2;11	1;10	Moderate-severe	Long/c	2.32	2	15	4.76
yawen_18	M	8;7	5+	CI	7;2	1;3	Profound	Long/c	2.80	2	15	21.43
yawen_19	M	5;1	5	CI	2;11	2;2	Severe	Long	0.44	3	6	0.00
yawen_20	F	3;3	4	HA	2;1	1;0	Moderate	Simple	1.84	2	6	4.76
yawen_21	M	5;9	4	CI	3;9	2;0	Profound	Long	3.65	1	4	17.65
yawen_22	M	4;6	4	CI	3;8	0;7	Profound	Simple	1.98	2	8	2.56

(*Continued*)

Appendix 1. (Continued)

Subject	Sex	Age at test	Level of auditory memory	Sensory aid	Duration of receiving auditory-verbal therapy	Age of receiving sensory aids	Degree of hearing loss	Sentence complexity	Vowel space size	Vowel space grouping	Speech intelligibility	% of spike
yawen_23	F	3;9	4	HA	2;7	1;0	Mild	Long	2.76	2	9	12.82
yawen_24	M	4;11	4	HA	2;4	2;5	Severe	Long	2.45	2	6	7.69
yawen_25	M	7;3	5+	HA	2;5	3;2	Mild	Long/c	1.62	2	12	10.00
yawen_26	M	5;3	4	HA	1;7	3;7	Moderate	Simple	2.84	2	12	0.00
yawen_27	F	11;2	4	CI	8;0	3;3	Profound	Long/c	4.27	1	8	16.67
yawen_28	F	7;0	5+	HA	4;7	2;5	Moderate-severe	Long	2.18	2	8	9.38
yawen_29	F	4;2	4	HA	1;5	2;10	Severe-profound	Long	2.83	2	9	3.13
yawen_30	F	9;2	5+	CI	2;11	5;1	Profound	Long/c	0.64	3	14	7.32
yawen_31	F	6;0	5	CI	2;11	3;1	Profound	Long	1.37	2	10	19.44
yawen_32	M	4;5	5+	HA	0;1	2;0	Moderate-severe	Long/c	6.42	1	13	5.13
yawen_33	F	10;5	5+	HA	7;2	2;8	Severe	Long/c	5.25	1	11	2.56
yawen_34	M	6;0	5+	HA	1;7	3;10	Mild	Long	2.95	2	13	5.00
yawen_35	M	6;0	5+	CI	2;6	2;0	Profound	Simple	0.98	3	11	0.00
yawen_36	M	5;7	5	CI	4;2	1;4	Severe-profound	Long	0.19	3	11	19.44
yawen_37	M	7;1	5+	HA	3;11	3;1	Severe	Long/c	7.69	1	15	0.00
yawen_38	M	7;1	5+	HA	4;11	2;2	Severe-profound	Long/c	1.52	2	8	2.50
yawen_39	F	6;0	5+	HA	3;5	2;7	Severe-profound	Long/c	0.48	3	6	7.50
yawen_40	M	6;0	5	HA	4;10	1;2	Severe	Long/c	1.62	2	8	2.50
yawen_41	M	9;5	5	HA	3;10	2;6	Moderate	Long/c	1.23	2	6	15.15
yawen_42	M	4;1	5	CI	3;0	0;11	Profound	Long/c	7.60	1	12	7.14
yawen_43	F	7;11	5+	HA	2;9	5;3	Moderate	Long	3.06	1	15	14.29
yawen_44	F	8;0	5+	HA	1;7	4;8	Moderate-severe	Long/c	3.79	1	13	0.00
yawen_45	M	6;4	5+	HA	0;4	5;0	Moderate-severe	Long	3.00	1	12	25.93

Note: Long/c, long sentences with connectives; HA, hearing aid; CI, cochlear implant.

Appendix 2.

Sentence	Translation
Wo ke yi kan dian shi	I can watch TV
Wo men ke yi kan dian shi	We can watch TV
Qing ni ba shu gei wo	Please pass me the book (recipient singular)
Qing ni men ba shu gei wo	Please pass me the book (recipients plural)
Ta xi huan dan gao	He likes cakes
Ta men xi huan tiao wu	They love dancing
San ge da ren	Three adults
Yi ge xiao hai	A child
Yi zhi xiao ji	A chicken
Yi he cai se bi	A box of colouring pencils
Bu hui lai	(Someone) won't come
Wo hui xie zi	I can write
Wo hui zi ji chi fan	I can eat by myself
Hei se de gou hen piao liang	Black dogs are beautiful
Xi bu xi huan kong long	Do you like dinosaurs or not?
Ta xi huan huang se	He likes yellow
Ta xi huan huang se de dan gao	He likes yellow cakes
Bu ke yi lin yu	Don't stay in rain

Notes: The above sentences are transcribed in Hanyu Pinyin, which is a transcription convention widely used in China and in academic papers in Taiwan. Plosives /p, t, k, p^h, t^h, k^h/ and affricates /ts, tʂ, tɕ, ts^h, $tʂ^h$, $tɕ^h$/ are transcribed as *b, d, g, p, t, k* and *z, zh, j, c, ch, q*, respectively.