Published in final edited form as:

Int J Pediatr Otorhinolaryngol. 2017 June ; 97: 143–149. doi:10.1016/j.ijporl.2017.04.012.

Mandarin compound vowels produced by prelingually deafened children with cochlear implants

Jing Yang¹ and Li Xu²

¹Communication Sciences and Disorders, University of Central Arkansas, Conway, AK 72035, USA

²Communication Sciences and Disorders, Ohio University, Athens, OH 45701, USA

Abstract

Objective—Compound vowels including diphthongs and triphthongs have complex, dynamic spectral features. The production of compound vowels by children with cochlear implants (CIs) has not been studied previously. The present study examined the dynamic features of compound vowels in native Mandarin-speaking children with CIs.

Methods—Fourteen prelingually deafened children with CIs (aged 2.9 – 8.3 years old) and 14 age-matched, normal-hearing (NH) children produced monosyllables containing six Mandarin compound vowels (i.e., /aI/, /aŮ/, /uo/, /ie/, /iaŮ/, /ioŮ/). The frequency values of the first two formants were measured at nine equidistant time points over the course of the vowel duration. All formant frequency values were normalized and then used to calculate vowel trajectory length and overall spectral rate of change.

Results—The results revealed that the CI children produced significantly longer durations for all six compound vowels. The CI children's ability to produce formant movement for the compound vowels varied considerably. Some CI children produced relatively static formant trajectories for certain diphthongs whereas others produced certain vowels with greater formant movement than did the NH children. As a group, the CI children roughly followed the NH children on the pattern of magnitude of formant movement, but they showed a slower rate of formant change than the NH children.

Conclusions—The findings suggested that prelingually deafened children with CIs, during the early stage of speech acquisition, had not established appropriate targets and articulatory coordination for compound vowel productions. This preliminary study may shed light on rehabilitation of prelingually deafened children with CIs.

Keywords

cochlear implants; prelingual; compound vowels; production; acoustic analysis; Mandarin Chinese

* Author to whom correspondence should be addressed: Li Xu, M.D., Ph.D., Communication , Sciences and Disorders, Ohio University, Grover Center – W229, Athens, OH 45701, Tel.: +1 740 593 0310, xul@ohio.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Cochlear implants (CIs) have provided unprecedented auditory sensation to patients with severe to profound hearing loss. While numerous studies have examined speech perception of the CI users, only a few studies have focused on their speech production, in particular, development of vowel and consonant inventory [1-7]. A number of acoustic studies were undertaken to examine the basic acoustic features of selected monophthongal vowels in CI children [8-15]. There was a consensus regarding the facilitating role of implantation on the acoustic development of vowel production in the hearing-impaired children. However, it remained controversial in whether the CI children could reach the age-matched norms in the detailed acoustic-phonetic characteristics [8,14,15]. Unlike monophthongs that involve one articulatory position, diphthongs have two distinct articulatory targets connected by a marked transition within one syllable [16]. Unlike two adjacent vowels belonging to two different syllables, diphthongs involve a smooth continuous change of the position of the tongue body and a distinct rate of formant movement [17]. Therefore, production of diphthongs requires smooth vowel-to-vowel transition within one single syllable, which involves complex motor planning and coordination of articulatory gestures.

Auditory input plays an important role in the development of speech motor control [18,19]. Previous studies have shown that the absence of auditory feedback in deaf speakers causes a distortion of speech production [20,21]. Compared to deafened speakers, CI recipients regain partial auditory sensation that helps them better regulate speech motor processes. However, compared to normal-hearing (NH) speakers, CI users still lack a full range of auditory information. Some recent studies have suggested that children with CIs who have no developmental or cognitive impairment still show delays in motor development, especially in complex motor [22] and fine motor skills [23]. The delayed motor skills in CI children may cause difficulty in compound vowel production that requires more complex planning than single vowel production. Therefore, CI children's phonetic-acoustic features might be different from those of NH children. However, little research has been done to examine the production of compound vowels in children with CIs.

The present study aimed to expand our knowledge of speech development in CI children by comparing the acoustic features of selected Mandarin diphthongs and triphthongs in CI children with those in age-matched NH children. Mandarin differs from English on both segmental and suprasegmental levels. Regarding the vowel systems, Mandarin has a smaller inventory of monophthongal vowel phonemes than English. However, Mandarin contains a large inventory of compound vowels that includes nine diphthongs (/aI/, /eI/, /aŮ/, /oŮ/, /ua/, /uo/, /ia/, /ie/, /ye/) and four triphthongs (/iaŮ/, /ioŮ/, /uaI/, /ueI/). So far, a number of studies have documented the vowel acquisition in normally developing Mandarin-speaking children [24-27]. But no study has investigated the production of compound vowels in Mandarin-speaking children with CIs.

Although there is a shortage of comparable studies on compound vowel production in CI children, some early research on the acoustic features of compound vowels in the deaf population were of relevance to the present study [20,28]. Monsen [20] examined the acoustic properties of the vowels /i, a, ə/ and the diphthong /ai/ produced by 36 deaf

adolescents and 4 NH adolescents. The authors found a noticeable reduction of the phonological space due to the reduced formant frequency values of the first two formants, especially in F2 in deaf adolescents. Plant and Hammarberg [28] compared vowel features of connected discourse in native Swedish speakers with an acquired profound hearing loss with those in NH speakers. They found elongated vowel duration and the absence of diphthongization in the deafened speakers. More recently, Palethorpe *et al.* [29] compared vowel duration, formant position, and formant trajectories of 11 monophthongs and 7 diphthongs in Australian English between the postlingually deaf and NH speakers. The analysis of the diphthong production revealed that the deaf speakers produced significantly longer vowel durations than did the NH speakers. In addition, less formant movement was also observed in diphthongs produced by the deaf speakers. The relative positions of the first target of the diphthongs showed a retraction of F2 and a reduction of F1 similar to the changes of F1 and F2 in corresponding monophthongs.

These studies suggested that hearing-impaired speakers tend to produce vowels, especially compound vowels, with less formant change in a more restricted frequency range, which was attributable to the lack of auditory input. In the present study, we investigated how prelingually deafened children who started obtaining language experience after implantation would produce compound vowels. Two research questions were specifically addressed: first, whether the prelingually deafened children with CIs showed reduced formant movements for their diphthongs and triphthongs relative to the age-matched NH children, and second, to what extent the CI children's acoustic features of compound vowels were similar to or different from those of the age-matched NH children.

Methods

Participants

The participants included 14 prelingually deafened Mandarin speaking children with CIs aged between 2.9 and 8.3 years old and 14 age-matched NH children. All participants were originally recruited for a study of lexical tone production in Mandarin-speaking children with CIs [30]. They were raised in Mandarin-speaking families in the Beijing area. All of the CI children were nonverbal prior to implantation and they all received post-surgery rehabilitation services that addressed basic speech and language skills at a professional rehabilitation center in Beijing. Detailed demographic information is presented in Table 1. However, the relevant auditory information including the etiology of hearing loss and auditory threshold was not available. The 14 age-matched NH children were reported as having no impairments of language, speech, or hearing by their parents or teachers.

Speech materials

The speech materials included a list of nine Mandarin monosyllables (“ai, bao, pao, duo, tuo, jie, qie, yao, you” in Pinyin) containing four diphthongs (/aɪ/, /aʊ/, /uo/, and /ie/) and two triphthongs (/iaʊ/ and /ioʊ/). The target vowels either followed obstruent consonants or occurred in an environment without initial consonant. None of the recorded monosyllables had a coda. The onglides of the compound vowels covered the three corner vowels /a, i, u/ in

Mandarin Chinese. Each syllable was produced in all four Mandarin tones one single time. Therefore, each participant produced $9 \times 4 = 36$ tokens.

Recording

For each participant, the recording session was conducted in a quiet room using a Sony portable DAT recorder (Model TCD-D100) connected to an Electro Voice omnidirectional microphone (Model RE50B). The sampling rate was set at 44.1 kHz and the quantization rate was set at 16 bits. The experimenter produced each syllable in Mandarin tone 1 and required the participants to articulate the same syllable in all four tones. Each syllable in each tone was produced once and the experimenter made no corrections during the recording session.

Data analysis

All recorded tokens were transferred to a computer hard disk drive. The recorded tokens were then segmented into separate wave files with each file containing one syllable. To capture the nature of the dynamic change of vowel quality in diphthongs and triphthongs, formant frequency values were extracted at nine equidistant time locations (10-20-30-40-50-60-70-80-90% point) over the course of vowel duration using the spectrographic analysis program TF32 [31]. The landmarks of vowel onset and offset were located by hand through a visual check of the waveform accompanied with spectrographic display. In particular, vowel onset was defined at the beginning of vowel periodicity and vowel offset was defined at the end of vowel periodicity where both F1 and F2 were present.

Due to the relatively large age range of the child speakers, all measured formant frequency values were normalized prior to further analysis. Given that only selected compound vowels were investigated in the present study, a vowel-intrinsic approach of the Bark transformation was used and all formant values of each token were converted to the Bark scale following the formula of Traunmüller [32]:

$$Z_i = 26.81 / (1 + 1960 / F_i) - 0.53,$$

$Z_i = 26.81 / (1 + 1960 / F_i) - 0.53$, in which F_i is the formant frequency value of a given formant i and Z_i is the Bark value of formant i . Based on the bark values at multiple measurement points, we calculated trajectory length (TL) and spectral rate of change (TL_roc) [33]. Trajectory length defined the sum of Euclidean distance between each two consecutive time points (10-20%, 20-30%, 30-40%, etc.), which was calculated by:

$$TL = \sum_{n=1}^9 VSL_n.$$

The Euclidean distance of each vowel section (VSL) was calculated by:

$$VSL_n = \sqrt{(F1_{n+1} - F1_n)^2 + (F2_{n+1} - F2_n)^2}.$$

TL_{roc} defined the average rate of formant change over a certain period of time. We calculated the average spectral rate of change over the tested vowel portion using the formula:

$$TL_{\text{roc}} = TL / (0.8d),$$

where d is the duration of the vowel. The acoustic measures of TL and TL_{roc} enabled us to quantify the dynamic feature of formant trajectories in terms of the magnitude and the velocity of articulatory movement during the course of vowel production.

Statistical analysis

Preliminary analysis showed no apparent differences on formant trajectories among the four tones. In addition, for a given vowel that was produced in more than one word, the preceding consonants shared the same place of articulation. Therefore, the data were collapsed across tones and words. Subject mean values of each acoustic measurement were obtained prior to further statistical analyses. For the comparison of formant frequency values, a three-way repeated-measures ANOVA was used on the F1 and F2 data respectively with vowel quality and measurement point as the within-subject factors and group as the between-subject factor. For the comparison of other acoustic measures (e.g., duration, TL, and roc), two-way repeated measures ANOVA tests were implemented with vowel quality as the within-subject factor and group as the between-subject factor. Because the effects of vowel quality on formant frequency values and vowel spectral change were anticipated and have been reported in previous studies, which were not of interest in the present study, only significant results related to the group effect were reported and discussed. One-way ANOVAs were conducted afterwards to compare the group difference in each vowel. The alpha level was adjusted for multiple comparisons.

Results

Formant trajectories in F1 × F2 space

Figure 1 shows the comparison of the formant trajectories between individual CI children and the group mean data of the NH children. Considerable individual variation was present in formant trajectory patterns in the CI children. Some of the CI children, such as CI8 and CI9, showed similar patterns of formant trajectories to those of the NH children. In particular, these CI children approximated the NH children in terms of the relative position of the compound vowels in the acoustic space, the direction, and the magnitude of the formant movement. However, some CI children, such as CI5 and CI7, produced all six compound vowels in a restricted, centralized region of the vowel space and demonstrated a distinctive pattern of vowel organization from that of the NH children. Many CI children differed from the NH children in one or more aspects of the formant trajectory features. In

the top-left panel for the NH children, the two triphthongs /iaʊ/ and /ioʊ/ showed curved trajectories that indicated a change of vowel quality corresponding to three vowel targets. Some CI children, such as CI7 and CI14, showed less curved formant trajectories, which suggested that these CI children did not show as much change of vowel quality as the NH children did. Some CI children, such as CI3 and CI14, showed a higher position of /iaʊ/ than /ioʊ/ while some other CI children, such as CI6 and CI7, showed a formant trajectory moving from high back to mid-high back position for the vowel /ioʊ/. These patterns were dramatically different from the pattern of front-to-back formant movement in the NH children, which suggested that these CI children pronounced the triphthongal vowels with different articulatory movement patterns.

In terms of the formant trajectories of diphthongs, some CI children, such as CI4, CI6, and CI12, showed very little formant movement in the acoustic space for the vowel /aɪ/. CI5 produced this vowel with the trajectory moving from a central position to a high front position. CI7 produced /aɪ/ with a “U-shape” trajectory which was located at a position similar to his /aʊ/ in the acoustic space. For the vowel /ie/, some CI children, such as CI2 and CI14, produced it with a relatively static formant trajectory which indicated that this CI child monophthongized this diphthong. In contrast, CI12 produced /ie/ with a formant trajectory moving down to a low central position which indicated that this CI child mispronounced /ie/ as /ia/. In addition, the formant trajectory of /ie/ produced by CI7 showed a front-to-back movement which also substantially differed from the direction of formant trajectory in the NH children. For the other two diphthongs /aʊ/ and /uo/, some CI children, such as CI6, CI12, and CI13, produced /uo/ with the formant trajectories moving from a high back position to a low central position, which suggested that these CI children produced this vowel as /ua/ rather than a typical /uo/. For the vowel /aʊ/, some CI children, such as CI1, CI3, CI5, and CI6, showed little formant frequency change which indicates a static vowel quality rather than a diphthong moving from a low mid position to mid-high back position.

The three-way repeated-measures ANOVA revealed no significant main effect of group for either F1 or F2. However, a significant interaction effect of vowel by group was yielded for F1 ($F(5,130)=7.075, p<0.001$) and a significant interaction effect of measurement point by vowel by group was yielded for F2 ($F(40, 1040)=4.157, p=0.001$). These significant interaction effects suggested that the two groups of children showed different formant frequency values on certain vowels at certain measurement points. In addition to the comparison of relative position defined by F1 and F2 between the NH and CI children, a measure of overall acoustic distance was calculated to show the potential positional deviation of the CI children relative to NH targets. Following the procedure described in Yang et al. [15], for each vowel, the mean F1 and F2 were calculated for each measurement point in the NH children. Then, the Euclidean distance between each participant and the NH target was calculated for each measurement point. Finally, the Euclidean distances of the nine points were summed up, which served as an index to quantify the potential positional deviation of the formant trajectories of each participant from the NH targets.

As shown in Figure 2, the CI children produced all six compound vowels with greater acoustic distance from the NH targets. This observation was confirmed by the significant

group effect on the acoustic distance ($F(1, 26)=12.158, p=0.002$). This result suggested that the CI children, as a group, deviated from the NH targets in their vowel production to a greater extent than the NH children did. Note that the CI children demonstrated a great amount of individual variability in the vowel acoustic distances. The subsequent one-way ANOVAs revealed that the CI children showed significantly greater acoustic distance for the vowels /aɪ/, /uo/, and /ioʊ/. In addition, the NH children produced the vowels /aɪ/ and /au/ with larger variability than the other vowels, which indicated that the NH children did not always produce these vowels in a consistent manner.

Vowel duration

Figure 3 displays the duration of individual compound vowels. Not surprisingly, the two triphthongs showed much longer vowel durations than the diphthongs for both CI and NH children. The CI children generally matched the NH children on the pattern of vowel duration, but they produced significantly longer vowel durations than did the NH children for all six compound vowels. The two-way repeated measures ANOVA revealed a significant main effect of group ($F(1, 26)=14.877, p=0.001$). Subsequent one-way ANOVAs confirmed our observation of longer vowel duration in the CI children than in the NH children for each compound vowel. Note that the NH children presented a bimodal distribution of duration for certain vowels. A nonparametric test was conducted to validate the ANOVA results. The Mann-Whitney U test results yielded significant longer duration in the CI children than in the NH children for all six vowels.

Vowel trajectory length

Figure 4 displayed the trajectory lengths (TLs) over the 80% portion of vowel duration. Not surprisingly, the two triphthongs showed longer TLs than the four diphthongs in the NH children. In general, the CI children showed a pattern compatible with the NH children although the variation of the CI children was greater than that in the NH children for certain vowels such as /aʊ/, /iaʊ/, and /ioʊ/. A two-way mixed-factor repeated measures ANOVA was implemented with vowel as the within-subject factor and group as the between-subject factor. The results revealed a significant main effect of vowel quality ($F(5, 125)=101, p<0.001$). No group difference or vowel by group interaction was found.

Vowel spectral roc

While VSLs and TLs characterize the magnitude of formant change for each vowel section and the entire tested vowel portion (80% of vowel in the present study), it fails to capture the feature of the rate of formant change over time. Figure 5 shows overall spectral rocs (TL_rocs) over the 80% portion of vowel duration. The TL_rocs of the two triphthongs were larger than those of diphthongs in the NH children. This indicates that the articulators move faster for the triphthongs than for the diphthongs. The CI children generally followed the pattern of TL_roc over the six compound vowels in the NH children. However, they showed evidently smaller TL_rocs than did the NH children for most of the compound vowels. In addition, the CI children did not show significantly larger TL_rocs for the triphthongs than the diphthongs, as the NH children did. The two-way mixed-factor repeated measures ANOVA revealed a significant main effect of vowel by group interaction effect

($F_{5,130}=5.737$, $p<0.001$). Subsequent one-way ANOVAs revealed a significantly smaller TL_{roc} of /uo/, /iaŋ/, and /ioŋ/ in the CI children relative to the NH children.

Discussion

The present preliminary study examined the acoustic-phonetic features of six selected Mandarin diphthongs and triphthongs produced by Mandarin-speaking prelingually deafened children with CIs. On the basis of normalized formant values, we compared the shape of the formant trajectories, the relative position of each measurement point along the formant trajectories, and the magnitude and rate of formant frequency change between the CI and aged-matched NH children.

Regarding the question of whether CI children show reduced formant movements for the compound vowels, the results varied in different CI children and different vowels. Some CI children tended to reduce diphthongs to monophthongs for certain vowels. For example, CI3 and CI5 showed little formant movement for the vowel /aŋ/, CI6 and CI12 showed little formant movement for the vowel /aI/, and CI2 and CI14 showed little formant movement for the vowel /ie/ (see Fig. 1). However, this trend was not consistent for all analyzed vowels across all CI children. Some CI children such as CI8 and CI9 exhibited compatible patterns of formant trajectories to the NH children. Some CI children such as CI6, CI10 and CI12 produced /uo/, /aI/, and /aI/, respectively, with a greater magnitude of formant movement in comparison to the NH children. The remarkable inter-subject variation in vowel dynamics in the CI children was consistent with previous literature [34,35], which demonstrated large individual differences in performance of pediatric CI users. Further examination of CI children's demographic information and the vowel dynamic patterns suggested that the production of compound vowels in the CI children was affected by the factors of age at implantation and length of use. We observed that the CI children such as CI8 and CI9, who received implantation younger than two years old and have used the device for a relatively longer period of time, showed formant trajectories more similar to those of the NH children. In contrast, those children who received implantation at an older age such as CI5 and CI7, or only used the device for a short period of time such as CI14, showed distinctive patterns of vowel dynamics from that of the NH children. These findings provide additional evidence supporting the positive influence of early age of implantation and long implantation use on speech production of CI children [36-38].

Although some CI children produced similar formant trajectories to those of the NH children, the majority of the CI group, still differed from the NH children in certain aspects of acoustic features. As shown in Fig. 1, most of the CI children either undershoot the formant trajectories by producing some compound vowels with much less formant movement, or overshoot the formant trajectories by producing some compound vowels with greater formant movement. The relative positions of the compound vowels and the organization of the vowel space in the CI children were different from the NH targets (as shown in Fig.1 and 2). The off-target onset and offset of the compound vowels as well as the deviation of the formant trajectories in the CI children relative to the NH peers suggested that the CI children were less likely to initiate or complete the articulatory gestures in a

similar manner to the NH peers, which reflected more difficulty in establishing precise tongue configuration for the compound vowel targets in the CI children.

An early study pointed out that the hearing-impaired speakers produced vowels within a more restricted frequency range for both F1 and F2 [20]. Among the 14 CI children in the present study, CI5, CI7, and CI14 produced the compound vowels in a centralized area and presented substantial acoustic overlaps among these vowels. Note that CI5 and CI7 had implantation after 5 years of age and had relatively short length of device use. CI14 had received implantation for less than 4 months. The highly clustered production of compound vowels in these participants suggested that these participants were still at the emergent stage of speech development and performed similarly to deaf children reported in [20]. In a more recent study [29], Palethorpe and colleagues reported retraction of F2 for diphthong targets in some deafened speakers. In the present study, some CI children produced certain vowels with retracted F2. However, we also observed fronted F2 in some vowels of the CI children. For example, CI2 and CI5 produced the vowel /uo/ with a fronted F2. CI5, CI6, CI7, and CI13 produced the vowel /au/ with a fronted F2. These inconsistent patterns of positional deviation in the CI children relative to the NH children revealed diverse articulatory mechanisms used by CI children.

Consistent with previous studies [15,39], the present study showed that the CI children produced all compound vowels with longer duration than the NH peers, which suggested that the CI children may need more time to form the articulatory gestures and to travel from one target to the other. Earlier research on NH speakers reported that lengthened vowel duration and greater formant frequency change corresponded to higher speech clarity [40]. The CI children might use elongated vowel durations as a compensatory strategy to increase the clarity and intelligibility of their speech. The similar overall trajectory lengths in the CI children relative to the NH children suggested that the CI children were able to move the articulators with similar magnitude of movement as the NH children did. However, the significantly lower TL_rocs in the CI children demonstrated that the CI children did not move the articulators as quickly as the NH children. Moreover, the result of significantly longer TL_rocs for both triphthongs in the CI children relative to the NH children suggested the difficulty of articulatory coordination for multi-target phonetic segment in the CI children.

The findings of the present study, although preliminary, bear important implications in clinical training and habilitation for children with CIs. Currently, diverse strategies and techniques have been proposed to improve CI user's perceptual and recognition ability in a variety of listening conditions. But, we still lack an effective evidence based oral training program that improves the articulatory accuracy for pediatric CI users, especially for non-English speaking CI children. The present study demonstrated both undershooting and overshooting of the vowel targets during the process of compound vowel production in CI speakers. To address these deficits, effort should be made to identify the specific pattern of vowel organization in individual CI children and then develop a more targeted plan to help the CI children establish appropriate vowel targets for the compound vowels. Moreover, the present study revealed that not all compound vowels showed equally poor performance in the CI children. Some CI children showed more difficulty in producing certain vowels than

other vowels. Triphthongs were generally more challenging than the diphthongs for most CI children. Therefore, an overall impressionistic and acoustic evaluation of the production of individual speech sounds is needed for individual CI children.

These results are informative and might help guide rehabilitation effort in prelingually deafened children with CIs. However, the results should be interpreted with caution due to the large individual variation of the CI population and small sample size of our participants. The overshooting and undershooting of formant trajectories and the slower rate of change in our CI children did not occur ubiquitously. The NH children also presented great variability in certain measures (e.g., duration). In the future research, the sample size of participants in both CI and NH groups needs to be expanded so that we can better group the children based on their developmental stages. Specifically, CI children with a wider range of age at implantation and varying length of device use should be included to examine how these factors affect the general shape of formant trajectories and other types of fine-grained acoustic features. In addition to recruiting NH children as the control group, we are interested in comparing CI children's production with NH adults to better understand the deviation of CI children's vowel production from the adult norms. Moreover, a perception test of CI children's vowel production by NH adult listeners should also be carried out to examine how the acoustic deviation of CI children's vowel production will affect the recognition of the intended vowels by the NH listeners.

Conclusions

In general, the findings of the present study suggested that prelingually deafened children with CIs, during the early stage of speech acquisition, had not established appropriate targets and demonstrated more difficulty of articulatory coordination for compound vowel productions.

Acknowledgments

Emily Hahn and Lauren Muscari provided editorial assistance during the preparation of the manuscript. The study was supported in part by an NIH NIDCD Grant No. R15-DC014587.

References

1. Blamey PJ, Barry JG, Jacq P. Phonetic inventory development in young cochlear implant users 6 years post-operation. *J Speech Lang Hear Res.* 2001; 44(1):73–79. [PubMed: 11218111]
2. McCaffrey HA, Davis BL, MacNeilage PF, von Hapsburg D. Multichannel cochlear implantation and the organization of early speech. *Volta Rev.* 1999; 101(1):5–28.
3. Ertmer DJ, Mellon IA. Beginning to talk at 20 months: Early vocal development in a young cochlear implant recipient. *J Speech Lang Hear Res.* 2001; 44(1):192–206. [PubMed: 11218103]
4. Warner-Czyz AD, Davis BL. The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants Int.* 2008; 9(3):143–166. [PubMed: 18925564]
5. Serry TA, Blamey PJ. A 4-year investigation into phonetic inventory development in young cochlear implant users. *J Speech Lang Hear Res.* 1999; 42:141–154. [PubMed: 10025550]
6. Tye-Murray N, Kirk KI. Vowel and diphthong production by young users of cochlear implants and the relationship between the phonetic level evaluation and spontaneous speech. *J Speech Hear Res.* 1993; 36:488–502. [PubMed: 8331906]

7. Warner-Czyz AD, Davis BL, Morrison HM. Production accuracy in a young cochlear implant recipient. *Volta Rev.* 2005; 105:151–173.
8. Chuang HF, Yang CC, Chi LY, Weismer G, Wang YT. Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant. *Int J Speech-Lang PA.* 2012; 14(2):119–129.
9. Manning WH, Moore JN, Dunham MJ, Lu FL, Domico E. Vowel production in prelinguistic child following cochlear implantation. *J Am Acad Audiol.* 1992; 3:16–21. [PubMed: 1571581]
10. Ertmer DJ. Emergence of a vowel system in a young cochlear implant recipient. *J Speech Lang Hear Res.* 2001; 44:803–813. [PubMed: 11521773]
11. Löfqvist A, Sahlen B, Ibertsson T. Vowel space in Swedish adolescents with cochlear implants. *J Acoust Soc Am.* 2010; 128:3064–3069. [PubMed: 21110601]
12. Neumeyer V, Harrington J, Draxler C. An acoustic analysis of the vowel space in young and old cochlear-implant speakers. *Clin Linguist Phon.* 2010; 24:734–741. [PubMed: 20645857]
13. Uchanski RM, Geers AE. Acoustic characteristics of the speech of young cochlear implant users: a comparison with normal-hearing age-mates. *Ear Hear.* 2003; 24:90S–105S. [PubMed: 12612484]
14. Horga D, Liker M. Voice and pronunciation of cochlear implant speakers. *Clin Linguist Phon.* 2006; 20:211–217. [PubMed: 16428239]
15. Yang J, Brown E, Fox RA, Xu L. Acoustic properties of vowel production in prelingually deafened Mandarin-speaking children with cochlear implants. *J Acoust Soc Am.* 2015; 138(5):2791–2799. [PubMed: 26627755]
16. Papakyrtsis, I., Granese, A. Cross linguistic study of vowel systems. In: Ball, MJ., Gibbon, FE., editors. *Handbook of Vowels and Vowel Disorders.* Psychology Press; New York: 2013. p. 186–206.
17. Kent RD, Moll KL. Tongue body articulation during vowel and diphthong gestures. *Folia Phoniatr Logop.* 1972; 24(4):278–300.
18. Fairbanks G. Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem. *J Speech Hear Disord.* 1954; 19:133–139. [PubMed: 13212816]
19. Perkell J, Matthies M, Lane H, Guenther F, Wilhelms-Tricarico R, Wozniak J, Guio P. Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Commun.* 1997; 22(2):227–250.
20. Monsen R. Normal and reduced phonological space: The production of English vowels by deaf adolescents. *J Phon.* 1976; 4:189–198.
21. Goffman L, Ertmer DJ, Erdle C. Changes in speech production in a child with a cochlear implant: Acoustic and kinematic evidence. *J Speech Lang Hear Res.* 2002; 45(5):891–901. [PubMed: 12381047]
22. Schlumberger E, Narbona J, Manrique M. Nonverbal development of children with deafness with and without cochlear implants. *Dev Med Child Neurol.* 2004; 46:599–606. [PubMed: 15344519]
23. Horn DL, Pisoni DB, Miyamoto RT. Divergence of fine and gross motor skills in prelingually deaf children: Implications for cochlear implantation. *Laryngoscope.* 2006; 116:1500–1506. [PubMed: 16885761]
24. Jeng, H. The 9th International Congress of Phonetic Sciences. Copenhagen: 1979. The acquisition of Chinese phonology in relation to Jakobson's laws of irreversible solidarity.
25. Hsu JH. A study of the stages of development and acquisition of Mandarin Chinese by children in Taiwan. Crane Publishing. 2003
26. Si Y. Putonghua ertong yuyin de ge'an yanjiu [The child's acquisition of Puthonghua Phonology: A case study], *Dangdai Yuyanxue* [Contemporary Linguistics]. 2006; 1:93–94.
27. Zhu H, Dodd B. The phonological acquisition of Putonghua (modern standard Chinese). *J Child Lang.* 2000; 27(1):3–42. [PubMed: 10740966]
28. Plant G, Hammarberg B. Acoustic and perceptual analysis of the speech of the deafened. *Speech Trans Lab Q Prog Stat Rep.* 1983; 2(3):85–107.
29. Palethorpe S, Watson CI, Barker R. Acoustic analysis of monophthong and diphthong production in acquired severe to profound hearing loss. *J Acoust Soc Am.* 2003; 114(2):1055–1068. [PubMed: 12942984]

30. Zhou N, Xu L. Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *J Acoust Soc Am*. 2008; 123:1653–1664. [PubMed: 18345853]
31. Milenkovic, P. University of Wisconsin; Madison, WI: 2017. TF32 software program. <http://userpages.chorus.net/cspeech/> [19.01.17]
32. Traunmüller H. Analytical expressions for the tonotopic sensory scale. *J Acoust Soc Am*. 1990; 88(1):97–100.
33. Fox RA, Jacewicz E. Cross-dialectal variation in formant dynamics of American English vowels. *J Acoust Soc Am*. 2009; 126:2603–2618. [PubMed: 19894839]
34. Pisoni DB, Cleary M, Geers AE, Tobey EA. Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *Volta Rev*. 1999; 101(3):111–164. [PubMed: 21666760]
35. Ertmer DJ, Goffman LA. Speech production accuracy and variability in young cochlear implant recipients: Comparisons with typically developing age-peers. *J Speech Lang Hear Res*. 2011; 54(1):177–189. [PubMed: 20689030]
36. Connor CM, Hieber S, Arts HA, Zwolan TA. Speech, vocabulary, and the education of children using cochlear implants: Oral or total communication? *J Speech Lang Hear Res*. 2000; 43(5): 1185–1204. [PubMed: 11063240]
37. Tobey EA, Geers AE, Brenner CB, Altuna D, Gabbert G. Factors associated with development of speech production skills in children implanted by age five. *Ear Hear*. 2003; 24:36S–45S. [PubMed: 12612479]
38. Montag JL, AuBuchon AM, Pisoni DB, Kronenberger WG. Speech intelligibility in deaf children after long-term cochlear implant use. *J Speech Lang Hear Res*. 2014; 57(6):2332–2343. [PubMed: 25260109]
39. Monsen RB. Durational aspects of vowel production in the speech of deaf children. *J Speech Lang Hear Res*. 1974; 17(3):386–398.
40. Tasko SM, Greilick K. Acoustic and articulatory features of diphthong production: A speech clarity study. *J Speech Lang Hear Res*. 2010; 53:84–99. [PubMed: 19948757]

Fig. 1.

Formant trajectories of selected Mandarin diphthongs and triphthongs from the NH and CI children. The top left panel represents the group mean data of the 14 NH children. Each of the following 14 panels shows data for an individual CI child. In each panel, the formant movement trajectories were plotted on the basis of normalized F1 and F2 values in barks at nine equidistant measurement points (i.e., 10-20-30-40-50-60-70-80-90%) over the course of vowel duration. The filled larger symbols represent the 90% point.

Figure 2.

Box plot showing the sum of acoustic distance across nine measurement points for individual vowels in the CI and NH children. The acoustic distance was defined as the Euclidean distance between a measurement point of a vowel trajectory in each CI or NH child to the corresponding measurement point of the group mean trajectory of the corresponding vowel in the NH children. Each dot in the boxes represents one participant. Some jitters along the abscissa were applied to the data for better visual representation. Each box shows horizontal lines at the lower quartile, median, and upper quartile values. The whiskers show 99.3% of the data and the data points out of the whiskers were outliers.

Figure 3.

Box plot of vowel duration for selected Mandarin diphthongs and triphthongs produced by the NH and CI children. Each data point represents one subject. Each box shows horizontal lines at the lower quartile, median, and upper quartile values. The whiskers show the range of the data. The data points out of the whiskers are outliers.

Figure 4.

Box plot of overall trajectory lengths (TLs) for selected Mandarin diphthongs and triphthongs produced by the NH and CI children. TL is defined as the sum of Euclidean distance between each two consecutive time points, which represents the magnitude of formant change. Each data point represents one subject. Each box shows horizontal lines at the lower quartile, median, and upper quartile values. The whiskers show the range of the data.

Figure 5.

Box plot of overall vowel spectral rate of change (TL_rocs) for selected Mandarin diphthongs and triphthongs produced by the NH and CI children. Spectral roc is defined as the average rate of formant change over the tested 80% portion of vowel duration, which represents the velocity of formant change. Each data point represents one subject. Each box shows horizontal lines at the lower quartile, median, and upper quartile values. The whiskers show the range of the data. The data points out of the whiskers are outliers.

Table 1

Demographic information of the CI children. The order of the subject number matches that in Fig. 1.

Subject number	Gender	Age at recording (years)	Age at implantation (years)	Duration of implant use (years)	CI device
1	F	7.53	5.19	2.34	Clarion CII
2	F	4.41	2.59	1.82	N24M
3	F	4.27	2.05	2.22	N24M
4	F	5.51	3.1	2.41	N24M
5	F	6.53	5.56	0.97	N24R
6	F	5.35	4.55	0.80	N24R
7	M	7.67	7.09	0.58	N24R
8	M	4.27	1.7	2.57	N24M
9	M	3.60	1.5	2.10	N24M
10	M	8.33	5.73	2.60	N24M
11	M	3.38	1.32	2.06	N24M
12	M	5.5	3.9	1.60	Clarion CII
13	M	2.93	1.16	1.77	Clarion CII
14	M	2.91	2.6	0.31	N24R