International Journal of Pediatric Otorhinolaryngology 97 (2017) 143–149

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: http://www.ijporlonline.com/

Mandarin compound vowels produced by prelingually deafened children with cochlear implants

Jing Yang [a], Li Xu [b,*]

[a] Communication Sciences and Disorders, University of Central Arkansas, Conway, AK 72035, USA
[b] Communication Sciences and Disorders, Ohio University, Athens, OH 45701, USA

ARTICLE INFO

Article history:
Received 22 February 2017
Received in revised form
4 April 2017
Accepted 6 April 2017
Available online 10 April 2017

Keywords:
Cochlear implants
Prelingual
Compound vowels
Production
Acoustic analysis
Mandarin Chinese

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., /aɪ/, /aʊ/, /uo/, /iɛ/, /iaʊ/, /ioʊ/). 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 did 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.

1. 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 whether or not 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-

* Corresponding author. Communication Sciences and Disorders, Ohio University, Grover Center – W229, Athens, OH 45701, USA.
E-mail address: xul@ohio.edu (L. Xu).

http://dx.doi.org/10.1016/j.ijporl.2017.04.012
0165-5876/

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 (/aɪ/, /eɪ/, /aʊ/, /oʊ/, /ua/, /uo/, /ia/, /iɛ/, /yɛ/) and four triphthongs (/iaʊ/, /ioʊ/, /uaɪ/, /ueɪ/). So far, a number of studies have documented the vowel acquisition in normally developing Mandarin-speaking children [24–27]. However, 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, ɔ/ 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 an elongated vowel duration and an 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 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.

2. Methods

2.1. 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.

2.2. 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 /iɛ/) 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 once in each of the four Mandarin tones. Therefore, each participant produced $9 \times 4 = 36$ tokens.

2.3. 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 ElectroVoice 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.

2.4. 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]:

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

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