International Journal of Pediatric Otorhinolaryngology 92 (2017) 146–150

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

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

Audio-visual speech perception in noise: Implanted children and young adults versus normal hearing peers

Riki Taitelbaum-Swead [a, b, *], Leah Fostick [a]

[a] Ariel University, Department of Communication Disorders, Israel
[b] Meuhedet Health Services, Israel

ARTICLE INFO

Article history:
Received 29 August 2016
Received in revised form
22 November 2016
Accepted 23 November 2016
Available online 25 November 2016

Keywords:
Cochlear implant
Audio visual speech perception
Children
Speech reading

ABSTRACT

Objective: The purpose of the current study was to evaluate auditory, visual and audiovisual speech perception abilities among two groups of cochlear implant (CI) users: prelingual children and long-term young adults, as compared to their normal hearing (NH) peers.
Methods: Prospective cohort study that included 50 participants, divided into two groups of CI (10 children and 10 adults), and two groups of normal hearing peers (15 participants each). Speech stimuli included monosyllabic meaningful and nonsense words in a signal to noise ratio of 0 dB. Speech stimuli were introduced via auditory, visual and audiovisual modalities.
Results: (1) CI children and adults show lower speech perception accuracy with background noise in audiovisual and auditory modalities, as compared to NH peers, but significantly higher visual speech perception scores. (2) CI children are superior to CI adults in speech perception in noise via auditory modality, but inferior in the visual one. Both CI children and CI adults had similar audiovisual integration.
Conclusions: The findings of the current study show that in spite of the fact that the CI children were implanted bilaterally, at a very young age, and using advanced technology, they still have difficulties in perceiving speech in adverse listening conditions even when adding the visual modality. This suggests that adding audiovisual training might be beneficial for this group by improving their audiovisual integration in difficult listening situations.

1. Introduction

Cochlear implants (CI) are now the standard of care for hearing rehabilitation of severe to profound hearing loss in children and adults [1,2]. A large number of studies have shown the success of cochlear implants in providing better sound accessibility and enabling better speech perception and production, among both children and adults [3–6]. This success, and the progress of technology, has resulted in an expansion in the candidacy criteria for CI, including a decrease in age at implantation (resulting in an increasing number of infants being implanted in the second half of their first year of life [7]), and more cases of bilateral implants [8]. The combination of the availability of good hearing devices available at a very young age, and, when needed, to both ears, produces improved accessibility of auditory information [3,9]. Some studies have even shown that many implanted children can achieve age-appropriate levels of speech perception, similar to their normal hearing peers, when tested under optimal conditions (e.g. with no background noise) [7]. However, under adverse listening conditions, such as the presence of noise, the performance of cochlear implanted individuals is still reduced [10,11].

The enormous range of environmental listening conditions during daily life communication places huge demands on the process of speech perception. In transit from speaker to listener, speech signals are often altered by background noise and other interfering signals. The performance of CI users under these conditions has been found to be deficient [5]. This deficiency is especially significant, since many implanted children are educated in mainstream settings where the signal to noise ratios (SNR's) in classrooms may be as high as −6 dB [12]. Such conditions can make it very difficult for CI users to concentrate and learn new academic material.

Previous studies have tried to quantify the performance of prelingual CI users under different listening conditions [5,6]. Most

* Corresponding author. Ariel University, Department of Communication Disorders, Kiryat HaMada 3, Ariel, 40700, Israel.
E-mail addresses: rikits@ariel.ac.il (R. Taitelbaum-Swead), leahfo@ariel.ac.il (L. Fostick).

http://dx.doi.org/10.1016/j.ijporl.2016.11.022

of these studies, however, focused on examining speech perception via the auditory modality. Yet, under natural conditions, the CI user often experiences face to face situations in which visual information is also available. Indeed, speech perception accuracy have been demonstrated to improve when speech reading is added to auditory input both among the normal hearing population [13–15] and CI users [16–19]. Only few studies tested both CI and NH participants on *audiovisual* speech perception. One study used background noise condition, but only for the normal hearing participants (to avoid ceiling effect) [16]. Therefore, no group comparison was made. Another study used under different presentation levels (speech detection and recognition thresholds) for CI and NH participants [19]. Identifying potential differences in audiovisual speech perception among these two populations would help shed light on how prelingual CI users employ each modality in perceiving speech, compared to NH participants. This was the aim of the current study.

When studying perceptual achievements of CI users, age has important implications. First, age of implantation has a huge effect on speech perception: those who were implanted at an earlier age perform better than those implanted at older age [4]. Second, the current age of the CI user needs to be considered. Guidelines for implantation have changed over the years with regard to the minimum age of implantation, amount of residual hearing required, and the option of bilateral implantation. These factors contribute to making the population of less recent implantees (long term CI users who are now young adults) inherently different from the more recent ones (children). A third implication of age is the sensory system maturation that occurs with time. We and others [13,20–23] have previously shown that each modality develops at a different rate; while auditory and audiovisual speech perception matures already by the age of 8–9, visual speech perception matures only after this age. Therefore, in the present study, we compare auditory, visual, and audiovisual speech perception between prelingual CI children and long-term CI user adults, and their NH peers, in order to determine how CI affect auditory, visual and audiovisual speech perception in different age groups.

2. Material and methods

2.1. Participants

The study included 50 individuals, divided into two groups of CI (children and long term young adult users), and two groups of normal hearing individuals. The adults were matched by age and the children by hearing age. All participants possessed Hebrew as a native language, normal or corrected visual ability, no reported developmental, cognitive or neurological problems, and normal speech and language abilities (based on parental report, in the case of children).

2.1.1. CI groups

The CI groups included 20 implanted participants: 10 children (CI children) and 10 young adults (CI adults). All implanted participants met the following inclusion criteria: (1) Onset of severe to profound hearing impairment before age 3; (2) Usage of hearing aids prior to implantation; (3) Mainstream education and oral communication; (4) At least 50% in monosyllabic open set test in quiet; (5) CI devices of Cochlear, Advanced Bionics, or Med-El.

The CI children's mean chronological age was 6.5 years (S.D. 0.9 years). All children used two implants. The mean age at first implantation was 16.3 months (S.D. 5.6 months), and the mean age of second implantation was 32.6 months (S.D. 19.6 months). Etiology of the hearing loss was genetic for seven of the CI children and unknown for three. The CI young adults' mean chronological age was 22.6 years (S.D. 2.05 years). Their mean age at first implantation was 9.1 years (S.D. 7.0); five of the participants used two implants, three participants used one implant with no hearing aid in the second ear, and two participants used an implant on one side and a hearing aid on the other side. The mean age at second implantation was 13.9 (S.D 7.1). Etiology of hearing loss was genetic for six of the CI adults and unknown for four.

2.1.2. Normal hearing groups

Normal hearing participants included 15 children aged 4–5 years and 15 young adults aged 20–30 years. All participants had normal hearing thresholds (pure-tone air-conduction thresholds less than 15 dB HL bilaterally at octave frequencies from 250 to 4000 Hz [24]).

2.2. Speech perception tests

Speech perception tests were monosyllabic meaningful and nonsense words, which include mainly acoustic information and minimal linguistic redundancy.

2.2.1. Meaningful words

Monosyllabic meaningful Hebrew AB lists (based on [25]) were used in the present study. This test includes 12 lists (narrated in a film, as described in the sub-section on Apparatus below), each consisting of ten monosyllabic words. Each list contains ten syllables in a consonant-vowel-consonant (CVC) pattern in which the 19 consonants of the Hebrew language appear either at the initial or the final position, and each of the five Hebrew vowels (/a/, /e/, /i/, /o/, /u/) appears twice.

2.2.2. Nonsense words

This test resembles the structure of the meaningful words test and also includes 12 lists of ten monosyllabic CVC syllables. However, in this test, the syllables are nonsensical but contain some phonological redundancy, in accord with Hebrew linguistic rules (for example, the consonants /b/ and /p/ never appear in the final position).

2.3. Apparatus

A female native Hebrew speaker with intelligible articulation and clear facial movements was filmed and recorded. The speaker looked directly into the camera, starting and ending each utterance with a natural face/closed mouth position. The speaker was recorded against a bright background in a quiet, well-lit recording studio. Her face appeared in full on the entire screen. The audiovisual recordings were digitized using Apple Final Cut Pro X software with 64-bit resolution.

The words were recorded in a studio using a SONTRONICS TCS-6 microphone and Samplitude classic 8.1 recording software. They were edited using the Sound Forge program, which digitized (16-bit) at a sampling rate of 44 kHz. Word level was normalized using the overall Root Mean Square (RMS). White noise generated by the Sound Forge program was added to the normalized words in a signal-to-noise ratio (SNR) of 0 dB. The noise was added to words in the audiovisual and auditory conditions, while in the visual condition the words were presented in quiet.

The words were presented using the Winamp Media Player 5.7 software, via S-Tech supra aural headphones for the NH groups, or via sound field for the CI groups. Words intensity was 70 dBSPL, as measured by a TA 1350A Sound Level Meter were used for the NH groups. Participant responses were recorded using a SONY ICD-PX312 recording device placed in close proximity to the participants. An inter-acoustic AD229B audiometer was used to screen