Hearing Research 316 (2014) 110–121

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Event-related potentials for better speech perception in noise by cochlear implant users

Takahiro Soshi [a], Satoko Hisanaga [a], Narihiro Kodama [b], Yori Kanekama [a], Yasuhiro Samejima [b], Eiji Yumoto [b], Kaoru Sekiyama [a,*]

[a] *Division of Cognitive Psychology, Faculty of Letters, Kumamoto University, 2-40-1, Kurokami, Chuo-ku, Kumamoto 860-8555, Japan*
[b] *Department of Otolaryngology-Head and Neck Surgery, Graduate School of Medicine, Kumamoto University, 1-1-1, Honjo, Chuo-ku, Kumamoto 860-0811, Japan*

A R T I C L E I N F O

Article history:
Received 2 December 2013
Received in revised form
25 July 2014
Accepted 5 August 2014
Available online 23 August 2014

A B S T R A C T

Speech perception in noise is still difficult for cochlear implant (CI) users even with many years of CI use. This study aimed to investigate neurophysiological and behavioral foundations for CI-dependent speech perception in noise. Seventeen post-lingual CI users and twelve age-matched normal hearing adults participated in two experiments. In Experiment 1, CI users' auditory-only word perception in noise (white noise, two-talker babble; at 10 dB SNR) degraded by about 15%, compared to that in quiet (48% accuracy). CI users' auditory-visual word perception was generally better than auditory-only perception. Auditory-visual word perception was degraded under information masking by the two-talker noise (69% accuracy), compared to that in quiet (77%). Such degradation was not observed for white noise (77%), suggesting that the overcoming of information masking is an important issue for CI users' speech perception improvement. In Experiment 2, event-related cortical potentials were recorded in an auditory oddball task in quiet and noise (white noise only). Similarly to the normal hearing participants, the CI users showed the mismatch negative response (MNR) to deviant speech in quiet, indicating automatic speech detection. In noise, the MNR disappeared in the CI users, and only the good CI performers (above 66% accuracy) showed P300 (P3) like the normal hearing participants. P3 amplitude in the CI users was positively correlated with speech perception scores. These results suggest that CI users' difficulty in speech perception in noise is associated with the lack of automatic speech detection indicated by the MNR. Successful performance in noise may begin with attended auditory processing indicated by P3.

1. Introduction

Nowadays, a CI is the most effective neural prosthesis for delivering auditory information to patients with profound deafness by bypassing the damaged inner ear and directly stimulating the auditory nerves (Zeng, 2004). With the use of a CI, post-lingual deaf patients rapidly improve speech perception within the first year of surgery (Hamzavi et al., 2003; Rouger et al., 2007; Ruffin et al., 2007). On the other hand, speech perception in noise is still difficult for CI users even after several years of device use (Tyler et al., 1995; Nelson et al., 2003; Nelson and Jin, 2004; Fu and Nogaki, 2005; Davidson et al., 2010). It is an immediate issue to be clarified as to what behavioral and neural foundations are responsible for speech perception in noise with CI use.

Neurophysiological studies have investigated the neural foundations for CI-dependent auditory performance in quiet, mainly using two event-related potentials (ERPs), that is, mismatch negativity (MMN) and P300 (P3) (Kaga et al., 1991; Kraus et al., 1993; Ponton and Don, 1995; Groenen et al., 2001).

The MMN is a negative ERP, appearing around 200 ms after stimulus onset, observed for deviant auditory stimuli compared with standard frequent stimuli (Näätänen et al., 1978; Kraus et al., 1992). The MMN may originate mainly from the superior and

Abbreviations: 2T, two-talker; AEP, auditory evoked potential; ANOVA, analysis of variance; AO, auditory-only; AV, auditory-visual; CI, cochlear implant; DF, deafness; EEG, electroencephalogram; EOG, electro-oculogram; ERP, event-related potential; LSD, least significant difference; MMN, mismatch negativity; MNR, mismatch negative response; N, noise; N1, N100; N2, N200; NH, normal hearing; P2, P200; P3, P300; RT, response time; SD, standard deviation; SNHL, sensorineural hearing loss; SNR, signal-to-noise ratio; SPL, sound pressure level; Q, quiet; WN, white noise

* Corresponding author. Tel.: +81 96 342 2845.
E-mail address: sekiyama@kumamoto-u.ac.jp (K. Sekiyama).

http://dx.doi.org/10.1016/j.heares.2014.08.001
0378-5955/

middle temporal areas (Marco-Pallarés et al., 2005; Näätänen et al., 2007) and reflects automatic auditory detection of deviant stimuli (Näätänen and Gaillard, 1983; Näätänen et al., 2007). Under attended conditions, MMN is overlapped by an attention-related posterior negativity (N2b) that peaks at around 250 ms (Näätänen and Gaillard, 1983; Novak et al., 1992; Cowan et al., 1993; Näätänen et al., 2007).

The MMN has been observed for good CI performers, but not for poor CI performers. Kraus et al. (1993) recorded the MMN response from good CI performers, using a passive auditory oddball task with speech. Similar findings about MMN elicitation for good CI performers have been reported in several studies (adult/speech: Groenen et al., 1996b; children/speech: Singh et al., 2004; adult/tone: Kelly et al., 2005; Zhang et al., 2011; Lonka et al., 2013).

P3 is another ERP component used in CI-related ERP studies. The P3 is the third positive component typically observed for attended rare targets in an active oddball task (Squires et al., 1975; Picton, 1992). Because P3 does not appear for an undetected change of stimulus properties, the elicitation is associated with an attentional evaluation of stimulus change (Donchin et al., 1978). The latency has a wide range from about 300 ms to over 600 ms after stimulus onset. The scalp distribution has a centro-posterior maximum.

P3 is also observed for good CI performers, but not for poor CI performers (Kaga et al., 1991; Oviatt and Kileny, 1991; Micco et al., 1995; Groenen et al., 1996a, 2001). Oviatt and Kileny (1991) observed that one poor CI performer could not detect stimulus change in an active oddball task, not showing the P3 to the deviant tone, while the other nine CI users could detect stimulus change, eliciting the P3.

In contrast to speech perception in quiet, very little is known about CI users' neurophysiological foundations for auditory speech perception in noise. The current study investigates the neurophysiological responses of CI users to auditory speech in noise. Participants were post-lingual adult CI users having at least 2 years of CI use, with NHs as controls. As with previous studies, we also used an auditory oddball paradigm with consonant-vowel syllables (/ba/and/ga/), comparing neurophysiological responses between deviant and non-deviant stimuli.

The main predictions of ERP results are as follows: present CI users having already used a CI device for more than 2 years, likely show good syllable detection in quiet (Hamzavi et al., 2003; Rouger et al., 2007; Ruffin et al., 2007). Accordingly, they will elicit the MMN and the N2b ('N2 deflection' noted together hereafter as 'mismatch negative response: MNR') (Näätänen and Gaillard, 1983) to deviant stimuli in quiet, similar to the NH controls (Groenen et al., 1996b). The P3 to deviant stimuli may not appear, because syllable detection in quiet may be easy for both groups; thus, the selective evaluation of deviant stimuli as a task-relevant rare target may be attenuated (Picton, 1992).

In noise, the CI users with good syllable detection performance and the NH controls may also show MNRs to deviant stimuli. They may also elicit the P3, because speech in noise probably promotes attentional stimulus evaluation (Wong et al., 2008), enhancing evaluation of deviant stimuli as a rare target. On the other hand, poor CI performers may elicit neither MNR nor P3, because degraded speech perception at a poor SNR did not elicit either MNR or P3 even for NH people (Martin et al., 1997; Whiting et al., 1998; Kaplan-Neeman et al., 2006).

We also behaviorally tested auditory-only (AO) and auditory-visual (AV) word perception in quiet and noise for the purpose of delineating an overview of noise effects on CI-dependent speech perception (Experiment 1). Experiment 1 used two types of noise (white noise (WN) and two-talker babble (2T)). Talker noise is suitable to examine noise interference effects to CI users' speech perception in ordinary communicative situations. A two-talker babble may work as not only an energetic masker such as white noise, but also as an information masker of the target speech (Brungart et al., 2001; Freyman et al., 2004; Nelson and Jin, 2004; Cooke et al., 2008; Mattys et al., 2009). As a result, the talker noise may more severely affect CI-dependent speech perception, providing the significant information that CI users are vulnerable in speech perception at two levels of noise masking. The present CI users may be weak in AO word perception in noise, in general (Nelson et al., 2003; Fu and Nogaki, 2005). In addition, the CI users' AV word perception is likely to be more degraded in the 2T noise condition than in the WN condition (Carhart et al., 1969; Brungart et al., 2001 for review of NHs' AO performances in two types of noise) because differences in AO noise interference may be enhanced in AV word perception in multiplicative ways, as suggested by a previous study (Sumby and Pollack, 1954). Therefore, Experiment 1 included not only AO, but also AV conditions. The results of Experiment 1 will be reported first.

2. Methods

2.1. Experiment 1: behavioral measure of word perception

2.1.1. Participants

Seventeen CI and twelve NH participants took part in the experiment. The CI users were post-lingually deafened (>90 dB hearing level at all test frequencies), and were monaurally implanted. Mean age of the CI users was 63.2 ± 10.6 years old (41–80 years old). Mean duration of CI use was 8.0 ± 5.5 years (2.4–19.7 years). Mean duration of deafness (DF) was 6.3 ± 7.1 years (0.3–24 years). The etiology included sudden sensorineural hearing loss (SNHL), idiopathic progressive SNHL, mitochondrial disease, sequelae of chronic otitis media, and mumps. Their primary communication method was oral, and none of them used a hearing aid on a non-implanted ear. The CI users used their individual standard comfortable device settings throughout the experiments. Table 1 summarizes the main demographical and clinical properties of the CI users.

The NH participants matched to the CI users in age (NH: 62.3 ± 9.0 years old, range from 43 to 76 years old; $t_{(27)} = 0.262$, $p = 0.796$), and male-to-female ratio (CI: female:male, 12:5; NH: 8:4; $\chi^2_{(1)} = 0.051$, $p = 0.822$). The NH participants possessed normal hearing ability (<25 dB of average hearing level at 500, 1000, 2000, and 4000 Hz as defined by the World Health Organization: left, 14.2 ± 5.0 dB; right: 15.0 ± 5.7 dB). All of the CI and NH participants had normal or corrected-to-normal visual acuity, and were right-handed. They reported no cortical and psychiatric deficits. All of the participants provided written informed consent prior to participation. All of the procedures were approved by the Human Subjects Ethics Committee of Kumamoto University.

2.1.2. Stimuli

The stimulus set consisted of 8 lists of 25 Japanese words (each word contained about 3 morae, e.g.,/ha-shi-ra/, "pilar"; /shi-ro-i/, "white"). These lists were from the CI 2004 list set (Technical Committee on Cochlear Implants in Japan, 2004).

The experimental conditions consisted of AO and AV conditions. In the AO condition, stimuli consisted of auditory speech and a visual fixation point (+). The visual cross was presented 900 ms before the onset of auditory stimuli and provided the cue for the duration of the auditory speech. In the AV condition, stimuli contained both auditory and visual speech. The visual speech used actual facial articulatory movements. These two conditions had sub-conditions of quiet and noise: in the quiet (Q) conditions (AO-Q, AV-Q), auditory words were presented without background noise. In the noise conditions, two types of noise were added to