



The P300 Auditory Event-Related Potential May Predict Segregation of Competing Speech by Bimodal Cochlear Implant Listeners

Duo-Duo Tao^{1*†}, Yun-Mei Zhang^{2†}, Hui Liu¹, Wen Zhang¹, Min Xu¹, John J. Galvin III³, Dan Zhang² and Ji-Sheng Liu^{2*}

¹ Department of Ear, Nose, and Throat, Shaanxi Provincial People's Hospital, Xi'an, China, ² Department of Ear, Nose, and Throat, The First Affiliated Hospital of Soochow University, Suzhou, China, ³ House Institute Foundation, Los Angeles, CA, United States

OPEN ACCESS

Edited by:

Xin Luo,
Arizona State University, United States

Reviewed by:

Yingjiu Nie,
James Madison University,
United States
Zilong Xie,
University of Kansas Medical Center,
United States

*Correspondence:

Duo-Duo Tao
entdtao@163.com
Ji-Sheng Liu
ljswwq@sina.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 03 March 2022

Accepted: 16 May 2022

Published: 10 June 2022

Citation:

Tao D-D, Zhang Y-M, Liu H, Zhang W,
Xu M, Galvin JJ III, Zhang D and
Liu J-S (2022) The P300 Auditory
Event-Related Potential May Predict
Segregation of Competing Speech by
Bimodal Cochlear Implant Listeners.
Front. Neurosci. 16:888596.
doi: 10.3389/fnins.2022.888596

Compared to normal-hearing (NH) listeners, cochlear implant (CI) listeners have greater difficulty segregating competing speech. Neurophysiological studies have largely investigated the neural foundations for CI listeners' speech recognition in quiet, mainly using the P300 component of event-related potentials (ERPs). P300 is closely related to cognitive processes involving auditory discrimination, selective attention, and working memory. Different from speech perception in quiet, little is known about the neurophysiological foundations for segregation of competing speech by CI listeners. In this study, ERPs were measured for a 1 vs. 2 kHz contrast in 11 Mandarin-speaking bimodal CI listeners and 11 NH listeners. Speech reception thresholds (SRTs) for a male target talker were measured in steady noise or with a male or female masker. Results showed that P300 amplitudes were significantly larger and latencies were significantly shorter for the NH than for the CI group. Similarly, SRTs were significantly better for the NH than for the CI group. Across all participants, P300 amplitude was significantly correlated with SRTs in steady noise ($r = -0.65$, $p = 0.001$) and with the competing male ($r = -0.62$, $p = 0.002$) and female maskers ($r = -0.60$, $p = 0.003$). Within the CI group, there was a significant correlation between P300 amplitude and SRTs with the male masker ($r = -0.78$, $p = 0.005$), which produced the most informational masking. The results suggest that P300 amplitude may be a clinically useful neural correlate of central auditory processing capabilities (e.g., susceptibility to informational masking) in bimodal CI patients.

Keywords: cochlear implant, competing speech, informational masking, event-related potentials, P300

INTRODUCTION

While cochlear implants (CIs) provide sufficient spectro-temporal resolution for speech recognition in quiet by deaf individuals, masked speech recognition is often difficult for CI users. Steady noise is thought to largely produce "energetic" masking; the spectro-temporal overlap between the target and masker occurs at the periphery (e.g., Brungart, 2001; Kidd et al., 2002). Competing speech is thought to produce some combination of energetic masking, "envelope"

masking (target and masker envelope interference even when there is no spectral overlap; e.g., Stone and Canavan, 2016), and “informational” masking (e.g., lexical interference, target/masker similarities, etc.; Brungart, 2001; Kidd et al., 2002, 2016). Different from normal-hearing (NH) listeners, who have greater difficulty with competing noise than with competing speech, CI listeners have greater difficulty with competing speech than with competing noise (e.g., Stickney et al., 2004; Cullington and Zeng, 2008; Tao et al., 2018). The coarse spectro-temporal resolution is thought to limit CI users’ segregation of target and masker speech (e.g., Friesen et al., 2001; Shannon et al., 2004; Fu and Nogaki, 2005; Luo and Fu, 2009).

Cortical measures have been used to characterize NH and CI listeners’ auditory processing. Auditory event-related potentials (ERPs) reflect the brain’s response to changes in an ongoing stimulus (e.g., deviant stimuli in the context of frequent stimuli in an oddball paradigm). Exogenous, pre-attentive responses (e.g., P1, N1, P2, N2 peaks) typically occur within the first 250 ms and do not reflect cognitive processing (e.g., Martin et al., 2008; Lightfoot, 2016). The latency of the endogenous P3 (or P300) response is typically between 250 and 400 ms, and is thought to reflect attention and/or arousal (e.g., Polich and Kok, 1995; Kok, 2001). Recording of P300 responses requires some sort of behavioral response to the deviant stimulus (e.g., counting the number of deviant stimuli during a test run, indicating when a deviant stimulus was heard, etc.). P300 latency has been shown to be related to the speed of information processing (e.g., Ritter et al., 1972; Kutas et al., 1977; Parasuraman and Beatty, 1980; Donchin and Coles, 1988). P300 amplitude has been shown to decrease with increasing task difficulty (e.g., Parasuraman and Beatty, 1980). Uncertainty in discrimination of sounds may be reflected in reduced P300 amplitude (e.g., Sutton et al., 1965; Hillyard et al., 1971; Squires et al., 1973; Picton, 2011).

There is great variability in CI outcomes that is largely unexplained but may be related to individual central auditory processing capacities (e.g., Dunn et al., 2005). In CI users, ERPs may be used to observe detection (exogenous components) and discrimination (endogenous components) of stimulus contrasts. Assuming there are no cognitive deficits, device-related factors (e.g., the number of implanted electrodes, frequency allocation) and patient-related factors (e.g., the electrode-neural interface, patterns of neural survival, etc.) may affect ERP responses. As such, it is important to select stimuli that are sufficiently contrastive when measuring ERPs. Some studies have used pure-tone contrasts (e.g., Groenen et al., 2001; Beynon et al., 2002; Sasaki et al., 2009; Obuchi et al., 2012; Calderaro et al., 2020; Van Yper et al., 2020; Wedekind et al., 2021) while others have used phonemic contrasts (e.g., Groenen et al., 2001; Beynon et al., 2002, 2005; Beynon and Snik, 2004; Henkin et al., 2009; Micco et al., 1995). ERPs have also been used to observe the evolution of auditory processing after cochlear implantation in longitudinal studies (e.g., Kubo et al., 2001).

P300 is closely related to cognitive processes involving auditory discrimination, selective attention, and working memory (e.g., Polich, 2007). Segregation of competing speech has been shown to involve cognitive processes (e.g., Francis, 2010). Some CI studies have compared P300 responses to

standard clinical measures such as word recognition in quiet (e.g., Kileny et al., 1997; Groenen et al., 2001; Grasel et al., 2018; Abrahamse et al., 2021; Amaral et al., 2021). Others have compared P300 responses to phoneme recognition in quiet (e.g., Groenen et al., 2001; Beynon et al., 2002) or to speech recognition in steady noise (e.g., Iwaki et al., 2004). Kileny et al. (1997) found a significant correlation between P300 amplitude and sentence recognition in pediatric CI users. Groenen et al. (2001) found a significant correlation between P300 amplitude and word/phoneme recognition in quiet in adult CI users.

Bimodal listening [CI in one ear, hearing aid (HA) in the other ear] provides important low-frequency temporal fine-structure cues that benefit pitch-mediated perception (e.g., music, talker identity, prosody) and segregation of target speech and maskers (e.g., Gifford et al., 2007; Cullington and Zeng, 2008; Dorman et al., 2008; Yoon et al., 2012; Crew et al., 2015; Liu et al., 2019). Previous studies have shown more robust P300 responses with bimodal than with CI-only listening. Iwaki et al. (2004) found that sentence recognition in noise was significantly better and P300 latency was significantly shorter with bimodal than with CI-only listening. Sasaki et al. (2009) also reported shorter P300 latency and better word recognition in quiet with bimodal than with CI-only listening. However, the relationship between P300 responses and segregation of competing speech with bimodal listening remains unclear.

In this study, P300 responses to pure-tone stimuli were recorded in NH listeners and bimodal CI users; speech recognition was measured in the presence of steady noise or competing speech. Given that the present participants used bimodal listening in daily life, only bimodal listening was tested. Also, previous studies have shown more robust P300 responses with bimodal than with CI-only listening (e.g., Iwaki et al., 2004; Sasaki et al., 2009). Consistent with previous studies (e.g., Beynon et al., 2005; Obuchi et al., 2012; Grasel et al., 2018), we expected greater P300 amplitudes and shorter P300 latencies in NH than in CI listeners. Given the great variability in speech performance among CI users (e.g., Stickney et al., 2004; Cullington and Zeng, 2008) and given that P300 is sensitive to auditory task difficulty (Parasuraman and Beatty, 1980; Polich, 1987; Causse et al., 2016), we expected that P300 responses would be related to masked speech recognition, especially for the more difficult segregation of competing speech by CI users.

METHODS

Participants

Eleven Mandarin-speaking CI listeners (six females, five males) participated in the study; the mean age at testing was 21.5 ± 9.2 years. All were users of Med-El devices. All except for CI-4 were implanted with the Sonata ti10 device with the Standard electrode array (31.5 mm); CI-4 was implanted with Concerto device and the Flex 28 electrode array (28 mm). All used the Opus 2 processor, and all used the FS4 strategy. All were bimodal listeners, using a CI in one ear and a hearing aid in the other ear in every day listening. The mean duration of deafness prior to implantation was 12.8 ± 6.5 years. The mean CI experience was 2.0 ± 1.9 years.

TABLE 1 | Demographic information of CI participants.

Participant	Sex	Age at test (yrs)	Dur deaf (yrs)	CI exp (yrs)	CI ear	Etiology	PTA (dB HL)	Mean HA gain (dB)
CI-C1	M	6.5	3.4	3.3	R	Congenital	94.2	33.3
CI-C2	F	10.6	8.0	0.8	L	Congenital	81.7	40.8
CI-C3	F	14.9	7.9	0.5	R	Unknown	82.5	29.2
CI-A1	F	20.1	15.0	0.8	R	Progressive	91.7	38.3
CI-A2	F	20.3	20.3	3.5	R	Congenital	90.8	29.2
CI-A3	F	20.7	20.0	0.6	L	Congenital	85.8	25.0
CI-A4	M	23.8	14.8	1.2	L	Unknown	85.0	23.3
CI-A5	M	23.9	12.9	1.0	L	Unknown	69.2	22.5
CI-A6	M	24.1	21.0	6.7	R	Congenital	81.7	20.0
CI-A7	M	31.3	15.0	0.6	R	Progressive	86.7	41.7
CI-A8	F	40.2	2.5	2.5	R	Sudden	83.3	17.5

All participants were users of Med-El devices, and all were everyday bimodal listeners (CI in one ear, hearing aid in the other ear).

Unaided PTA thresholds were calculated across 0.25, 0.5, 1.0, 2.0, 4.0, and 8.0 kHz. The mean HA gain was calculated as the mean difference between unaided and aided PTA thresholds. Dur deaf, duration of deafness before cochlear implantation; CI exp, experience with the CI device; CI-C, CI children; CI-A, CI adult.

CI participants C1, C2, A2, A3, A6 were prelingually deaf, and C3, A1, A4, A5, A7, A8 were postlingually deaf. **Table 1** shows demographic information for the CI participants. Eleven Mandarin-speaking NH listeners (seven females, four males) also participated in the study; the mean age at testing was 22.1 ± 9.5 years. A *t*-test showed no significant difference in age at testing between the CI and NH groups [$t(20) = 0.2$, $p = 0.882$]. All participants were recruited from Department of Ear, Nose, and Throat, The First Affiliated Hospital of Soochow University. The Ethical Committee from The First Affiliated Hospital of Soochow University specifically approved this study (Approval number 2021122). All participants provided written informed consent before participating in the study; parental approval was obtained for pediatric CI and NH listeners.

Speech Perception

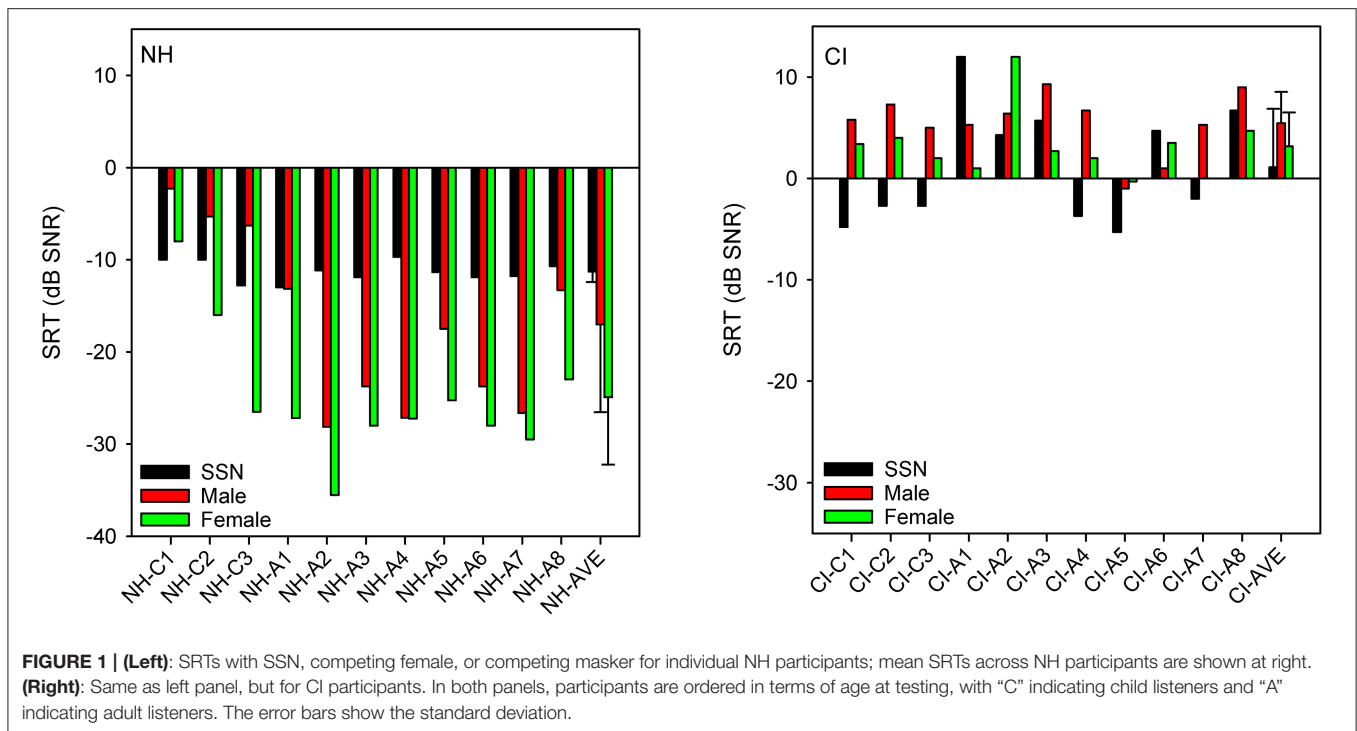
The Closed-set Mandarin Speech (CMS; Tao et al., 2017) test materials were used to test speech recognition with the different maskers. The CMS test materials consist of familiar words selected to represent the natural distribution of vowels, consonants, and lexical tones found in Mandarin Chinese. Ten keywords in each of five categories (Name, Verb, Number, Color, and Fruit) were produced by native Mandarin talkers.

Speech reception thresholds (SRTs), defined as the target-to-masker ratio (TMR) that produced 50% correct keyword recognition, were adaptively measured using a modified coordinate response matrix test (Brungart, 2001). Two target keywords (randomly selected from the Number and Color categories) were embedded in a five-word carrier sentence uttered by a male target talker [mean fundamental frequency (F0) across all words = 136 Hz]. The first word in the target sentence was always the Name “Xiaowang,” followed by randomly selected words from the remaining categories. Thus, the target sentence could be (translated from Mandarin) “**Xiaowang** sold **Three Red**

strawberries” or “**Xiaowang** chose **Four Brown** bananas,” etc. (Name to cue target talker in bold; target keywords in bold italic).

Recognition of the target keywords was measured in the presence of steady state noise (SSN) or competing speech; maskers were co-located with the target (0° azimuth). The spectrum of the SSN was matched to the long-term average spectrum of the target talker, averaged across all words. For competing speech, the masker was a female talker (mean F0 across all words = 248 Hz) or a different male talker (mean F0 = 178 Hz). Masker sentences were randomly generated for each test trial; words were randomly selected from each category, excluding the words used in the target sentence. Thus, the masker sentence could be “Xiaozhang saw **Two Blue** kumquats,” “Xiaodeng took **Eight Green** papayas,” etc. (competing keywords in italic).

All stimuli were presented in the sound field at 65 dBA via a single loudspeaker; subjects were seated in a sound-attenuated booth, directly facing the loudspeaker at a 1-m distance. For CI participants, SRTs were measured using the clinical settings for their devices, which were not changed throughout the study. During each test trial, a sentence was presented at the desired TMR; the initial TMR was 10 dB. Participants were instructed to listen to the target sentence (produced by the male target talker and beginning with the name “Xiaowang”) and then click on one of the 10 response choices for each of the Number and Color categories; no selections could be made from the remaining categories, which were grayed out. If the subject correctly identified both keywords, the TMR was reduced by 4 dB (initial step size); if the subject did not correctly identify both keywords, the TMR was increased by 4 dB. After two reversals, the step size was reduced to 2 dB. The SRT was calculated by averaging the last six reversals in TMR. If there were fewer than six reversals within 20 trials, the test run was discarded and another run was measured. Two test runs were completed for each condition and the SRT was averaged across runs. The masker conditions were randomized within and across participants.



P300 Recordings

P300 ERPs were recorded using the Smart EP software (Intelligent Hearing System, Miami, FL, USA) and a multichannel recording paradigm. Disposable electrodes were placed at the high forehead (non-inverting electrode), both sides of the mastoid (inverting electrode), and low forehead (ground electrode). Absolute impedances and inter-electrode impedances were <5 and $3 \text{ k}\Omega$, respectively. Responses were filtered online using a band-pass filter between 1 and 100 Hz. Pure-tone acoustic stimuli (1 or 2 kHz) with 50-ms duration and 5-ms rise and decay times were presented to the subjects every 1 s. Pure-tone stimuli were used instead of speech stimuli because pure-tone stimuli show better P300 reproducibility (e.g., Perez et al., 2017). The intensity of the stimuli was 20–30 dB above the aided PTA thresholds at 1 or 2 kHz to ensure that stimuli were clearly and comfortably audible for all participants.

Participants were seated in an electrically-shielded, sound-attenuated examination room. The stimuli were presented *via* two loudspeakers placed at ear level, 1 m away, $\pm 45^\circ$ relative to center. The probability was set at 80% for the frequent stimulus (1 kHz tone) and 20% for the rare stimulus (2 kHz tone). Participants were instructed to count the number of 2 kHz stimuli (oddball paradigm). All participants were able to discriminate between 1 and 2 kHz with 100% accuracy. In each run where all 20 oddball stimuli were identified, 20 ERPs for the rare stimuli were averaged. The recording window was comprised of a pre-stimulus baseline of 200 ms and a 500 ms post-stimulus epoch with a sampling rate of 1,000 Hz. Artifact rejection level was set at 100 mV. To avoid artifacts due to eye blinks, participants were instructed to close their eyes during the recording (Groenen et al., 2001). To reduce unwanted

alpha rhythm, the inter-stimulus-interval was jittered by $\pm 0.1 \text{ s}$ ($\pm 10\%$), which made stimulus presentation less predictable and participants more attentive. Also, alpha rhythm was partially canceled out during the average processing because the onset of the P300 ERP is random relative to the phase of the alpha wave (Talsma and Woldorff, 2005).

P300 amplitude was calculated between the most positive point in the waveform between ≈ 250 –400 ms and the following most negative point. This approach was chosen because the following most negative point was more distinct than the previous negative point. P300 latency was identified according to the P300 positive point. A minimum of three runs were tested, with more as needed if the participant did not identify all 20 oddball stimuli; only test runs where all 20 oddball stimuli were identified were included in analyses. Rest periods were taken between sessions to keep the participants alert.

RESULTS

Figure 1 shows SRTs with SSN or with a competing male or female talker for the NH and CI listeners. SRTs were much lower (better) for NH than for CI listeners. For NH listeners, mean SRTs progressively improved from SSN ($-11.3 \pm 1.1 \text{ dB}$) to the male masker ($-17.0 \pm 9.0 \text{ dB}$) and then to the female masker ($-24.9 \pm 7.3 \text{ dB}$). For CI listeners, mean SRTs were poorer with the competing male ($5.5 \pm 3.1 \text{ dB}$) or female talker ($3.2 \pm 3.3 \text{ dB}$) than with SSN ($1.1 \pm 5.9 \text{ dB}$). A mixed-design analysis of variance (ANOVA) was performed on the SRT data, with masker (SSN, male, female) as the within-subject factor and group (NH, CI) as the between-subject factor. Results showed significant effects of

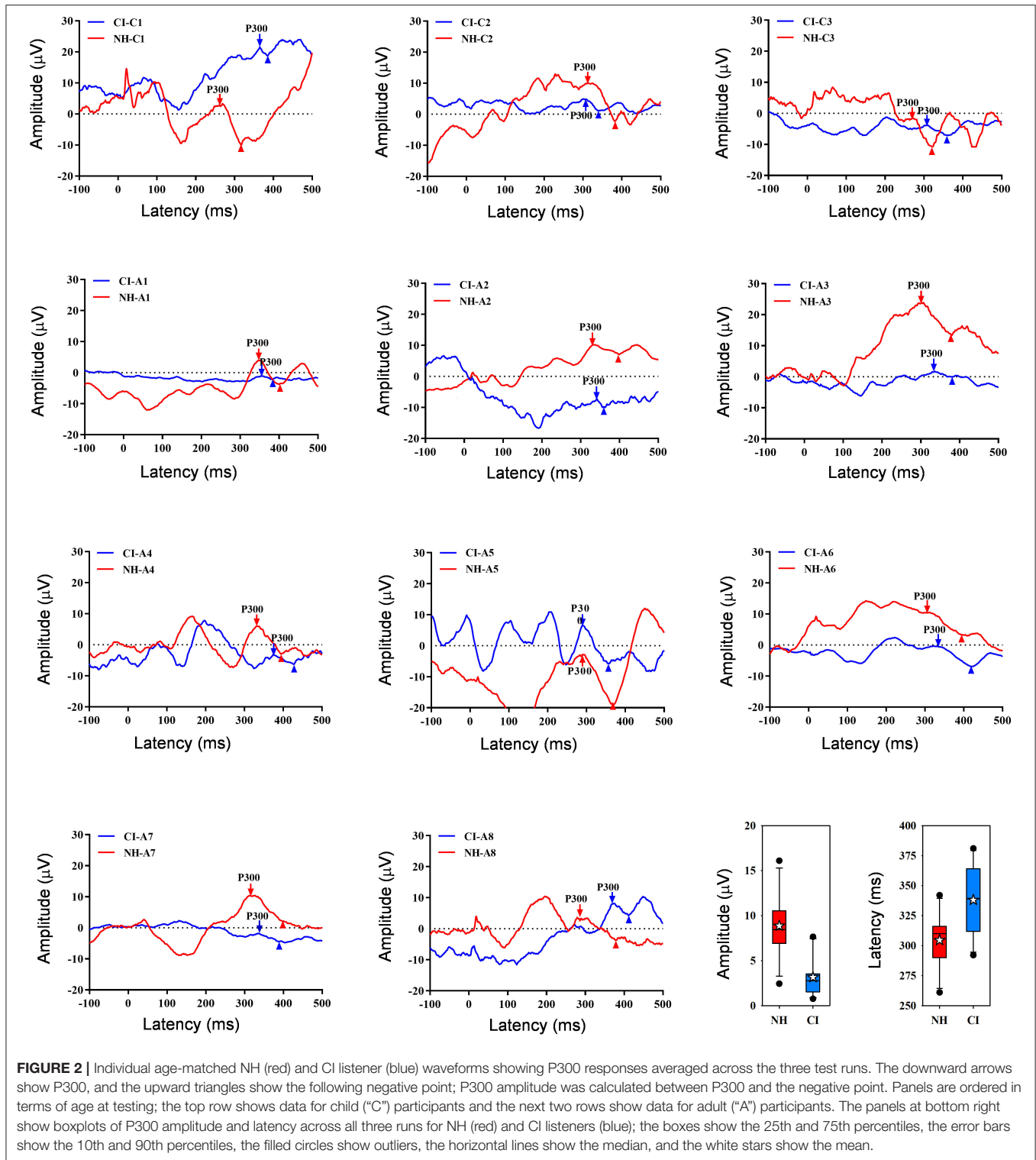


FIGURE 2 | Individual age-matched NH (red) and CI listener (blue) waveforms showing P300 responses averaged across the three test runs. The downward arrows show P300, and the upward triangles show the following negative point; P300 amplitude was calculated between P300 and the negative point. Panels are ordered in terms of age at testing; the top row shows data for child (“C”) participants and the next two rows show data for adult (“A”) participants. The panels at bottom right show boxplots of P300 amplitude and latency across all three runs for NH (red) and CI listeners (blue); the boxes show the 25th and 75th percentiles, the error bars show the 10th and 90th percentiles, the filled circles show outliers, the horizontal lines show the median, and the white stars show the mean.

group [$F_{(1,40)} = 125.1, p < 0.001$] and masker [$F_{(2,40)} = 10.7, p < 0.001$]; there was a significant interaction [$F_{(2,40)} = 16.8, p < 0.001$]. *Post-hoc* Bonferroni pairwise comparisons showed that for the NH group, SRTs were significantly higher (poorer) with SSN than with the male ($p = 0.016$) or female masker

($p < 0.001$), and significantly higher with the male than with the female masker ($p < 0.001$). There were no significant differences among the maskers for the CI group. SRTs were significantly lower (better) for the NH than for the CI group for all maskers ($p < 0.001$ for all comparisons).

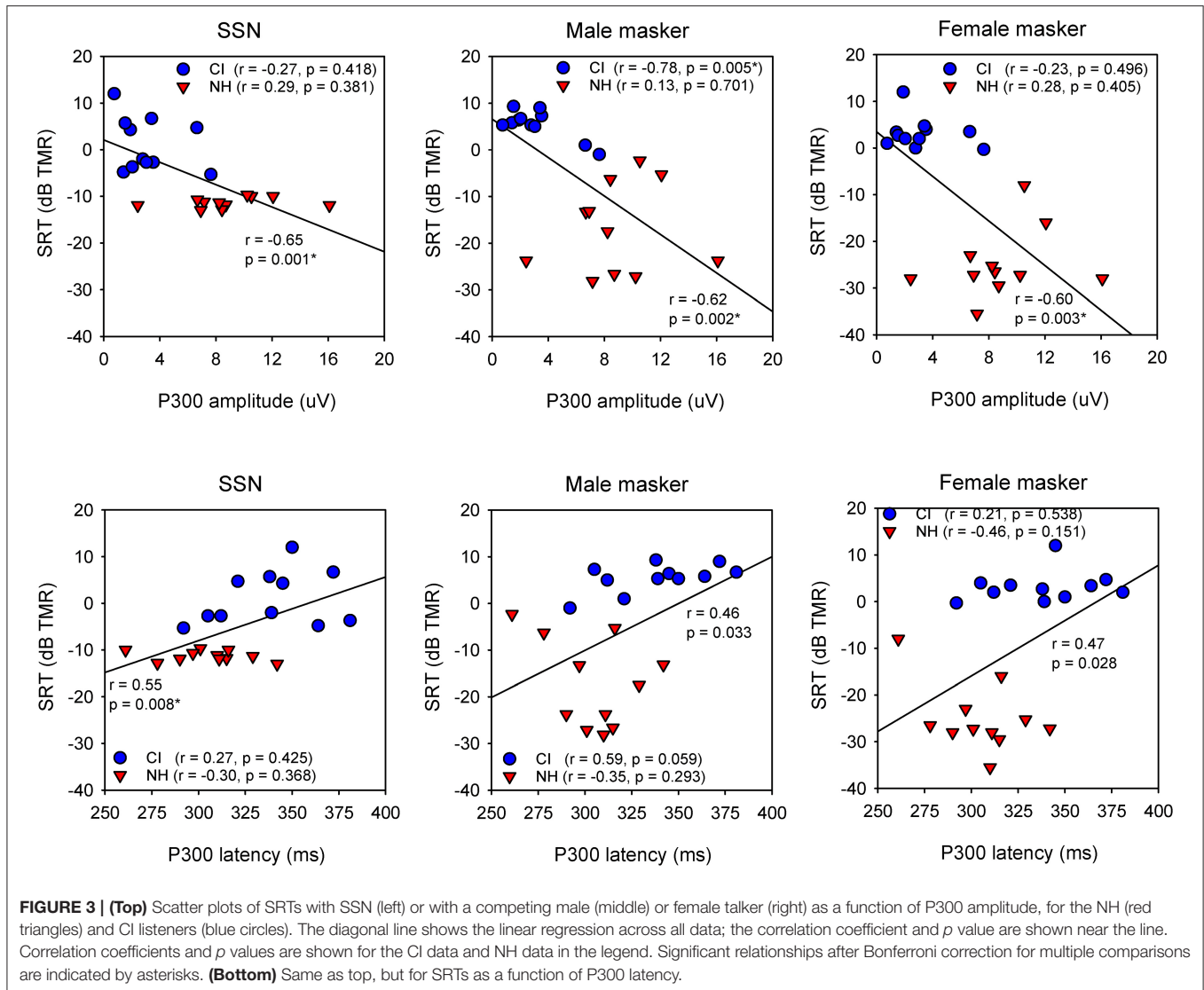


Figure 2 shows waveforms with the peak P300 response averaged across the three test runs for individual NH and CI listeners. Note that the intra-class correlation coefficient was 0.99 and 0.97 for P300 amplitude and latency, respectively, suggesting good test-retest reliability across the three runs. Because RM ANOVAs showed no significant effect of test run for NH or CI participants ($p > 0.05$ for all analyses), data were averaged across runs. Mean P300 amplitude was higher for the NH group ($8.9 \pm 3.5 \mu\text{V}$) than for the CI group ($3.2 \pm 2.2 \mu\text{V}$); mean P300 latency was shorter for the NH group ($305 \pm 23 \text{ ms}$) than for the CI group ($338 \pm 28 \text{ ms}$). *T*-tests showed that P300 amplitude was significantly higher for the NH than for the CI group [$t(20) = 4.6$, $p < 0.001$], and that P300 latency was significantly shorter for the NH than for the CI group [$t(20) = -3.1$, $p = 0.006$].

Figure 3 shows SRTs with SSN or with a competing male or female talker for the NH and CI groups as a function of P300 amplitude and latency; each data point shows the mean across three test runs. When all NH and CI data were combined, and

after Bonferroni correction for multiple comparisons (adjusted $p = 0.016$), Pearson correlation analysis showed significant relationships between P300 amplitude and SRTs with SSN ($r = -0.65$, $p = 0.001$), and with the male ($r = -0.62$, $p = 0.002$) and female maskers ($r = -0.60$, $p = 0.003$). A significant relationship was observed between P300 latency and SRTs with SSN ($r = 0.60$, $p = 0.008$), but not for SRTs with the male or female masker. For the CI group, Pearson correlation analysis showed a significant relationship only between P300 amplitude and SRTs with the male masker ($r = -0.78$, $p = 0.005$); the correlation remained significant after controlling for age at testing, duration of deafness, and CI experience ($r = -0.81$, $p = 0.016$). No significant correlations were observed between P300 amplitude and SRTs with SSN or with the female masker, or between P300 latency and SRTs with any of the maskers. For the NH group, no significant relationships were observed between P300 amplitude or latency and SRTs with any of the maskers. For the CI group, a significant correlation was observed between P300 amplitude

and unaided PTA thresholds (across all frequencies; $r = -0.87$, $p < 0.001$); there were no significant correlations between P300 amplitude and aided PTA thresholds. Significant correlations were observed between P300 latency and unaided PTA thresholds ($r = 0.67$, $p = 0.025$) and aided PTA thresholds ($r = 0.68$, $p = 0.021$). Note that statistical power was >0.80 for all of the above correlations, except for P300 latency vs. unaided PTA thresholds (power = 0.63) or aided PTA thresholds (power = 0.65).

DISCUSSION

Consistent with previous studies (e.g., Kubo et al., 2001; Beynon et al., 2005; Obuchi et al., 2012; Soshi et al., 2014; Grasel et al., 2018; Han et al., 2020), P300 amplitudes were significantly larger and latencies were significantly shorter for the NH group than for the CI group. For the present bimodal CI listeners, mean P300 amplitude and/or latency values were comparable to those observed in previous studies with CI listeners (e.g., Iwaki et al., 2004; Sasaki et al., 2009; Grasel et al., 2018; Abrahamse et al., 2021; Calderaro et al., 2020; Van Yper et al., 2020). P300 responses were elicited in all CI participants, consistent with Obuchi et al. (2012).

Mean SRTs for all maskers were lower (better) for the NH group than for the CI group, and values were comparable to those in previous studies using similar methods and stimuli (Tao et al., 2018; Zhang et al., 2020). Different from previous CI studies that showed lower SRTs in SSN than in competing speech (e.g., Cullington and Zeng, 2008; Croghan and Smith, 2018; Tao et al., 2018; Liu et al., 2019), there was no significant difference in SRTs between the SSN and competing speech maskers within the CI group. Note that CI listeners were tested while wearing contralateral hearing aids, which likely aided in segregation of competing speech, thereby reducing the deficit relative to SSN.

Across all NH and CI listeners, significant correlations were observed between P300 amplitude and SRTs with the SSN, male, and female maskers; a significant correlation was also observed between P300 latency and SRTs with SSN. These correlations were largely driven by across-group differences in speech performance and P300 responses. In general, higher P300 amplitude and shorter P300 latency were associated with better masked speech recognition.

In the NH group, there were no significant correlations between P300 responses and SRTs with any of the maskers. In the CI group, a significant correlation was observed only between P300 amplitude and SRTs with the male masker, the most challenging listening condition with the greatest informational masking. The correlation between P300 amplitude and SRTs with the male masker suggests some common relation to informational masking, a central auditory process. With the female masker, informational masking was reduced, and SSN produced largely energetic masking. Given the correlations between unaided PTA thresholds and P300 amplitude and latency and between aided PTA thresholds and P300 latency, differences in P300 response across CI listeners may have represented differences in segregation of the competing male

talkers with residual acoustic hearing that provided low-frequency pitch cues.

Different from Soshi et al. (2014), we observed a significant correlation between P300 amplitude and SRTs with the male masker, but not between P300 amplitude and SRTs with SSN. Differences in cortical measure stimuli (1 vs. 2 kHz contrasts; consonant contrast), speech tests, methods, and CI patients (bimodal vs. CI-only listening) may have contributed to differences in results across studies. The 1 and 2 kHz stimuli used for ERP recording were presented at 20–30 dB above the aided thresholds, meaning that the aided acoustic hearing should have contributed to the response.

In the present study, ERPs and speech performance were measured only with bimodal listening. Some studies have shown greater P300 response and speech performance with bimodal than with CI-only listening (e.g., Iwaki et al., 2004; Sasaki et al., 2009). Interestingly, Wedekind et al. (2021) found no significant difference in P300 response between the NH ear and the CI ear in unilaterally deaf CI recipients; speech recognition in noise was better with the CI on than off. While it was not directly measured in Wedekind et al. (2021), speech performance would be expected to be much poorer with the CI ear alone than with the NH ear alone (e.g., Galvin et al., 2019). It is unclear why the P300 response would be similar across ears when speech performance would be different. As shown in **Figure 3**, significant relationships were observed between P300 amplitude and masked SRTs, presumably due to the underlying spectro-temporal resolution that was much better for NH than for CI listeners. However, some caution is warranted regarding the correlational analyses, given the limited number of participants and test runs. ERPs and speech performance were not measured with the acoustic-hearing ear alone or the CI ear alone in this study. It is possible that strong P300 responses may have been elicited within the acoustic-hearing ear alone, despite the expectedly poor speech performance. In future studies, it would be worthwhile to collect ERPs and speech performance with each ear alone and both ears together to better understand how the peripheral representations might affect the relationship between ERPs and speech performance.

The present results show some evidence that ERPs may be a useful objective measure to predict complex perception such as segregation of competing speech. However, eliciting P300 also requires a behavioral component in the oddball presentation, and the magnitude of the response may depend on the strength of the stimulus contrast. Obuchi et al. (2012) showed increasing P300 amplitude in CI listeners as the stimulus frequency contrast was increased from 1.5 to 4 kHz. Depending on the acoustic-to-electric frequency allocation and the electrode-neural interface (electrode position relative to healthy neurons), small contrasts (e.g., 1 vs. 1.5 kHz) may be perceived differently among CI listeners. The 1 vs. 2 kHz contrast in this study appeared to be sufficiently large to be discriminated by the present MED-EL CI users, most likely resulting in stimulation of electrodes 6 and 8, given the default frequency allocation. Note that there may have been some contribution from residual acoustic hearing for discrimination of the stimuli contrast.

CONCLUSIONS

Auditory ERPs and speech recognition in steady noise or competing speech were measured in NH and bimodal CI listeners. P300 amplitude was larger and latency was shorter in the NH group than in the CI group. Similarly, speech performance was better for the NH group than for the CI group. Significant correlations were observed across all participants between P300 amplitude and SRTs with steady noise and the male and female maskers. Within the CI group, P300 amplitude was significantly correlated with SRTs with the male masker, suggesting some relation between cortical response and informational masking.

DATA AVAILABILITY STATEMENT

The raw data from the study are included in the supplementary material; further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee from the First Affiliated Hospital of Soochow University (Approval number 2021122).

REFERENCES

- Abrahamse, R., Beynon, A., and Piai, V. (2021). Long-term auditory processing outcomes in early implanted young adults with cochlear implants: the mismatch negativity vs. P300 response. *Clin. Neurophysiol.* 132, 258–268. doi: 10.1016/j.clinph.2020.09.022
- Amaral, M. S. A. D., Calderaro, V. G., Pauna, H. F., Massuda, E. T., Reis, A. C. M. B., and Hyppolito, M. A. (2021). Is there a change in P300 evoked potential after 6 months in cochlear implant users? *Braz. J. Otorhinolaryngol.* 10, S1808-8694(21)00182-8. doi: 10.1016/j.bjorl.2021.10.002
- Beynon, A. J., and Snik, A. F. (2004). Use of the event-related P300 potential in cochlear implant subjects for the study of strategy-dependent speech processing. *Int. J. Audiol.* 43(Suppl. 1), S44–S47.
- Beynon, A. J., Snik, A. F., Stegeman, D. F., and van den Broek, P. (2005). Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *J. Am. Acad. Audiol.* 16, 42–53. doi: 10.3766/jaaa.16.1.5
- Beynon, A. J., and Snik, A. F., and van den Broek, P. (2002). Evaluation of cochlear implant benefit with auditory cortical evoked potentials. *Int. J. Audiol.* 41, 429–435. doi: 10.3109/14992020209090420
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696
- Calderaro, V. G., Amaral, M. S. A. D., Luz, B. A. B. D., Bernal, S. C., Hyppolito, M. A., and Reis, A. C. M. B. (2020). Behavioral and electrophysiological assessment of adults who underwent cochlear implantation after hearing aid experience. *Int. Arch. Otorhinolaryngol.* 24, e132–e139. doi: 10.1055/s-0039-1695022
- Causse, M., Peysakhovich, V., and Fabre, E. F. (2016). High working memory load impairs language processing during a simulated piloting task: an ERP and pupillometry study. *Front. Hum. Neurosci.* 10:240. doi: 10.3389/fnhum.2016.00240
- Crew, J. D., Galvin, J. J. 3rd, Landsberger, D. M., and Fu, Q. J. (2015). Contributions of electric and acoustic hearing to bimodal speech and music perception. *PLoS ONE*. 10, e0120279. doi: 10.1371/journal.pone.0120279

All participants provided written informed consent before participating in the study; approval was obtained for pediatric CI and NH participants from their parents or adult next of kin.

AUTHOR CONTRIBUTIONS

D-DT: study design, data analysis, and writing of manuscript. Y-MZ and DZ: data analysis and writing of manuscript. JG: data analysis, data visualization, and writing of manuscript. HL, WZ, and MX: study supervision. J-SL: study design. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81970877), the Jiangsu Provincial Natural Science Foundation—Outstanding Youth Foundation (BK20200054), the Jiangsu Provincial Key Research and Development Program Special Funds (BE2019670), and the Suzhou Science and Technology Project (SYS2019049).

ACKNOWLEDGMENTS

We thank the participants and their families for their time and contribution to this study.

- Croghan, N. B. H., and Smith, Z. M. (2018). Speech understanding with various maskers in cochlear-implant and simulated cochlear-implant hearing: effects of spectral resolution and implications for masking release. *Trends Hear.* 22, 2331216518787276. doi: 10.1177/2331216518787276
- Cullington, H. E., and Zeng, F. G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *J. Acoust. Soc. Am.* 123, 450–461. doi: 10.1121/1.2805617
- Donchin, E., and Coles, M. G. H. (1988). On the conceptual foundations of cognitive psychology. *Behav. Brain Sci.* 1, 406–417. doi: 10.1017/S0140525X00058246
- Dorman, M. F., Gifford, R. H., Spahr, A. J., and McKarns, S. A. (2008). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol. Neurootol.* 13, 105–112. doi: 10.1159/000111782
- Dunn, C. C., R. S., and Tyler, and, S. A., Witt (2005). Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. *J. Speech Lang. Hear. Res.* 48, 668–680. doi: 10.1044/1092-4388(2005/046)
- Francis, A. L. (2010). Improved segregation of simultaneous talkers differentially affects perceptual and cognitive capacity demands for recognizing speech in competing speech. *Atten. Percept. Psychophys.* 72, 501–516. doi: 10.3758/APP.72.2.501
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.* 110, 1150–1163. doi: 10.1121/1.1381538
- Fu, Q. J., and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *J. Assoc. Res. Otolaryngol.* 6, 19–27. doi: 10.1007/s10162-004-5024-3
- Galvin, J. J. 3rd, Fu, Q. J., Wilkinson, E. P., Mills, D., Hagan, S. C., Lupo, J. E., et al. (2019). Benefits of cochlear implantation for single-sided deafness: data from the House Clinic-University of Southern California-University of California, Los Angeles clinical trial. *Ear Hear.* 40, 766–781. doi: 10.1097/AUD.0000000000000671
- Gifford, R. H., Dorman, M. F., McKarns, S. A., and Spahr, A. J. (2007). Combined electric and contralateral acoustic hearing: word and sentence

- recognition with bimodal hearing. *J. Speech Lang. Hear. Res.* 50, 835–843. doi: 10.1044/1092-4388(2007/058)
- Grasel, S., Greters, M., Goffi-Gomez, M. V. S., Bittar, R., Weber, R., Oiticica, J., et al. (2018). P3 cognitive potential in cochlear implant users. *Int. Arch. Otorhinolaryngol.* 22, 408–414. doi: 10.1055/s-0037-1613687
- Groenen, P. A., Beynon, A. J., Snik, A. F., and van den Broek, P. (2001). Speech-evoked cortical potentials and speech recognition in cochlear implant users. *Scan Audiol.* 30, 31–40. doi: 10.1080/010503901750069554
- Han, J. H., Lee, J., and Lee, H., J. (2020). Noise-induced change of cortical temporal processing in cochlear implant users. *Clin. Exp. Otorhinolaryngol.* 13, 241–248. doi: 10.21053/ceo.2019.01081
- Henkin, Y., Tetin-Schneider, S., Hildesheimer, M., and Kishon-Rabin, L. (2009). Cortical neural activity underlying speech perception in postlingual adult cochlear implant recipients. *Audiol. Neurootol.* 14, 39–53. doi: 10.1159/000153434
- Hillyard, S. A., Squires, K. C., Bauer, J. W., and Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science.* 172, 1357–1360. doi: 10.1126/science.172.3990.1357
- Iwaki, T., Matsushiro, N., Mah, S. R., Sato, T., Yasuoka, E., Yamamoto, K., et al. (2004). Comparison of speech perception between monaural and binaural hearing in cochlear implant patients. *Acta Otolaryngol.* 124, 358–362. doi: 10.1080/000164803100005484
- Kidd, G. Jr, Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., and Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *J. Acoust. Soc. Am.* 140, 132. doi: 10.1121/1.4954748
- Kidd, G. Jr., Arbogast, T. L., Mason, C. R., and Walsh, M. (2002). Informational masking in listeners with sensorineural hearing loss. *J. Assoc. Res. Otolaryngol.* 3, 107–119. doi: 10.1007/s101620010095
- Kileny, P. R., Boerst, A., and Zwolan, T. (1997). Cognitive evoked potentials to speech and tonal stimuli in children with implants. *Otolaryngol. Head Neck Surg.* 117, 161–169. doi: 10.1016/s0194-5998(97)70169-4
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiol.* 38, 557–577. doi: 10.1017/s0048577201990559
- Kubo, T., Yamamoto, K., Iwaki, T., Matsukawa, M., Doi, K., and Tamura, M. (2001). Significance of auditory evoked responses (EABR and P300) in cochlear implant subjects. *Acta Otolaryngol.* 121, 257–261. doi: 10.1080/000164801300043749
- Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science.* 197, 792–795. doi: 10.1126/science.887923
- Lightfoot, G. (2016). Summary of the N1-P2 cortical auditory evoked potential to estimate the auditory threshold in adults. *Semin. Hear.* 37, 1–8. doi: 10.1055/s-0035-1570334
- Liu, Y. W., Tao, D. D., Chen, B., Cheng, X., Shu, Y., Galvin, J. J., et al. (2019). Factors affecting bimodal benefit in pediatric Mandarin-speaking Chinese cochlear implant users. *Ear Hear.* 40, 1316–1327. doi: 10.1097/AUD.0000000000000712
- Luo, X., and Fu, Q. J. (2009). Concurrent-vowel and tone recognitions in acoustic and simulated electric hearing. *J. Acoust. Soc. Am.* 125, 3223–3233. doi: 10.1121/1.3106534
- Martin, B. A., K. L., and Tremblay, and, P., Korczak (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear Hear.* 29, 285–313. doi: 10.1097/AUD.0b013e3181662c0e
- Micco, A. G., Kraus, N., Koch, D. B., McGee, T. J., Carrell, T. D., Sharma, A., et al. (1995). Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am. J. Otol.* 16, 514–520.
- Obuchi, C., Harashima, T., and Shiroma, M. (2012). Auditory evoked potentials under active and hearing conditions in adult cochlear implant users. *Clin. Exp. Otorhinolaryngol.* 5(Suppl. 1), S6–S9. doi: 10.3342/ceo.2012.5.S1.S6
- Parasuraman, R., and Beatty, J. (1980). Brain events underlying detection and recognition of weak sensory signals. *Science.* 210, 80–83. doi: 10.1126/science.7414324
- Perez, A. P., Ziliotto, K., and Pereira, L. D. (2017). Test-retest of long latency auditory evoked potentials (P300) with pure tone and speech stimuli. *Int. Arch. Otorhinolaryngol.* 21, 134–139. doi: 10.1055/s-0036-1583527
- Picton, T. W. (2011). “Human auditory evoked potentials,” in *Endogenous Auditory Evoked Potentials: Attention Must be Paid*, ed. T. W. Picton (San Diego, CA: Plural Publishing), 399–448.
- Polich, J. (1987). Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalogr. Clin. Neurophysiol.* 68, 311–320. doi: 10.1016/0168-5597(87)90052-9
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Polich, J., and Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biol. Psych.* 41, 103–146. doi: 10.1016/0301-0511(95)05130-9
- Ritter, W., Simson, R., and Vaughan H. G. Jr. (1972). Association cortex potentials and reaction time in auditory discrimination. *Electroenceph. Clin. Neurophysiol.* 33, 547–555. doi: 10.1016/0013-4694(72)90245-3
- Sasaki, T., Yamamoto, K., Iwaki, T., and Kubo, T. (2009). Assessing binaural/bimodal advantages using auditory event-related potentials in subjects with cochlear implants. *Auris Nasus Larynx.* 36, 541–546. doi: 10.1016/j.anl.2008.12.001
- Shannon, R. V., Fu, Q. J., and Galvin, J. 3rd (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol. Suppl.* 552, 50–54. doi: 10.1080/03655230410017562
- Soshi, T., Hisanaga, S., Kodama, N., Kanekama, Y., Samejima, Y., Yumoto, E., et al. (2014). Event-related potentials for better speech perception in noise by cochlear implant users. *Hear Res.* 316, 110–121. doi: 10.1016/j.heares.2014.08.001
- Squires, K. C., Hillyard, S. A., and Lindsay, P. H. (1973). Vertex potentials evoked during auditory signal detection: relation to decision criteria. *Percept. Psychophys.* 14, 263–272. doi: 10.3758/BF03212388
- Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *J. Acoust. Soc. Am.* 116, 1081–1091. doi: 10.1121/1.1772399
- Stone, M. A., and Canavan, S. (2016). The near non-existence of “pure” energetic masking release for speech: extension to spectro-temporal modulation and glimpsing. *J. Acoust. Soc. Am.* 140, 832–842. doi: 10.1121/1.4960483
- Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science.* 150, 1187–1188. doi: 10.1126/science.150.3700.1187
- Talsma, D., and Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* 17, 1098–1114. doi: 10.1162/0898929054475172
- Tao, D. D., Fu, Q. J., Galvin, J. J. 3rd, and Yu, Y. F. (2017). The development and validation of the Closed-set Mandarin Sentence (CMS) test. *Speech Comm.* 92, 125–131. doi: 10.1016/j.specom.2017.06.008
- Tao, D. D., Liu, Y. W., Fei, Y., Galvin, J. J. 3rd, Chen, B., and Fu, Q. J. (2018). Effects of age and duration of deafness on Mandarin speech understanding in competing speech by normal-hearing and cochlear implant children. *J. Acoust. Soc. Am.* 144, EL131. doi: 10.1121/1.5051051
- Van Yper, L. N., Dhooge, I. J. M., Vermeire, K., De Vel, E. F. J., and Beynon, A. J. (2020). The P300 auditory event-related potential as a method to assess the benefit of contralateral hearing aid use in bimodal listeners: a proof-of-concept. *Int. J. Audiol.* 59, 73–80. doi: 10.1080/14992027.2019.1656346
- Wedekind, A., Távora-Vieira, D., Nguyen, A. T., Marinovic, W., and Rajan, G. P. (2021). Cochlear implants in single-sided deaf recipients: near

- normal higher-order processing. *Clin. Neurophysiol.* 132, 449–456. doi: 10.1016/j.clinph.2020.11.038
- Yoon, Y. S., Li, Y., and Fud Q. J. (2012). Speech recognition and acoustic features in combined electric and acoustic stimulation. *J. Speech Lang. Hear. Res.* 55, 105–124. doi: 10.1044/1092-4388(2011/10-0325)
- Zhang, J., Wang, X., Wang, N. Y., Fu, X., Gan, T., Galvin, J. J. 3rd, et al. (2020). Tonal language speakers are better able to segregate competing speech according to talker sex differences. *J. Speech Lang. Hear. Res.* 63, 2801–2810. doi: 10.1044/2020_JSLHR-19-00421

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Tao, Zhang, Liu, Zhang, Xu, Galvin, Zhang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.