INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

The Implication of Sound Level on Spatial Selective Auditory Attention for Cochlear Implant Users: Behavioral and Electrophysiological Measurement

Sara Akbarzadeh[1], *Sungmin Lee*[2], *Chin-Tuan Tan*[1]

[1]University of Texas at Dallas
[2] Tongmyong University

`Sara.akbarzadeh@utdallas.edu, Slee18@tu.ac.kr, Chin-tuan.tan@utdallas.edu`

Abstract

This study investigated the effect of sound level on the spatial selective auditory attention of Normal Hearing (NH) and Cochlear Implant (CI) listeners behaviorally and electrophysiologically. Three spatially separated speech streams consisting of one target and two maskers were presented to the subjects. While keeping the same target to masker ratio, the target stimuli were presented at three different sound levels. In the behavioral test, subjects were instructed to attend the target speech, and speech perception score was calculated based on correctly recognized words in percentage. In the electrophysiological test, electroencephalography (EEG) signals were recorded while the subjects were listening for target speech in the presence of two maskers. The attended speech envelope was reconstructed from EEG using a linear regression model. The spatial auditory attention was detected through comparison of the reconstructed speech envelope with the original envelopes associated with the three speech streams presented. Outcome of both behavioral and electrophysiological experiments showed that an increase in the sound level decreases the spatial speech recognition performance of CI listeners, but not NH listeners.

Index Terms: selective auditory attention, speech detection accuracy, speech reconstruction, EEG

1. Introduction

The human auditory system is able to attend a target sound source and ignore the competing sounds in multi-speaker environments, which is also known as the cocktail party scenario [1]. Since this scenario was first described by Cherry [2], there has been a significant increase in auditory neuroscience studies attempting to explain the underlying mechanism of the auditory system under this effect [3][4][5][6]. This area of emphasis has also inspired the research community to further examine the central issue of the selective auditory attention with hearing impairment [7][8][9][10], particularly with severely impaired Cochlear Implant (CI) listeners [11][12][13][14].

Spatial separation between the target speech and competing sounds helps listeners to segregate the sound sources and consequently improves the speech intelligibility. Different factors affecting spatial selective auditory attention have been studied with Normal Hearing (NH) and Hearing Impaired (HI) listeners and one common conclusion is that reduced audibility limits the spatial advantage in their speech recognition performance [15][16][17][18][19]. Glyde et al. [20] found that the spatial hearing deficit in HI listeners (compared to NH listeners) persists even when their impaired hearing has been compensated. To date, we found that there are not many studies on the effect of sound level on spatial speech recognition performance for CI users.

Our previous study [21] investigated the effect of sound level on the speech quality perception in NH and CI listeners. The study showed that when the Signal to Noise Ratio (SNR) was kept the same at different conversational speech levels, the pattern of speech quality perception was different for NH and CI listeners; NH listeners preferred higher speech levels, and CI listeners preferred lower speech levels. Our CI listeners have chosen lower noise levels at the expense of reduced speech audibility. This perceptual difference between the two groups has inspired us to study the effect of speech level on the spatial speech recognition performance by NH and CI listeners.

Spatial hearing is mostly measured from behavioral response but can also be observed in the electrophysiological response. Some studies successfully determined the locus of attention from EEG signals in a two speaker (one target and one masker) cocktail party scenario [22][23][24][1][25][26][18]. In real cocktail party scenarios, people are capable of isolating the chosen speech from dozens of competing sounds presented in a room. To investigate the underlying neural basis for spatial segregation of sound sources in complex sound settings, we presented three spatially separated speech streams, one as a target and two as maskers in this study. We examined the effect of stimulus level on the spatial speech recognition performance for NH and CI listeners by varying the sound level of the stimuli while keeping the Target to Masker Ratio (TMR) constant, both in behavioral and electrophysiological experiments.

2. Method

2.1. Participants

Six NH listeners (four males, two females; mean age: 23, range: 21-29, SD: 3.8 years) and five bilateral CI users (two males, three females; mean age: 56, range: 24-76, SD: 23 years) participated in this study. All participants self-reported negative history of cognitive deficits prior to participation. All NH listeners were screened to have better hearing than 20 dB HL. Table 1 shows the demographic information of CI users. All hearing tests were conducted at octave frequency range from 250 to 8000 Hz. All procedures were approved by the institutional review board of University of Texas at Dallas.

2.2. Behavioral test

2.2.1. Stimuli

We adopted a matrix sentence corpus for spatial hearing test. On each trial, a set of stimuli consisting of one target and two masker sentences was presented simultaneously to the subject via different loudspeakers. Each sentence was chosen from a

`http://dx.doi.org/10.21437/Interspeech.2020-2836`

data set consisting of five groups of words (in the order of 1- name, 2- verb, 3- number, 4- adjective, and 5- noun). Each of the five groups consisted of 8 words, spoken by 18 female and 18 male talkers. The sentences were grammatically correct and sound natural, but conceptually unpredictable to minimize the effect of higher order language processing.

2.2.2. *Procedure*

Subjects were seated in the middle of a double-wall sound treated booth, while surrounded by five speakers at -90˚, -45˚, 0˚, 45˚, and 90˚ azimuth. The speakers were located at the radius of 1 meter away from the subjects with the height of the subjects' head.

In each trial, three different talkers were randomly selected from 36 talkers to utter the three different sentences with one as a target and two as maskers. An arrow appeared on the screen, placed in front of the subject, showing the location of the target speaker that subject should attend to. After presenting the stimuli, the subjects were instructed to select the words uttered by the target speaker, from the table presented on the screen.

To investigate the effect of speech level on spatial hearing, three different sound levels at fixed TMR were examined. For CI listeners, stimuli were presented at 10 dB TMR, target at 75, 65, and 55 dB SPL and maskers at 65, 55, and 45 dB SPL, respectively. For NH listeners, to avoid ceiling effect in their responses, stimuli were presented at 0 dB TMR, which means target at 75, 65, and 55 dB SPL and maskers at 75, 65, and 55 dB SPL.

Five speakers' configuration were examined, as shown in Figure 1. The experiments were conducted in three sessions, associated with three different stimuli levels, over a day. At a given session, 50 trials (five target speaker configurations * 10 repetitions) were randomly presented to the subject. Each subject participated in one training session to get familiar with the procedure. At the training session, targets were presented at 65 dB SPL, and maskers were presented at 55 dB SPL for CI listeners and 65 dB SPL for NH listeners.

2.3. Electro-physiological test

The electro-physiological measurements were acquired using EEG signals. The BrainVision system (actiCHamp amplifier) was used to obtain EEG signals via a 64-channel actiCAP Electrode Cap (the same set up used in [27]). EEG was digitized at the sampling rate of 1 kHz and stored for offline analysis. The primary EEG analysis method used in this study is cortical entrainment to the speech envelope. This requires a decoder which extracts the speech envelope from EEG signal. Thus, our EEG test consisted of two parts: one for training a decoder in a plain condition and the other for measuring the EEG associated with various spatial configuration and sound level conditions.

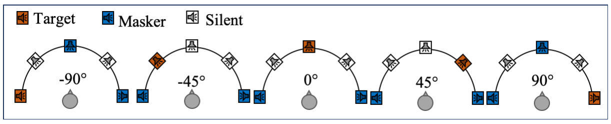

Figure 1: *Speakers' configuration*

2.3.1. *Training the decoder*

To collect the data for training the decoder, the subjects listened to two short stories. The first story was "lady or tiger", approximately 10 minutes long (632.7341 sec) narrated by a female speaker. The second story was "ambitious guest", approximately 10 minutes long (615.2858 sec) narrated by a male speaker. It should be noted that both stories were presented in American English and all silences longer than 300 msec in the speech were removed. The stories presented at 65 dB SPL from front speaker without any noise. EEG data was recording while the subject was listening to the stimuli. Subjects were instructed to sit calm with their eyes open and to look at a fixed point on the monitor.

2.3.2. *Test in spatial configuration*

We had one target and two maskers passages, which were selected randomly from Connected Speech Test (CST) [28] and Speech Intelligibility Rating (SIR) [29] data set. Half of the trials had a female speaker as target and the other half had a male speaker as target. While having a female target, two maskers were male, and while having a male target, two maskers were female. The length of the sentences was between 38 and 45 seconds. Three speaker configurations were examined for different permutation of target and masker locations using the speakers at -90, 0, and 90 azimuths. The stimuli were presented at three different levels with fixed TMR at 10 dB. Targets were presented at 75, 65, and 55 dB SPL and maskers at 65, 55, and 45 dB SPL, respectively. All 18 trials (3 speaker configurations*2 target genders*3 speech levels) were randomly presented to each subject.

At the beginning of each trial, an arrow appeared on the screen indicating the location of the target speech that the subject should attend to. After a one second pause, the stimuli were presented simultaneously. Subsequently, subjects should respond to two questions about the target passage. The aim of these questions was only to encourage the subjects to keep their attention to the target speech and the answers were not used for scoring purposes.

2.3.3. *Reconstruction of attended speech envelope*

In order to detect the auditory attention, we reconstructed the attended speech envelope from EEG signals. This requires establishing a model which maps the EEG signals to the

Table 1: *Demographic information of CI listeners*

Subject number	Age (years)	Gender	CI Model	CI use (years)	Speech processing strategy	Duration of HL (years)	Pure tone average of 500, 1000, and 2000 Hz (dB HL)
1	75	Male	Medel/ Sonnet	7	FS4	33	Right: NR /Left: NR
2	40	Female	Medel/ Sonnet	4	FS4	38.5	Right: NR /Left: NR
3	68	Male	Medel/ Sonnet	10	FS4	38	Right: NR /Left: NR
4	25	Female	Cochlear/Nucleus 6	21	ACE	25	Right: NR /Left: NR
5*	77	Female	Medel/Sonnet	70	FS4	6	Right: NR /Left: NR

* Subject 5 was not able to participate in EEG recording

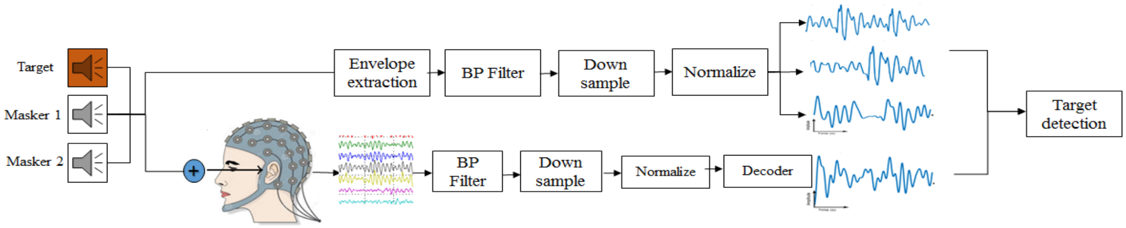

Figure 2: *EEG-based speech detection procedure*

corresponding speech envelopes. The following steps were initially taken to prepare the speech envelopes and EEG signals.

Speech envelope: The temporal envelopes of target and maskers stream were extracted using Hilbert transform, and band pass filtered to 0.3-30 Hz using filtfilt function. Speech envelopes were down sampled to 128 Hz and normalized to [0-1]. All the implementations were performed in MATLAB 2018b.

EEG: After preprocessing of the EEG data recorded from 64 channels and removing the artifacts using Independent Component Analysis (ICA), EEG signals were band pass filtered to 0.3-30 Hz using Fieldtrip toolbox [30]. The filtered EEG signals were subsequently down sampled to 128 Hz and normalized to [0-1].

A linear individualized model for each subject was trained using the speech envelope and corresponding EEG signals captured during the training sessions. We used mTRF toolbox (version 1.5) [31] to generate this model. The trained model was used to decode the EEG signal to reconstruct the speech envelope attended by the subject.

The linear mapping [31] from the instantaneous cortical response $r(t,n)$ to the speech envelope $s(t)$, sampled at times $t = 1 \dots T$ and at channel n, is represented as $d(\tau,n)$ which is the decoder that integrates the cortical response over a range of time lags τ. The linear convolution equation corresponding to this system could be expressed as:

$$\hat{s}(t) = \sum_n \sum_\tau r(t+\tau,n)\,d(\tau,n), \qquad (1)$$

Where the $\hat{s}(t)$ is the reconstructed speech envelope. The decoder is designed by minimizing the MSE between $s(t)$ and $\hat{s}(t)$.

$$\min \varepsilon(t) = \sum_t [s(t) - \hat{s}(t)]^2, \qquad (2)$$

Which results in the following formula for decoder:

$$d = (R^T R)^{-1} R^T s, \qquad (3)$$

In the above equation, R represents the lagged time series of neural response matrix r. To avoid overfitting of the model to the training data, Tikhonov regularization is applied on the optimization problem of equation 2, which results in:

$$d = (R^T R + \lambda I)^{-1} R^T s, \qquad (4)$$

In the above equation, λ is the regularization parameter.

To define the regularization parameter, we used leave one out cross validation. The training data (speech stimuli and corresponding EEG signal collected in training session) was split into K trials. To design a decoder that decodes trial k, all the other $K-1$ trials were used. We denote the preliminary decoders constructed from single trials as $\tilde{d}_i = (R_{(i)}^T R_{(i)} + \lambda I)^{-1} R_{(i)}^T S_{(i)}$, where $R_{(i)}$ and $S_{(i)}$ are the EEG response and stimuli of trial i, and $i \neq k$. The decoder d_k would be defined by averaging the preliminary decoders.

$$d_k = \frac{1}{K-1} \sum_{\substack{i=1 \\ i \neq k}}^{K} \tilde{d}_i \qquad k = 1, \cdots, K \qquad (5)$$

The correlation between the reconstructed speech envelope and the original speech envelope is a criterion to evaluate the accuracy of the decoder. The λ that resulted in maximum averaged correlation across K decoders was chosen to train individualized decoders. Here, we split the training data of each subject to 20 trials and examined a range of $\lambda = \{10^{-3}, 10^{-2}, \cdots, 10^{11}\}$. We defined individualized λ value for each subject using each individual's training data. The best λ for six NH listeners and two out of four CI listeners participated in electrophysiological test was 10^5 , and for the other two CI subjects was 10.

Having an individualized linear decoder built for each subject, the speech envelope attended by the subject was reconstructed from the EEG signal recorded during the test in spatial configuration. Pearson correlations between the reconstructed speech envelope and the envelope of each of the three stimuli (one target and two maskers) were calculated. The stimulus that yields the highest correlation with the reconstructed speech envelope was considered as the attended speech. Figure 2 summarizes this procedure. The ratio of the number of trials when the target stimulus was identified as the attended speech to the total number of trials is termed as the speech detection accuracy in this paper.

3. Results

3.1. Behavioral speech perception score

The speech perception scores were calculated based on the number of words correctly identified by the subjects over the total number of words presented. The scores are represented in percentage for CI and NH listeners in Figure 3-a, and Figure 3-b, respectively. For CI users, the speech perception score decreased as the target sound level increased in a constant TMR. The Analysis of Variance (ANOVA) was conducted on the speech perception score with a factor of target level. The results showed that there was a significant main effect of target level for CI users (F(2,2249)=15.98, p<0.01). A pairwise comparison with Bonferroni adjustment showed that speech perception score at 55 dB SPL target was significantly higher than that at 75 dB SPL target. The red asterisk in Figure 3 shows the significant difference between the speech perception score at different target levels.

The speech perception scores for NH group were remained nearly the same across all speech levels. The ANOVA test showed no significant effect of target speech level on their speech perception score (F(2,899)=0.13, p=0.88).

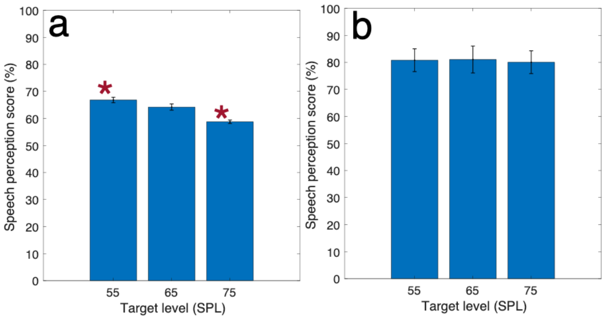

Figure 3: *Speech perception score for a) CI listeners, and b) NH listeners*

3.2. EEG based speech detection accuracy

Speech detection accuracy was computed at different target speech levels for CI and NH subjects (see Figure 4). Statistical analysis (ANOVA) shows a significant effect of target level on the speech detection accuracy in CI group (F(2,215)=7.17, p<0.01). Speech detection accuracy in CI decreases when the target level increase while TMR remains the same. The pairwise comparison with Bonferroni adjustment for CI group showed that the speech detection accuracy with target stimulus at 55 dB SPL was significantly higher than that at 75 dB SPL. But this trend was not seen within the NH group. (F(2,107)=2.31, p=0.1).

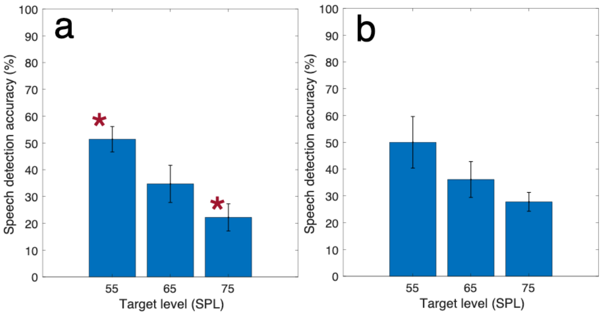

Figure 4: *EEG based speech detection accuracy for a) CI listeners, and b) NH listeners*

4. Discussion

Behaviorally, we found that CI's speech perception score decreases significantly when conversation speech levels increases, even though this trend was not found with NH's. Mixed results have been reported in the earlier studies. Firszt et al. [32] showed that CI listeners' speech recognition improves by increasing the speech level in quiet. However, our previous results [33] showed that increasing the speech level, in the presence of noise at constant SNR, negatively affects the perceived quality of speech by CI listeners.

Most CI users prefer a lower level of background noise even at the expense of lower speech intelligibility when engaged in their daily conversations. It is simply too difficult for CI listeners to attend to a target speech when masker level is high, even when the target and masker are spatially separated. By increasing the level of target speech while keeping the same TMR, our experiment results also showed that the preference of CI users was to lower the masker level rather than to raise the target level, particularly in spatial speech recognition task.

The outcome of the electrophysiological experiment was similar to that of the behavioral experiment. The speech detection accuracy computed for NH listeners decreases when the sound level increases, but the trend was not statistically significant. For CI users, the trend was statistically significant, which is also consistent with their behavioral outcome.

For both NH and CI listeners, the speech detection accuracy was around chance level (33%) at some speech levels. One possible reason may be attributed to the difficulty of the task. It may also suggest that more physiological metrics should be further explored to better quantify the auditory attention involved in spatial hearing.

5. Conclusion

Outcome of both our behavioral and physiological experiments support that the speech perception scores on target speech attended by CI users significantly decreases when the level of target speech increases while the TMR remains the same. CI users clearly prefer a lower noise levels at the expense of reduced speech audibility.

In the next phase of our study, we will also explore more behavioral and physiological measures that better discriminate the spatial speech perception performance between electrical and acoustic hearing by inclusion of acoustic hearing aid users.

6. References

[1] W. Biesmans, N. Das, T. Francart, and A. Bertrand, "Auditory-Inspired Speech Envelope Extraction Methods for Improved EEG-Based Auditory Attention Detection in a Cocktail Party Scenario," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 5, pp. 402–412, 2016.

[2] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.*, vol. 25, no. 5, pp. 975–979, 1953, doi: 10.1121/1.1907229.

[3] K. R. Pugh *et al.*, "Auditory Selective Attention : An fMRI Investigation OF ATTENTION," *Neuroimage*, vol. 4, no. 3, pp. 159–173, 1996.

[4] P. Lakatos, G. Musacchia, M. N. O. Connel, A. Y. Falchier, D. C. Javitt, and C. E. Schroeder, "The Spectrotemporal Filter Mechanism of Auditory Selective Attention," *Neuron*, vol. 77, no. 4, pp. 750–761, 2013, doi: 10.1016/j.neuron.2012.11.034.

[5] M. Giard, L. Collet, P. Bouchet, and J. Pernier, "Auditory selective attention in the human cochlea," *Brain Res.*, vol. 633, no. 1–2, pp. 353–356, 1994.

[6] O. Etard, M. Kegler, C. Braiman, A. Elia, and T. Reichenbach, "Decoding of selective attention to continuous speech from the human auditory brainstem response," *Neuroimage*, vol. 200, no. October, pp. 1–11, 2019, doi: 10.1016/j.neuroimage.2019.06.029.

[7] C. L. Mackersie, T. L. Prida, and D. Stiles, "The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss," *J. Speech, Lang. Hear. Res.*, vol. 44, pp. 19–28, 2001.

[8] N. Marrone, C. R. Mason, and G. Kidd, "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," vol. 124, no. 5, pp. 3064–3075, 2008, doi: 10.1121/1.2980441.

[9] V. Best, N. Marrone, C. R. Mason, G. Kidd, and B. G. Shinn-Cunningham, "Effects of Sensorineural Hearing Loss on Visually Guided Attention in a Multitalker Environment," *J. Assoc. Res. Otolaryngol.*, vol. 10, no. 1, pp. 142–149, 2008, doi: 10.1007/s10162-008-0146-7.

[10] L. Dai, V. Best, and B. G. Shinn-cunningham, “Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention,” *Proc. Natl. Acad. Sci.*, pp. 1–10, 2018, doi: 10.1073/pnas.1721226115.

[11] J. Murphy, A. Q. Summerfield, G. M. O. Donoghue, and D. R. Moore, “Spatial hearing of normally hearing and cochlear implanted children,” *Int. J. Pediatr. Otorhinolaryngol.*, vol. 75, no. 4, pp. 489–494, 2011, doi: 10.1016/j.ijporl.2011.01.002.

[12] M. Mok, L. Galvin, R. C. Dowell, and M. Mckay, “Spatial Unmasking and Binaural Advantage for Children with Normal Hearing , a Cochlear Implant and a Hearing Aid , and Bilateral Implants,” *Audiol. Neurotol.*, vol. 12, no. 5, pp. 295–306, 2007, doi: 10.1159/000103210.

[13] E. E. Bennett and R. Y. Litovsky, “Sound Localization in Toddlers with Normal Hearing and with Bilateral Cochlear Implants Revealed Through a Novel ‘“Reaching for Sound”’ Task,” *J. Am. Acad. Audiol.*, 2019, doi: 10.3766/jaaa.18092.

[14] H. M. Wright, W. Bulla, and E. W. Tarr, “Spatial release from masking and sound localization using real-time sensorineural hearing loss and cochlear implant simulation,” *J. Acoust. Soc. Am.*, vol. 145, no. 3, pp. 1877–1877, 2019.

[15] O. Strelcyk and T. Dau, “Relations between frequency selectivity , temporal fine-structure processing , and speech reception in impaired hearing,” *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3328–3345, 2009, doi: 10.1121/1.3097469.

[16] T. Y. C. Ching *et al.*, “Spatial release from masking in normal-hearing children and children who use hearing aids,” *J. Acoust. Soc. Am.*, vol. 129, no. 1, pp. 368–375, 2011, doi: 10.1121/1.3523295.

[17] F. J. Gallun, A. C. Diedesch, S. D. Kampel, K. M. Jakien, E. S. Sussman, and A. Einstein, “Independent impacts of age and hearing loss on spatial release in a complex auditory environment,” *Front. Neurosci.*, vol. 7, no. December, pp. 1–11, 2013, doi: 10.3389/fnins.2013.00252.

[18] T. L. Arbogast, C. R. Mason, G. Kidd, T. L. Arbogast, C. R. Mason, and G. Kidd, “The effect of spatial separation on informational and energetic masking of speech,” *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2086–2098, 2002, doi: 10.1121/1.1510141.

[19] G. Kidd *et al.*, “Stimulus factors influencing spatial release from speech-on-speech masking,” *J. Acoust. Soc. Am.*, vol. 128, no. 4, pp. 1965–1978, 2010, doi: 10.1121/1.3478781.

[20] H. Glyde, J. M. Buchholz, L. Nielsen, V. Best, H. Dillon, and S. Cameron, “Effect of audibility on spatial release from speech-on-speech masking,” vol. 138, no. 5, pp. 3311–3319, 2015, doi: 10.1121/1.4934732.

[21] S. Akbarzadeh, S. Lee, F. Chen, and C.T. Tan, “The effect of perceived sound quality of speech in noisy speech perception by normal hearing and hearing impaired listeners,” *2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 3119–3122, 2019, doi: 10.1109/EMBC.2019.8857952.

[22] C. Horton, R. Srinivasan, and M. D’Zmura, “Envelope responses in single-trial EEG indicate attended speaker in a ‘cocktail party,’” *J. Neural Eng.*, vol. 11, no. 4, p. 046015, 2014, doi: 10.1088/1741-2560/11/4/046015.

[23] N. Mesgarani and E. F. Chang, “Selective cortical representation of attended speaker in multi-talker speech perception,” *Nature*, vol. 485, no. 7397, p. 233, 2012, doi: 10.1038/nature11020.

[24] J. A. O. Sullivan *et al.*, “Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG,” *Cereb. Cortex*, vol. 25, no. 7, pp. 1697–1706, 2014, doi: 10.1093/cercor/bht355.

[25] Y. Wang, J. Zhang, N. Ding, J. Zou, and H. Luo, “Prior Knowledge Guides Speech Segregation in Human Auditory Cortex,” *Cereb. Cortex*, pp. 1–11, 2018, doi: 10.1093/cercor/bhy052.

[26] B. Mirkovic, S. Debener, M. Jaeger, and M. De Vos, “Decoding the attended speech stream with multi-channel EEG : implications for online , daily-life applications,” *J. Neural Eng.*, vol. 12, no. 4, p. 046007, 2015, doi: 10.1088/1741-2560/12/4/046007.

[27] S. Lee, S. Akbarzadeh, and C. T. Tan, “Neural Tracking of Band-Limited Sine-Wave Speech in Normal Hearing and Cochlear Implant Listeners,” *Int. J. Speech Lang. Pathol. Audiol.*, vol. 8, no. 1, pp. 38–49, 2020, doi: 10.12970/2311-1917.2020.08.06.

[28] R. M. Cox and C. G. Genevieve C. Alexander, “Development of the Connected Speech Test (CST).pdf,” *Ear Hear.*, vol. 8, no. 5, pp. 119S-126S, 1987.

[29] R. M. Cox and D. M. Mcdaniel, “Development of the Speech Intelligibility Rating (SIR) test for hearing aid comparisons,” *J. Speech, Lang. Hear. Res.*, vol. 32, no. 2, pp. 347–352, 1989.

[30] R. Oostenveld, P. Fries, E. Maris, and J. Schoffelen, “FieldTrip : Open Source Software for Advanced Analysis of MEG , EEG , and Invasive Electrophysiological Data,” *Comput. Intell. Neurosci.*, vol. 2011, p. 1, 2011, doi: 10.1155/2011/156869.

[31] M. J. Crosse, G. M. Di Liberto, A. Bednar, and E. C. Lalor, “The Multivariate Temporal Response Function (mTRF) Toolbox : A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli,” *Front. Hum. Neurosci.*, vol. 10, p. 604, 2016, doi: 10.3389/fnhum.2016.00604.

[32] J. B. Firszt *et al.*, “Recognition of Speech Presented at Soft to Loud Levels by Adult Cochlear Implant Recipients of Three Cochlear Implant Systems,” *Ear Hear.*, vol. 25, no. 4, pp. 375–387, 2004, doi: 10.1097/01.AUD.0000134552.22205.EE.

[33] S. Akbarzadeh, S. Lee, S. Singh, and C. T. Tan, “Implication of speech level control in noise to sound quality judgement,” in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2018, no. November, pp. 388–392, doi: 10.23919/APSIPA.2018.8659672.