AWARD NUMBER: W81XWH-19-1-0637

TITLE: Identifying the Sources of Degraded Speech-in-Noise Understanding and Individualized Therapeutic Options

PRINCIPAL INVESTIGATOR: Inyong Choi, Ph.D.

CONTRACTING ORGANIZATION: University of Iowa, Iowa City, IA

REPORT DATE: October 2021

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE October 2021	**2. REPORT TYPE** Annual	**3. DATES COVERED** 01Sep2020-31Aug2021
4. TITLE AND SUBTITLE Identifying the Sources of Degraded Speech-in-Noise Understanding and Individualized Therapeutic Options		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER W81XWH-19-1-0637
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Inyong Choi, Ph.D. Assistant Professor Department of Communication Sciences and Disorders E-Mail: inyong-choi@uiowa.edu		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Iowa 2 Gilmore Hall Iowa City, IA 52242-1320		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Development Command Fort Detrick, Maryland 21702-5012		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project will identify the peripheral and central factors that predict the variance of speech-in-noise (SiN) understanding ability among normal hearing listeners. To achieve this goal, in Aim 1, neural substrates of all the key stages along the proposed SiN processing model will be characterized through within-subject design experiments of human electrophysiology. In Aim 2, a correlational study, we will characterize a hierarchical regression model that adopts the measures of auditory neural processes as independent variables and electrophysiological responses and behavioral performance during phonological and lexical processing as dependent variables. Finally, Aim 3 will seek an optimal, clinically applicable set of test batteries and measures that identifies 1) the sources of degraded speech-in-noise understanding and 2) the most effective, individualized therapeutic options.

15. SUBJECT TERMS
None listed.

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unclassified	17	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION:

Even among young "normal hearing" people, there is great variation in the ability to understand speech in noisy real-world settings. However, current audiology practice does not have the methods to diagnose and treat speech-in-noise (SiN) problems in normal hearing listeners. This project will identify the peripheral and central factors that predict the variance of SiN understanding ability among normal hearing listeners. To achieve this goal, in Aim 1, neural substrates of all the key stages along the proposed SiN processing model will be characterized through within-subject design experiments of human electrophysiology. In Aim 2, our correlational study will characterize a hierarchical regression model that adopts the measures of peripheral encoding, auditory grouping, and selective attention processes as independent variables, cortical responses during speech unmasking as a mediator, and electrophysiological responses and behavioral performance during phonological and lexical processing as dependent variables. Finally, Aim 3 will seek an optimal, clinically applicable set of test batteries and measures that identifies 1) the sources of degraded speech-in-noise understanding and 2) the most effective, individualized therapeutic options. This study will characterize several key neural processes required for successful speech understanding in social settings, and will promote our understanding of how peripheral and central processing deteriorates in listeners with degraded ability.

2. KEYWORDS:

Speech in noise, peripheral auditory encoding, auditory grouping, auditory selective attention, phonological processing, lexical processing, electroencephalography (EEG), electrocorticography (ECoG)

3. ACCOMPLISHMENTS:

What were the major goals of the project?

This is the second-year annual report. Within the second year, there was no specific milestones / target dates listed in the approved SOW. Instead, the following activities were planned to be started and continued in the second year, aimed to be completed by the end of Year 3:

1. Aim 1: Within-subject study to characterize neural correlates of auditory grouping
2. Aim 1: Within-subject study to characterize neural correlates of selective attention
3. Aim 1: Within-subject study to investigate the effect of neurofeedback attention training on speech-in-noise perception
4. Aim 2: Correlational study to characterize the relationship between the fidelity of peripheral auditory encoding and speech-in-noise performance

5. Aim 2: Correlational study to characterize the relationship between auditory grouping ability and speech-in-noise performance
6. Aim 2: Correlational study to characterize the relationship between selective attention ability and speech-in-noise performance

What was accomplished under these goals?

1) Major activities

As it was planned, we have collected simultaneous behavioral (auditory grouping, selective attention, and speech-in-noise) and EEG data from 72 subjects, which completed the data collection goal at the end of Year 2. Currently we are at the 54.5% point of the total data collection goal.

Based on our interim results from the data we have collected so far, we submitted four manuscripts to peer-reviewed journals all of which were accepted or published. Key outcomes from the data analysis and the list of publications are described in the following sub-sections (including Section 3: Significant results and key outcomes).

2) Specific objectives

Objectives of the Aims 1 and 2 within- and across-subject studies were to identify neural markers of speech-in-noise variance in normal hearing listeners.

3) Significant results and key outcomes

For the Aim 1 within-subject studies, we explored the neural mechanisms of speaker identity-based selective attention during a speech-in-noise task.

Previous studies support this phenomenon with evidence for the presence of stronger speech-evoked brain activity (i.e., better release from maskgin) when listening to speech from a predicted voice. Based on the predictive coding theory, our within-subject study to characterize neural correlates of selective attention (Aim 1) investigated what neuro-cognitive processes are utilized while listeners prepare to recognize a target word spoken by a known voice or unknown.

We sought two critical components of predictive coding using EEG: The internal representation of the prediction and the incoming sensory inputs that update the prediction. Given that gamma-band oscillations reflect bottom-up processing for sensory sampling while beta-band oscillations reflect the top-down flow of prediction information, we hypothesized that tracking a target word spoken by a known voice will recruit stronger beta oscillation in auditory cognitive regions and weaker gamma oscillation in auditory sensory regions. To test this hypothesis, we measured 64-channel cortical EEG data from participants with normal hearing in a word-in-noise task with an auditory cue that either did or did not identify the upcoming speaker who will say a target word. The carrier phrase ("Choose the word") containing the auditory cue was spoken 1.5 seconds before the onset of the multi-talker babble noise that was continued for 2 seconds. In each trial, one of two types of auditory cues was randomly given: Either spoken by the same speaker (female or male) who will say a target word (henceforth referred to as "Cued" condition) or a mixture of both (female AND male) voices (i.e., "No-cue" condition). A target word was presented 1 second after the onset of the noise, which was 3 seconds after the auditory cue. After listening to all the sounds, four minimal-pair options appeared on the computer screen as the participant will have to select a key that corresponds to the target word. We analyzed event-related spectral perturbation during the period after the auditory cue and before the target word. The Cued condition exhibited stronger beta oscillations in the left inferior frontal gyrus than the No-cue condition, whereas the No-cue condition exhibited stronger gamma oscillations in the superior temporal gyrus of both hemispheres. The combination of these results supports the predictive coding theory and provides a potential neurophysiological substrate of speaker tracking during speech-in-noise recognition.

Figure 1 shows the trial structure of our speaker identity-based speech-in-noise experiment. At the start of every trial the subjects were asked to listen to the carrier phrase, "choose the word" which either was spoken by the speaker identity cue (single male or female voice) or by the ambiguous identity cue (both female and male voices at the same time). The carrier phrase would also prepare the subjects to listen for the target word presented in background noise. The carrier phrase was spoken 1.5 seconds before the onset of the multispeaker babbles noise that begun at 0 seconds. The target word that either matched or mismatched the auditory cue was presented in background noise at 1 second and last about .3-.4 seconds depending on the target word. Offset of noise and speech occurred at 2 seconds followed by silence.

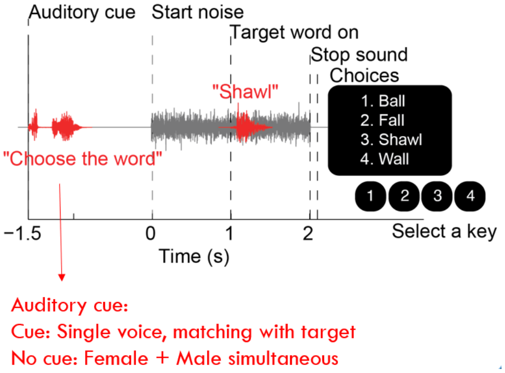

Figure 1. Trial structure; auditory cue or ambiguous cue was embedded in the carrier phrase, "choose the word" before the onset of noise. Multi-speaker babble noise started at zero seconds and target word that either matched or mismatched the auditory cue was presented in background noise. Offset of noise and speech was at 2 seconds. 4 minimal pair options appeared on the computer screen. The participant selected the key that corresponding to the target word they heard in the background noise.

Figure 2 shows the analysis across the time course from -1.5 seconds after the auditory cue to 1 second prior to the speech target (preparatory period). The three panels are grand-average magnitude spectrograms with time on the x axis, frequency on the y axis, and intensity is based on the relative darkness of the colors. The spectrograms are averaged across all 64 channels. Analyses were done across the different oscillation bands for both of the cued and non-cued condition. Additionally, the bottom panels show t-scores from a series of paired t-tests across all the time-frequency bins. T-scores are used to show the difference in effect size of the cued and no cued condition. The red color represents positive values which indicates that the cued condition had statistically significant and greater effects than the no cued condition. In contrast, the blue color is negative and shows that the no cued condition had statistically significant and stronger effects. The neutral green color surrounding the blue and red colors indicates p-values > 0.05 (uncorrected) and there was no significant difference. Referencing back to the first panel in Figure 2. there is a strong frequency band of a beta oscillation (~15 Hz) at .5 seconds in the cued condition. However, in the second panel spectrogram at 0 seconds there is a greater gamma band oscillation (~37-45 Hz). Therefore, in the last panel the gamma band oscillation was significantly larger for the no cue and the beta band oscillation was significantly larger than the cued condition. Black boxes indicate the clusters of time-frequency bins that exhibited significant differences between conditions from the cluster-based permutation analysis, indicating that the difference did not occur by coincidence at least in those black boxes. Both beta and gamma band oscillations are found in the "preparatory period" (time between the cue or no cue and the onset of speech).

For each cluster with a significant difference, paired t-tests were performed across all the cortical surface voxels to reveal source-level differences between conditions at each oscillation bands (i.e., gamma and beta). The cortical surface map at the bottom of Figure 2 shows the distribution of t scores. The gamma band oscillation that was significantly larger for the no cue condition was located in the supratemporal gyrus of both the left and right hemispheres of the brain. The beta band oscillation that was significantly greater for the cued condition was located in the left hemisphere only of the inferior frontal gyrus.

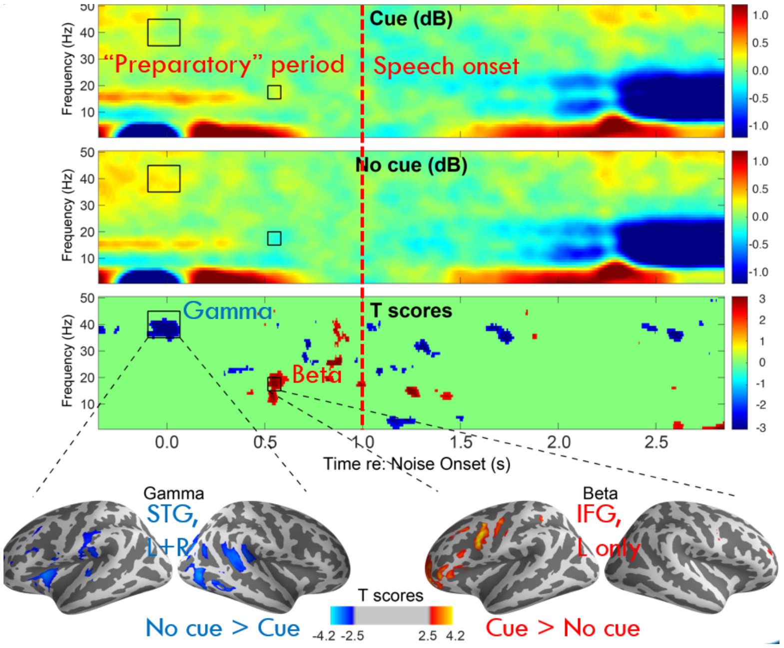

Figure 2. Spectrograms and cortical surface maps of induced activities. Black boxes around 0 and 0.5 seconds denote clusters of time-frequency bins that exhibited statistically significant differences between conditions. The first row: Grand-average spectrograms averaged across all 64 electrodes in the cue condition. The second row: Grand-average spectrograms averaged across all 64 electrodes in the no-cue condition. The between-condition t-scores in the third row is masked by p-value (<0.05, uncorrected). The cortical surface map at bottom shows the distribution of t-scores from paired t-tests performed within each cortical voxel between conditions within the significant clusters of the gamma (35 – 45 Hz, left panel) and beta (15-20 Hz, right panel) bands. Significant difference in gamma is found in

the supratemporal gyrus in both hemispheres and beta was only found in the left hemisphere of the inferior frontal gyrus.

These findings from the Aim 1 within-subject study provide neural markers of optimal brain activities that should look like in listeners with speech-in-noise deficits. In the following year, we will use this information to make clinical advancements to help mitigate normal-hearing listeners' communication barriers in noisy environments. These clinical implications will address their most frequent complaint and reduce their levels of strain and fatigue.

For Aim 2, to address our primary research question of characterizing contribution of speech unmasking processes to SiN performance, we conducted a correlation analysis in which the amplitude ratio between cortical responses to the target speech and background noise (i.e. henceforth referred to as "neural SNR") and normal-hearing listeners' benefit from reduced background noise combined with distorted speech spectrum (like most noise reduction: NR algorithms in hearing devices do).

People with higher noise tolerance (i.e., group 2 in Figure 3) who had higher neural SNR got less NR benefits in behavioral accuracy compared to the other group of people with lower noise tolerance (i.e., group 1 in Figure 3). For high-tolerant listeners, NR processing seemed redundant and potentially harmful due to distortions for speech cues. The current study results pointed out that NR might be redundant only for people with higher noise tolerance, who could unmask the target speech from background noise for themselves, suggesting the importance of examining individual noise tolerance in predicting any intervention benefits in each listener.

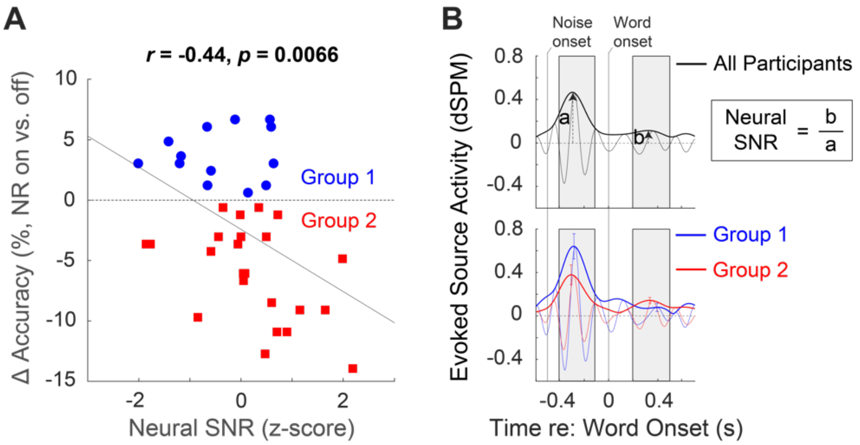

Figure 3. Individual differences in noise tolerance and noise reduction (NR) benefits. A. Individual noise tolerance indexed by the neural signal-to-noise ratio (SNR), the amplitude ratio of auditory-cortical responses to the target word and noise during speech-in-noise tasks, correlates with changes in behavioral accuracy due to NR. B. Illustration of the computation of neural SNR and differences in auditory-evoked responses between two groups divided based on NR-driven changes in behavioral accuracy. The thin lines indicate grand-average auditory-evoked responses, while the thick lines indicate their temporal envelope, with ±1 standard error of the mean at the peak-activity timing. dSPM: dynamic statistical parametric maps.

What opportunities for training and professional development has the project provided?

This project has been providing research experiences to one post-doctoral research scientist, one PhD student, two graduate (Doctor of Audiology) students, a post-bachelor research assistant, and four undergraduate students. The provided research experiences include EEG and behavioral data collection from human subjects, theoretical training of hearing science, and computational analyses of electrophysiological data.

How were the results disseminated to communities of interest?

Nothing to report.

What do you plan to do during the next reporting period to accomplish the goals?

In the next project year, we are planning to finish data collection for both Aim 1 within-subject and Aim 2 across-subject studies to find out neural markers that predict speech-in-noise performance and their relative contributions. Based on the findings, we will submit three additional manuscripts to peer-reviewed journals. The very next manuscript that describes neural mechanisms of target speaker tracking is currently being prepared for the submission. The second next manuscript will describe the individual differences in selective attention efficacy and its relationship with speech-in-noise performance. The third manuscript will expand our findings from single-word based speech-in-noise tasks to sentence-based tasks. We will also begin Aim 3 optimization studies.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

Our key findings so far demonstrated that the brain understands speech in background noise much more efficiently when it knows who will be speaking. We found a few regions in the brain that communicate each other to predict the acoustic features of forthcoming target sounds and correct any errors in the prediction by analyzing incoming sounds. We learned how to measure such brain processes using scalp electrodes. In the future we will be able use these measures as a clinical diagnostic tool, which will tell us why some listeners experience difficulty in understanding speech in background noise albeit their ears are normal.

What was the impact on other disciplines?

Nothing to report.

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Nothing to report.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them

Because human subject research was suspended university-wide due to the COVID-19 pandemic in Year 1, data collection was slower than expected. To achieve all the planned research goals with a less-than 12 month delay, we have updated the number of subject enrollment in Year 2. However, the human research data collection is still slow due to the continued social distancing. Under the revised plan, we will achieve the same 132 subjects by the end of Year 3. Then the final large-cohort regression analyses and disseminations of results will be performed in Year 4.

Changes that had a significant impact on expenditures

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Significant changes in use or care of human subjects

None

Significant changes in use or care of vertebrate animals

Not applicable.

Significant changes in use of biohazards and/or select agents

Not applicable.

6. PRODUCTS:

- **Publications, conference papers, and presentations**

  Journal publications. Four manuscripts have been submitted to peer reviewed journals; all of which have been accepted or published:

  1. Kim, S., Wu, Y.-H., Bharadwaj, H., Choi, I. (accepted). Effect of noise reduction on cortical speech-in-noise processing and its variance due to individual noise tolerance. Ear and Hearing. Acknowledgement of federal support: Yes.
  2. Geller, J., Holmes, A., Schwalje, A., Berger, J., Gander, P., Choi, I., McMurray, B. (accepted). Validation of the Iowa Test of Consonant Perception. Journal of the Acoustical Society of America, Acknowledgement of federal support: Yes. https://psyarxiv.com/wxd93/
  3. Kim, S., Emory, C., Choi, I. (Published: 2021). Neurofeedback training of auditory selective attention enhances speech-in-noise perception. Frontiers in Human Neuroscience, 15, 337. Acknowledgement of federal support: Yes. https://doi.org/10.3389/fnhum.2021.676992
  4. Kim, S., Schwalje, A., Liu, A., Gander, P., McMurray, B., Griffiths, T., Choi, I. (Published: 2021). Pre- and post-target cortical processes predict speech-in-noise performance. Neuroimage, 228, 117699. Acknowledgement of federal support: Yes. https://doi.org/10.1016/j.neuroimage.2020.117699

Books or other non-periodical, one-time publications.

Nothing to report.

Other publications, conference papers and presentations.

Presentations at the Association for Research in Otolaryngology (ARO) 44th Annual MidWinter Meeting (online):

1. Difference Analysis of Oscillatory Response to Continuous Speech Stimuli Between Native and Non-Native Speakers, Presenters/Authors: Luong Do Anh Quan, Youngmin Na, Inyong Choi, Jihwan Woo
2. Impact of Different Components on Spoken Sentence Comprehension Revealed by Speech-Evoked Electroencephalography, Presenters/Authors: Trang Le Thi, Youngmin Na, Inyong Choi, Jihwan Woo

- **Website(s) or other Internet site(s)**

Nothing to report.

- **Technologies or techniques**

Nothing to report.

- **Inventions, patent applications, and/or licenses**

.

Nothing to report.

- **Other Products**

  Nothing to report.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name: Inyong Choi
Project Role: PI
Researcher Identifier: ORCID 0000-0002-6663-9152
Nearest person months worked: 3
Contribution to Project: Conducting all the research activities.

Name: Phillip Gander
Project Role: Co-investigator
Researcher Identifier: ORCID 0000-0003-3945-8820
Nearest person months worked: 1
Contribution to Project: Developed stimuli and experiment scripts for the ECoG study.

Name: Adam Schwalje
Project Role: Resident
Nearest person months worked: 1
Contribution to Project: Developed speech materials for both EEG and ECoG studies.

Name: Sungyoung Kim
Project Role: Subaward PI
Nearest person months worked: 3
Contribution to Project: Developed a wireless EEG system (i.e., a preparation for Aim 3).

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to report.

What other organizations were involved as partners?

Organization Name: Rochester Institute of Technology
Location of Organization: Rochester, NY
Partner's contribution to the project: Collaboration.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

QUAD CHARTS:

9. APPENDICES: