Spring 2020

Effects of Speaker-Identity Cueing on Listening Effort During Speech-in-Noise

Meghan Wickham

Jason Geller
University of Iowa

Inyong Choi
University of Iowa

EFFECTS OF SPEAKER-IDENTITY CUEING ON LISTENING EFFORT DURING SPEECH-IN-NOISE

by

Meghan WickhamJason GellerInyong Choi

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in the Speech Pathology and Audiology

---

Inyong Choi
Thesis Mentor

Spring 2020

All requirements for graduation with Honors in the
Speech Pathology and Audiology have been completed.

Yu-Hsiang Wu
Speech Pathology and Audiology Honors Advisor

College of Liberal Art and Sciences

The University of Iowa

Iowa City, Iowa

EFFECTS OF SPEAKER-IDENTITY CUEING ON LISTENING EFFORT DURING SPEECH-IN-NOISE

By

Meghan Wickham

A thesis submitted in partial fulfillment of the requirements for graduation with Honors in the Department of Communication Sciences and Disorders

Dr. Inyong Choi

Thesis Mentor

Spring 2020

Abstract

The brain is an organ that performs a variety of intricate functions. Specifically, the brain has an amazing ability to recover the complexities of a speech signal within a mixture of sounds. The process of extracting the speech signal from background noise, however, is not necessarily straightforward or easy. Previous studies have developed the concept of "listening effort" as an umbrella term to include all cognitive demand listeners confront to understand speech. From a clinical standpoint, this term suggests that accuracy measurements alone are not sufficient, and a supplementary assessment of how hard a client must try in order to understand speech (especially when the speech is degraded due to background noise) must be conducted. Current research emphasis is on the post-speech-time compensatory processes in recovering speech cues. However, in this study, we claim pre-speech-time attentional processes also create a source of listening effort. To support this idea, we measured the cortical, behavioral, and pupillary responses of 19 normal-hearing participants to SiN conditions when speaker-identity cues were provided before speech. We found that such speaker-identity cues significantly increased alpha oscillations in fronto-temporal cortex during post-cue pre-target time. Cortical evoked responses to target speech exhibited significantly greater amplitude in the cued condition, indicating speaker-identity cues enhance attentional processes. Grand-mean pupil dilation was larger in the cued condition, albeit the difference was not significant. The speaker-identity cues did not alter accuracy significantly, which guaranteed that our comparisons on pupil and EEG responses were not affected by the ratio of correct trials in across-trial averages. Combining these results, we claim that listening effort is not always an inherently bad, fatiguing process, but rather, includes top-down brain mechanisms that help listeners better attend to a target speech signal in background noise.

Introduction

Understanding unclear speech is an ability that is critical for successful communication in our increasingly noisy, imprecise world. The clarity of a speech signal may become greatly compromised in environments with a high level of background noise. In listeners with hearing loss in noisy environments, the speech signal received is initially degraded by hearing loss or the limitations of a hearing device, and is further degraded by the background noise. As a result, individuals with hearing loss report high levels of fatigue and strain when listening to speech-in-noise (Finke, Büchner, Ruigendijk, Meyer, & Sandmann, 2016; Ohlenforst et al., 2017).

The ability to understand speech-in-noise is not only affected by the degradation of bottom-up acoustic cues, but also by top-down cognitive processes (Strait & Kraus, 2011). "Listening effort" is a loosely defined term that describes the top-down processes of listening as "the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task (Pichora-Fuller et al., 2016)." While bottom-up acoustic cues cannot necessarily be adjusted in a certain listening environment, the cognitive processes behind effortful listening are an area researchers seek to further characterize in order to address the age-old problem of listening fatigue and decreased accuracy in speech-in-noise conditions. This field of research aims to determine if being provided additional acoustic features of speech before physically hearing the noisy speech may help to improve accuracy and relieve listening effort. Researchers are conducting theoretical research to establish the neural correlates of listening effort, so that ultimately this knowledge may be applied in a clinical setting as an objective assessment or treatment tool for subjective reports of listening fatigue (Dimitrijevic, Smith, Kadis, & Moore, 2019).

Theoretical framework of attention and effort suggested by Kahneman (Kahneman, 1973) has served as a widely-accepted basis for the current model of cognitive effort. In this model, the term 'effort' is used interchangeably with the term 'attention,' and is explained to be an objective, cognitive attentional-effort (Bruya & Tang, 2018). The model outlines that allocation of attentional-effort changes with task difficulty, momentary intentions, and capacity limitations in the model. A feedback loop instantaneously directs the allocation of attentional-effort as a task unfolds. The allocation of attentional-effort is an involuntary process when performing a task, however, an individual retains the choice to discontinue partaking in the task at any point. In this instance, an individual can choose to allocate less effort, and therefore, suffer a decrease in performance. Conversely, allocating any more effort than demanded of the task is not possible. Lastly, Kahneman suggests that time-pressure is an important determinant of attentional-effort allocation in that it causes a sudden, increased allocation of attentional-effort to meet task demands (Kahneman, 1973, Chapter 2).

The attentional-effort required to understand speech-in-noise can be categorized as endogenous, externally-directed attention because inward attentional-effort is being applied to outward stimuli in a listener's adverse listening environment (Strauss & Francis, 2017). When listening to speech in a noisy background, the listener must first detect the sound, which immediately undergoes an initial grouping process based on spatial and temporal aspects of the signal. In the grouping process the speech stream is extracted from the background noise. Ascending the perception hierarchy, attentional-effort will now be allocated in the ensuing step, figural emphasis. At this level, the figure (speech) and ground (background noise) become distinct, and attentional-effort will affect which aspects of the figure get emphasized and further processed. The allocation of attentional-effort in this step is determined by innate dispositions to physical characteristics of the stimuli, collative factors of the stimuli, and selective intentions of the listener. Kahneman highlights studies demonstrating the deployment of attentional-effort during figural emphasis in the visual modality. Subjects showed improved accuracy scores when they were cued before the trial to pay attention to a designated object. In this way, certain features of a stimulus can be cued for emphasis during the figural emphasis stage of perception (Kahneman, 1973, Chapter 5).

Selective attention is a key process within figural emphasis. Kahneman explains, "It is reasonable to describe selective attention as a consistent emphasis on a class of perceived objects or perceived events in preference to others." It can be modulated to encoding of spatial and non-spatial acoustic features (Lee et al., 2012), reshape receptive fields in the primary auditory cortex (Fritz, Elhilali, David, & Shamma, 2007), and can be enhanced with stimulus continuity (Best, Ozmeral, Kopco, & Shinn-Cunningham, 2008). Work by Holmes et al. (E. Holmes, Kitterick, & Summerfield, 2018) demonstrated that endogenous preparatory auditory attention to a pre-target visual cue builds over time and reaches optimal configuration when the cue-target period was at least 2,000 ms. Further, accuracy in target selection improved as cue-target intervals neared 2,000 ms.

The effects of attentional-effort have primarily been studied using indirect measures such as pupillometry, subjective measures, and accuracy. The physiological measure of pupillometry has been used as an index of the effort exerted during a task, under the consensus that a larger pupil dilation indicates a more effortful task (Ohlenforst et al., 2017; Winn, Edwards, & Litovsky, 2015; Winn, Wendt, Koelewijn, & Kuchinsky, 2018), although pupillometry alone cannot reveal underlying neural substrates of effort that change quickly over time.

It has also been shown that a more direct measure, EEG, is sensitive to the allocation of attentional-effort across time. The amplitude of N1 and P2 event-related potential (ERP) components becomes greater when the sensory input is attended, and becomes smaller when the same input is ignored (Hillyard, Hink, Schwent, & Picton, 1973). Such modulation of N1-P2 amplitude is a consequence of attentional effort (Deng, Reinhart, Choi, & Shinn-Cunningham, 2019; Hansen & Hillyard, 1980). Further studies have shown that induced activity, especially in the alpha band, is indicative of effort (Dimitrijevic et al., 2019; Strauß, Wöstmann, & Obleser, 2014). Measures such as EEG and pupillometry can track cortical responses to stimuli across time and are used in this study to measure the allocation of attentional-effort across a time course of noisy speech.

Currently, the literature is rich with studies exploring attentional-effort in the visual modality, but less is known about the neuromodulation of attentional-effort in the auditory modality, specifically during a preparatory period following an auditory cue. Therefore, the aim of this study was to more concisely characterize the neural correlates of attentional-effort when pre-stimulus auditory cueing was provided. In this study, the allocation of attentional-effort was measured at the preparatory and target speech intervals when an auditory cue indicating speaker-identity was provided before the presentation of speech-in-noise. It was hypothesized that, in the cue-target interval, more effort (resulting in larger induced EEG activities) would be present when a speaker-identity cue was provided than when no cue was provided. We also hypothesized that, as the consequence of greater attentional effort, the cued condition would exhibit larger N1-P2 ERPs at the onset of the target speech in noise and smaller pupil dilations than when no cue was provided. Lastly, we hypothesized that accuracy would be greater when a cue was provided.

The pre-registration form for this experiment, which includes hypotheses, planned analyses, exclusion criteria, and sample size justification, can be found at: https://aspredicted.org/49e98.pdf

Methods

Stimuli

The stimulus words-in-noise in this study were selected from the Iowa Test of Consonant Confusion (ITCC) (A. Holmes, Geller, Schwalje, Choi, & McMurray, 2020). The ITCC is a four-alternative forced choice (4AFC) recognition task composed of 120 unique CVC rhyming keywords in confined "sets", or groups of four words with which they are always presented (e.g. ball, fall, shawl, and wall appear as choices when each word serves as the keyword). It is a phonetically-balanced, difficulty-balanced, and closed-set list.

The ITCC offers two male and two female voices for each test item. For this study, a female-voiced list and a male-voiced list were selected. The ITCC words were presented in multi-talker babble as the background noise with a fixed signal-to-noise-ratio(SNR) of 0 dB. The stimulus was presented in the sound field at 75db SPL.

Participants

Subjects with normal hearing and normal or corrected-to-normal vision were recruited through a mass email and by word of mouth from a population of students at a large Midwestern university. All subjects were volunteers who provided written informed consent and were compensated for their participation. Our screening criteria required each participant to i) be a native English speaker, ii) report no hearing threshold > 25dB at any tested frequency between 250 and 8000Hz during a hearing screening, iii) have no reported history of ADHD, brain injury, neurological conditions, or taking psychoactive medications (e.g., SSRIs). After the screening, 22 subjects participated in the experiment. After data collection, 3 subjects were excluded based on the following exclusion criteria: i) does not complete the entire experiment, ii) excessive data loss or artifacts (20% percent missing data), iii) speech-in-noise accuracy below 50% at 0 dB SNR. Finally, data from 19 subjects were used for further analyses. Participants (7 men, 12 women) had a mean age of 23 years old. The target sample size was determined by power analysis based on the pilot data. This study was approved by the University of Iowa Institutional Review Board.

Equipment

Subjects were fit with a 64 channel EEG cap from the BioSemi ActiView system. The Polhemus Patriot 3D scanner system was used to record the position of each electrode on individual listeners' heads. Subjects performed the task in a sound-treated, electrically shielded booth. The room was bright enough to ensure baseline pupil sizes were small. Subjects were seated in a wooden chair at a table, with the computer screen 90 centimeters in front of them at eye level. Their head was stabilized in a table-mounted chinrest 60 centimeters from the Eyelink 1000 desktop eye-tracker. Subjects responded to the task on a keypad that contained only keys denoting 1-4. The placement of the keys was manipulated by experiments so they formed a single horizontal row. Sound was presented from a single loudspeaker (model #LOFT40, JBL) placed 1 meter from the subject at a 0° azimuth angle. The task was implemented using Matlab.

Procedure

Scalp electrical activity (EEG) was recorded during the SiN task at a 2048 Hz sampling rate in the BioSemi ActiView system. Trigger signals were sent from Matlab to the BioSemi ActiView acquisition software. Sixty-four active electrodes were placed according to the international 10-20 configuration. Electrode positions on each listener's head were measured using Polhemus Patriot 3D scanner. Offset voltages were kept below 30mV.

The pupil data were collected with an EyeLink Duo. The distances and settings for the eye-tracking equipment were based on recommendations from the EyeLink 1000 Plus manual. Participants rested their chin in a chin-rest mount during the task to reduce movement. Eye-tracking data were collected at 500 Hz in monocular mode for the better-calibrated eye. Pupil tracking was set as ellipse and diameter. A standard nine-point calibration and validation was completed with drift correction every 30-50 trials. Calibration standards implemented were that average error was <0.30° and <1.00° max error.

In this task participants were asked to respond to a speech-in-noise task with which word they think they heard spoken in the noise. The two conditions in this study were cued speaker-identity and non-cued speaker-identity. In the cued condition, a cue phrase was spoken by either a single male or female voice, and that same voice spoke the following target word-in-noise. In the non-cued condition, a cue phrase was spoken by both the male and female voices at the same time, but only one speaker, either male or female, spoke the following target word-in noise.

Each participant was presented with each of the 120 items twice during the task: once in the cued condition and once in the non-cued condition. Within the cued condition, speaker-identity was counterbalanced across participants so that each item had been cued by the male speaker and the female speaker.

Figure 1 represents the trial structure. Each trial began with the presentation of a fixation cross ('+') on the screen to which listeners were asked to fix their gaze throughout the trial to minimize eye-movement artifacts. The task began with the cue phrase "choose the word," which enabled listeners to acquaint to the identity (or lack of identity) of the target word speaker and predict the timing of next acoustic event – the noise onset. After fixed-duration silence (3 seconds) that followed the cue phrase, multi-talker babble noise started and continued for 3 seconds. The target word was presented in the noise 2 seconds after the onset of the noise. The composite auditory stimulus (noise + word) was followed by a 3 second silent retention period to account for the sluggish pupil response. After the brief silent retention period, four foils were presented on the monitor, and subjects were instructed to select the word they heard using the keypad.

No feedback was given to subjects at the end of a trial. The next trial began 1 second after the button press. They had 10 seconds to respond before the trial advanced with no response. Subjects completed 240 trials with an optional break every 30 trials to prevent fatigue.

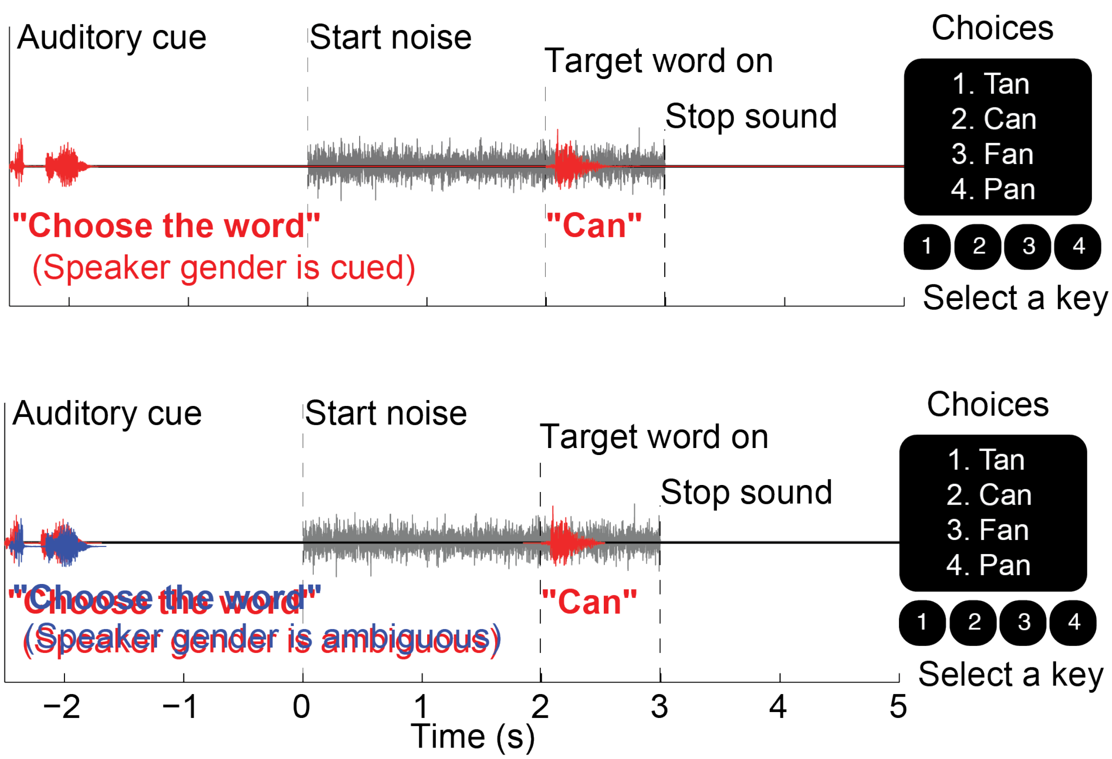

Figure 1: Trial structure of cued and non-cued conditions

Pre-Processing

Pupillometry

Pre-processing of the pupil data was done in R using the gazeR package (Geller, Winn, Mahr, & Mirman, n.d.). Trials with more than 20% data loss due to blinks and participants with more than 20% trial loss in either cueing condition were excluded. First, blinks were identified and extended 100 ms prior to and 100 ms after the period of missing data, and were then linearly interpolated. Next, a five-point moving average was used to smooth the data. The smoothed data was then normalized with subtractive baselining (Reilly, Kelly, Kim, Jett, & Zuckerman, 2019). The baseline for each trial was the median value of the 500 ms immediately preceding the beginning of the stimulus. Lastly, the data were time-binned to 50ms bins, reducing the sampling frequency from 500 Hz to 100 Hz.

EEG

The recorded EEG data from each channel were bandpass filtered from 1 to 20 Hz using a 2048-point FIR filter. Epochs were extracted twice per trial. The first epoch was taken from -500 ms to 2 s relative to the auditory cue onset. The second epoch started 500ms prior to noise onset and ended 5s after. After baseline correction using the average voltage between -200 and 0 ms, epochs were down-sampled to 256 Hz.

Eye blink-related artifacts were removed by signal space projection based on independent component analysis. After rejecting noisy epochs that exceeds 70μV instantly, 61 to 119 trials (mean 111 trials) were averaged to obtain event-related evoked potentials.

Accuracy

Accuracy of target word-in-noise identification was measured across the cued and non-cued condition. Responses were marked correct if the target foil was correctly identified from the four options, and responses were marked incorrect if any of the other three foils were selected.

Results

Effect of Auditory Cueing on Accuracy

Accuracy in determining the target word-in-noise was measured as the percent of correct responses out of all responses. A paired t-test revealed that there was no significant effect ($p=0.253$) of auditory speaker-identity cueing on accuracy of speech-in-noise identification (Figure 2). A ceiling effect was present (mean percent correct 79.83%, standard deviation 5.20% in cued condition and mean 79.02%, standard deviation 4.48 in non-cued condition). Given that the speaker-identity cues did not alter accuracy significantly, it is guaranteed that our comparisons on pupil and EEG responses were not affected by the ratio of correct trials in across-trial averages.

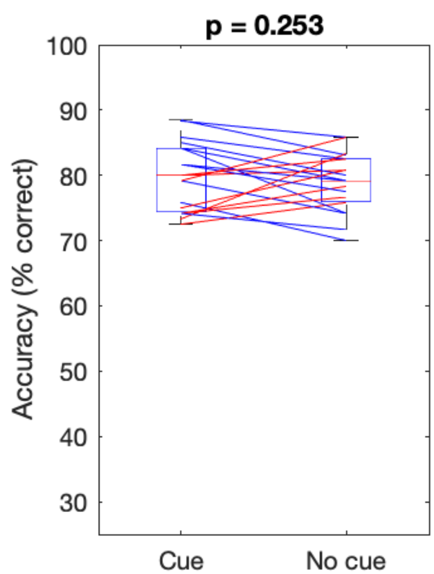

Figure 2: Effect of auditory cueing on identification of the target word-in-noise

EEG Analysis

Analysis across the time course from 0.5 seconds prior to noise onset (NO) to the retention period revealed 1) induced activities emerging from ~0,4 seconds prior to NO and 2) typical evoked ERPs at the onset (0s) and offset (3s) of the complex stimulus (Figure 3)(Figure 4).

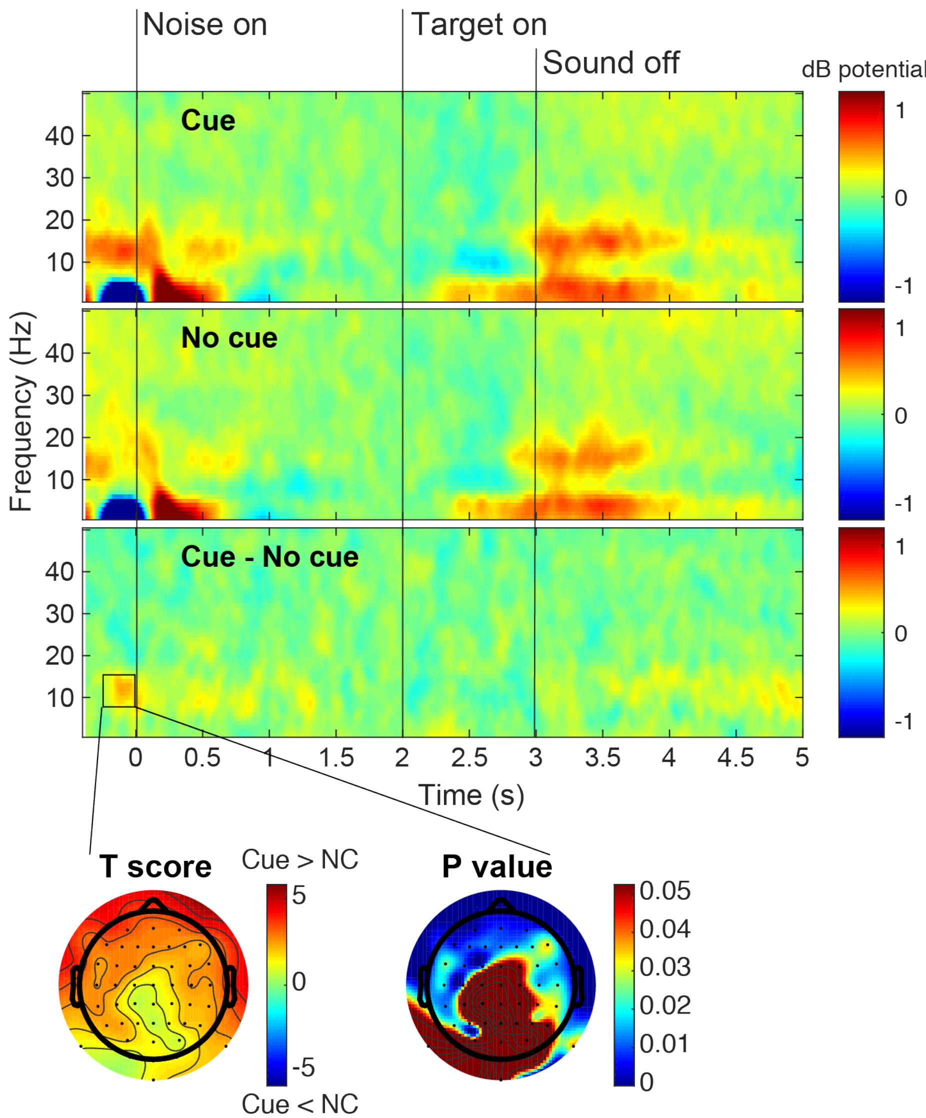

Figure 3: Spectrograms and topographies of induced activities. Statistically significant differences in alpha (~10Hz) band activity is found in post-cue, pre-stimulus period in frontal electrodes (i.e., stronger alpha in the cued condition, corrected p-values < 0.05 from paired t-test).

In the post-cue pre-stimulus period (i.e., -0.3 to 0.0 seconds in Figure 3 spectrograms), stronger alpha-band (~10Hz) induced activity was found at frontal electrodes when a cue was presented (paired t-test, FDR-corrected $p < 0.05$ in 21 out of 64 electrodes in both left and right frontal area: See t-value and p-value topographies at the bottom panel of Figure 3).

An additional ERP occurred at the onset of the target word-in-noise (Figure 4). The auditory N1 ERP component in the cued condition occurred at 2.25 s (i.e., ~250ms after target word onset) while the P2 ERP component was observed in both cued and un-cued conditions at 2.43 s (i.e., ~430ms after target word onset). The N1-P2 amplitude was stronger in the cued condition than the non-cued condition (paired t-test, $t(18) = 3.40$, $p=0.0032$) as shown in Figure 5. The average N1-P2 amplitude in the cued condition was 0.97μV compared to 0.27μV in the non-cued condition.

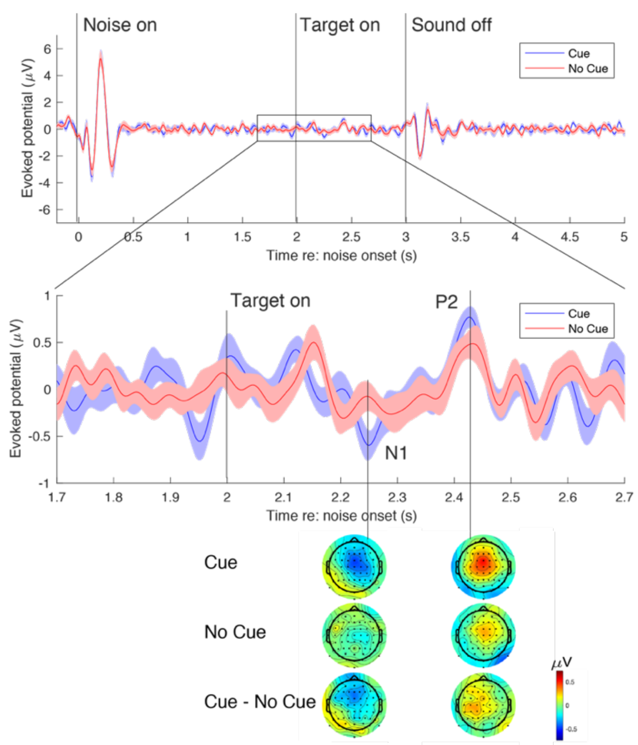

Figure 4: Evoked response potential across trial time course from noise onset to retention period. Expected ERPs at stimulus onset and offset were observed at 0s and 3s. Additional ERP to target-word-in noise observed at 2.3s represented by the highlighted portion of the time course. N1-P2 shown for cued condition in blue and non-cued condition in red.

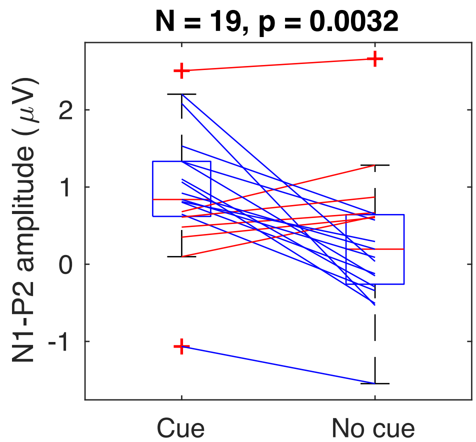

Figure 5: Difference in N1-P2 amplitude between cued and non-cued conditions. N1-P2 amplitude was stronger in the cued condition represented by blue lines (two-sided paired t-test, t(18) = 3.40, p=0.0032).

Pupillometry Analysis

Figure 6 presents the grand averaged pupillary time course from -100 ms before noise onset until the end of the trial. Looking at mean percent change in pupil size from baseline across that time region, no significant difference was observed between the cued and non-cued conditions, t(18) = 0.48, p = .638.

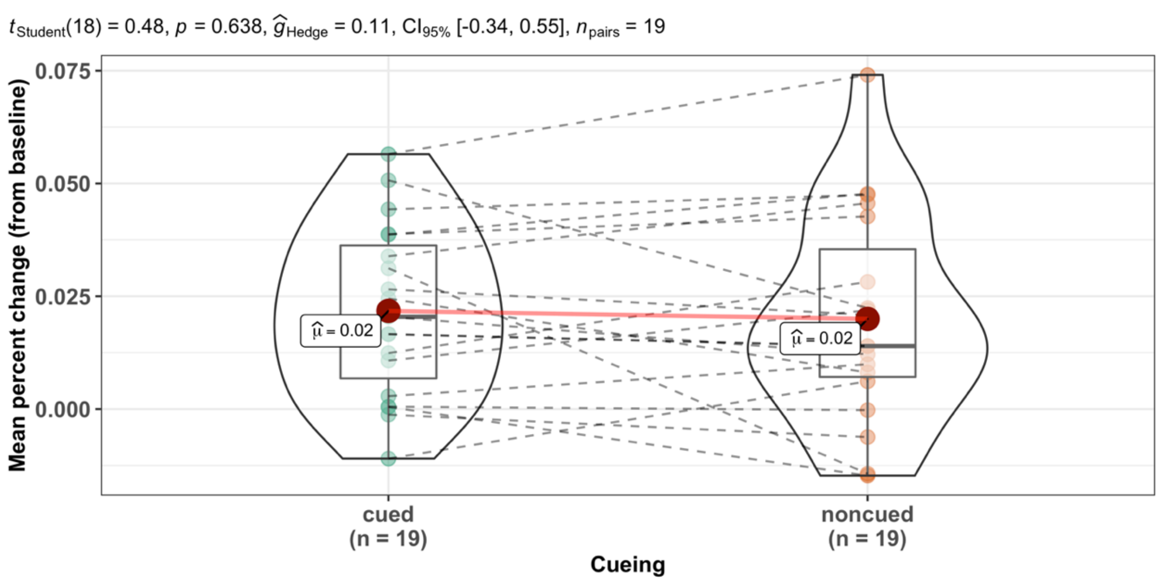

Figure 6: Mean percent pupil change from baseline at -100 ms before noise onset until the end of the trial. No significant difference was found between the cued and non-cued condition.

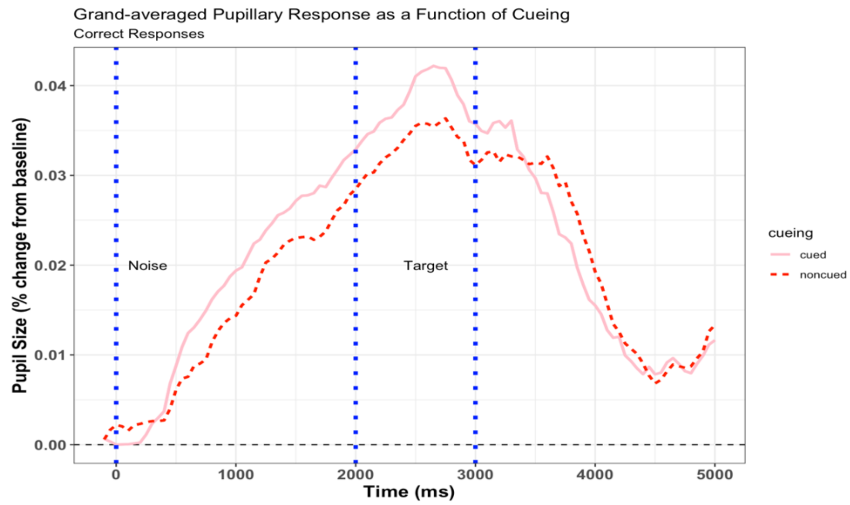

Figure 7: Pupil size percent change (from baseline) across the time course. The pink line indicates the percent pupil size change during the cued condition, and the dashed red line indicates the percent pupil size change during the non-cued condition. No significant difference was found (p=0.98).

Discussion

Results from this study revealed that speaker-identity auditory cueing is a source of preparatory attentional-effort, as indicated by increased induced activities during the cue-target interval in the cued condition (Figure 3). This preparatory attentional-effort in the cue-target period was shown to have an effect on later processing, causing a greater N1-P2 amplitude at the onset of noisy speech in the cued condition than the non-cued condition, indicative of increased attentional-effort (Figure 5). Surprisingly, the accuracy of identifying speech-in-noise was not improved by providing an auditory speaker-identity cue (Figure 2).

Results confirmed our hypotheses stating that providing an auditory cue will cause the allocation of attentional-effort before a stimulus is presented and will further increase attentional effort at the onset of noisy speech. These findings align with current work being done in the field of auditory selective attention and effort. Holmes et.al, 2018, suggests that auditory selective attention builds over time in the cue-target interval, which was demonstrated in this study by the induced activities occurring in the cue-target interval (Figure 3). The induced activities during this task occurred in the alpha band, which corroborates findings by Dimitrijevic et al, 2019, showing that greater oscillatory power change in the alpha band occurs in an attentive versus passive listening condition. Work by Best et al., 2008, shows that selective auditory attention can be enhanced by continuity of the target voice during successive stimulus presentations. This supports our finding of increased attentional-effort at the noisy speech in the cued condition because attention was enhanced as the target voice remained the same from the cue.

Findings from the current study contribute to the elaboration of Kahneman's widely-accepted model of perception and attention (Kahneman, 1973). This study has shown that auditory cueing does have an effect on the allocation policy of attentional-effort during the figural emphasis stage of perception. The attentional-effort allocated to the figure is determined by physical aspects of the stimulus—featural speaker-identity information, as well as intentions of the listener—wanting to understand the speech.

While accuracy differences between conditions were not found, this may further demonstrate the capacity limitations in the model of attentional-effort. With subjects scoring at ~80% accuracy on average, it is possible that minimal attentional-effort was required of subjects to complete this task. Given that the brain cannot allocate more effort than what is required of the task, it is possible that allocation of objective attentional-effort was not large enough to tax the capacity or further cause listener fatigue and accuracy differences. For this reason, future studies would utilize a more difficult SNR to eliminate ceiling effects. An alternate interpretation for the lack of accuracy differences could be due to the difference in sensitivity of the measures used. EEG is able to measure precisely time-locked cortical responses to stimuli, whereas accuracy simply measures an indirect behavioral response to the entire trial which may not be entirely reflective of the cognitive processes occurring.

Our hypothesis that smaller pupil dilation would be present after the target speech was not supported, and this has been similarly attributed to sensitivity of the measure to listening effort in this task. Pupillometry is an indirect measure of listening effort with a slow temporal response, so perhaps it was unable to reflect each stage of cognitive processing as the trial progressed. In an alternative hypothesis, we propose that pupil dilation is reflective of the sum of the preparatory attentional effort and the post-speech-time compensatory effort, while those two sources of effort are inconsistently modulated by the speaker-identity cueing. Previous studies have found little correlation between EEG and pupillometry measures during listening effort tasks, and suggest that these measures may reflect different cognitive processes involved in listening effort (Miles et al., 2017)(McMahon et al., 2016). This finding can be applied to the current study by attributing the insignificant pupillometry results to being a consequence of different cognitive processes occurring during the listening effort task.

The results of this study are limited by the small sample size. Underpowered analyses may have contributed to the insignificant pupillometry and accuracy results. Another limitation in this study was the SNR used. Future studies would utilize a more difficult SNR to reduce ceiling effects by taxing the attentional-effort capacity, without being so difficult the subject voluntarily withdraws effort. Task design would be adjusted to account for differences in temporal responses of pupillometry and EEG measures; thus, contributing to the expanding research field regarding best practices for measuring listening effort and teasing apart the various underlying cognitive processes.

The current task utilized a complex word-in-noise stimulus, but it would be worthwhile to utilize a greater structural level, such as a phrase- or sentence-in-noise, to make these research results more ecologically valid and applicable to a clinical population. Including a subjective measure of listening effort in later research will also be key in associating theoretical research of objective attentional-effort with the subjective listening effort of clinical populations. Future overall aims for this line of work include creating objective clinical assessment and treatment tools that validate the subjective listening experiences of nearly 37.5 million US adults reporting some degree of hearing loss (NIH, 2018).

Acknowledgements

This research was funded by the research grant awarded to Dr. Choi from the American Otological Society. Further funding was graciously provided by the Iowa Center for Research by Undergraduates.

We thank lab members Dr. Inyong Choi, PI, for his assistance in conducting EEG analyses and providing feedback that greatly improved the manuscript; Ann Holmes, lab manager, for her assistance in data collection; and Dr. Jason Geller, post-doctoral scholar, for his assistance with pupillometry analyses and study design.

References

Best, V., Ozmeral, E. J., Kopco, N., & Shinn-Cunningham, B. G. (2008). Object continuity enhances selective auditory attention. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(35), 13174–13178. https://doi.org/10.1073/pnas.0803718105

Bruya, B., & Tang, Y. Y. (2018). Is attention really effort? Revisiting Daniel Kahneman's influential 1973 book attention and effort. *Frontiers in Psychology*, *9*(SEP). https://doi.org/10.3389/fpsyg.2018.01133

Deng, Y., Reinhart, R. M., Choi, I., & Shinn-Cunningham, B. G. (2019). Causal links between parietal alpha activity and spatial auditory attention. *ELife*, *8*. https://doi.org/10.7554/eLife.51184

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2019). Neural indices of listening effort in noisy environments. *Scientific Reports*, *9*(1), 11278. https://doi.org/10.1038/s41598-019-47643-1

Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., & Sandmann, P. (2016). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. *Neuropsychologia*, *87*, 169–181. https://doi.org/10.1016/j.neuropsychologia.2016.05.019

Fritz, J. B., Elhilali, M., David, S. V, & Shamma, S. A. (2007). Auditory attention--focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*(4), 437–455. https://doi.org/10.1016/j.conb.2007.07.011

Geller, J., Winn, M. B., Mahr, T., & Mirman, D. (n.d.). GazeR_final. https://doi.org/10.31234/osf.io/gvcxb

Hansen, J. C., & Hillyard, S. a. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, *49*(3–4), 277–290. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6158404

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Holmes, A., Geller, J., Schwalje, A., Choi, I., & McMurray, B. (2020). Assessing the Reliability and

Validity of the Iowa Test of Consonant Confusion. In *Association for Research in Otolaryngology MidWinter Meeting*.

Holmes, E., Kitterick, P. T., & Summerfield, A. Q. (2018). Cueing listeners to attend to a target talker progressively improves word report as the duration of the cue-target interval lengthens to 2,000 ms. *Attention, Perception, and Psychophysics*, *80*(6), 1520–1538. https://doi.org/10.3758/s13414-018-1531-x

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Lee, A. K. C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-Cunningham, B. G. (2012). Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Frontiers in Neuroscience*, *6*(January), 190. https://doi.org/10.3389/fnins.2012.00190

McMahon, C. M., Boisvert, I., de Lissa, P., Granger, L., Ibrahim, R., Lo, C. Y., … Graham, P. L. (2016). Monitoring alpha oscillations and pupil dilation across a performance-intensity function. *Frontiers in Psychology*, *7*(MAY). https://doi.org/10.3389/fpsyg.2016.00745

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., de Lissa, P., Graham, P., & Lyxell, B. (2017). Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG. *Trends in Hearing*, *21*. https://doi.org/10.1177/2331216517706396

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., … Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing Research*, *351*, 68–79. https://doi.org/10.1016/j.heares.2017.05.012

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., … Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). In *Ear and Hearing* (Vol. 37, pp. 5S-27S). https://doi.org/10.1097/AUD.0000000000000312

Reilly, J., Kelly, A., Kim, S. H., Jett, S., & Zuckerman, B. (2019). The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behavior*

Research Methods, *51*(2), 865–878. https://doi.org/10.3758/s13428-018-1134-4

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(JUN). https://doi.org/10.3389/fpsyg.2011.00113

Strauß, A., Wöstmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, *8*(May), 350. https://doi.org/10.3389/fnhum.2014.00350

Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 809–825. https://doi.org/10.3758/s13415-017-0513-0

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, *36*(4), e153--e165. https://doi.org/10.1097/AUD.0000000000000145

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started. *Trends in Hearing*, *22*, 1–32. https://doi.org/10.1177/2331216518800869