PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

EEG alpha power as a measure of listening effort reduction in adverse conditions

Amy J. HALL[1]; Jan RENNIES-HOCHMUTH[1]; Axel WINNEKE[1]

[1] Fraunhofer IDMT, Branch Hearing, Speech and Audio technology, Oldenburg, Germany

ABSTRACT

Noise levels in everyday environments are typically low enough to comprehend speech, but over extended periods of time, listeners may experience tiredness and even fatigue. We measured the neurophysiological compensation required to separate speech from noise – listening effort – using electroencephalography, for normal and AdaptDRC enhanced speech.

AdaptDRC is a noise-dependent near-end-listening-enhancement algorithm which significantly improves the intelligibility of speech in noise and reduces subjectively rated listening effort (Rennies et al., 2018).

Participants (N=27) completed a listening effort and intelligibility task using a categorical listening effort scale, in which sentences were presented at five SNRs and in two types of noise.

Subjective listening effort was significantly lower, and speech intelligibility was significantly higher for AdaptDRC speech than unprocessed speech ($p<.001$). Analyses of the neurophysiological data show that there is a non-linear relationship between SNR and spectral alpha power with a peak between 0-5dB SNR, and that alpha power is sensitive to noise type and speech processing.

These findings provide insight into the neurophysiological correlates of listening effort and aid the development of an objective measure of cognitive load for speech in noise.

Keywords: Speech enhancement, listening effort, EEG

1. INTRODUCTION

Listening to and understanding speech is a task that normally hearing adults achieve effortlessly. However, in adverse conditions, the original speech signal may be distorted by reverberation, or disguised by noise. In such conditions, additional cognitive processing is required to separate the signal from the noise, requiring compensatory cognitive effort than listening to speech in quiet conditions, to maintain speech understanding at a high level. Pichora-Fuller et al. (1) defined cognitive effort as "the deliberate allocation of resources to overcome obstacles in goal pursuit when carrying out a task", with listening effort applying specifically to listening tasks.

Recent research using electroencephalography (EEG) has shown that neural oscillations in the alpha band region (8-12Hz) may modulate the suppression of task irrelevant information when listening to speech in noise (2). Winneke et al. (2016) found further evidence for this relationship; a hearing support algorithm significantly decreased spectral alpha power and subjective listening effort for call centre employees (3).

There is ongoing research regarding how listening effort is affected by speech enhancement algorithms (4). AdaptDRC is an example of a near-end-listening-enhancement algorithm, which adjusts the quality of a speech signal dependent on the noise in the environment (details in 5). The AdaptDRC algorithm increases the intelligibility of speech in noise in competing talker, speech-shaped noise, and cafeteria noise conditions (6,7). It also reduces the subjective experience of listening effort, even when word recognition performance was at ceiling (4), demonstrating that compensatory effort is required when listening conditions are intelligible but challenging.

In this experiment, participants performed a listening task while EEG was recorded. Participants were presented with sentences in background noise at different SNRs. 50% of trials were unprocessed and 50% were processed by the AdaptDRC algorithm.

[1] Amy.jane.hall@idmt.fraunhofer.de

2. METHODS

2.1 Participants

27 normally hearing adult native German speakers aged between 18-30 years (M = 23.4, SD = 2.6) were recruited for this study and paid €10 per hour for their participation. The experimental session lasted for a total of 2.5 hours. This research was approved by the Carl von Ossietzky Universität Oldenburg Kommission für Forschungsfolgenabschätzung und Ethik (Drs.81/2017).

2.2 Materials

Sentences were chosen from the Oldenburg Satztest (OLSA) set, which consists of five word sentences with the structure name-verb-number-adjective-noun, spoken by a male German speaker. At each position, there are ten possible words, giving rise to syntactically legal but semantically unpredictable combinations, for example 'Peter sieht acht nasse Steine' *Peter sees eight wet stones.*

Participants rated listening effort using a modified version of the Adaptive Listening Effort Scale (ACALES) (8), which is a 14-point scale with effort rating options ranging from 'mühelos' *effortless* to 'extrem anstrengend' *extremely effortful* and a further option of 'nur Störgeräusch' *just noise* for when they could not identify speech in the signal.

Speech was presented in two types of noise: speech shaped noise (SSN) and cafeteria noise. The SSN was created by layering the speech items, producing unintelligible noise with the same long term spectrum as the test stimuli. The cafeteria noise was recorded in a busy cafeteria and consists of the sound of many voices, chairs scraping, and cutlery being used (9).

Two types of speech were presented to participants: unprocessed or AdaptDRC processed speech. Details of the AdaptDRC algorithm can be found in Schepker et al., 2015 (5). The speech level was fixed at 60dB SPL and the noise level adjusted for each condition (-10, -5, 0, 5 & 10 dB SNR).

For each SNR there were four conditions (unprocessed+SSN, unprocessed+cafeteria noise, AdaptDRC+SSN, and AdaptDRC+cafeteria noise) giving a total of 20 conditions in a 5x2x2 design.

Sound mixing and speech enhancement processing was performed using MATLAB. Test stimuli were presented to participants over a single loudspeaker at 0° at a distance of 1.5m.

2.3 Experimental Procedure

2.3.1 Pure tone audiometry

Pure tone audiometry confirmed that participants had normal hearing (<10Hz HL).

2.3.2 Listening effort task

Throughout the listening effort task, participants kept their gaze directly ahead and kept movement to a minimum during the trials. There were 10 blocks of 30 trials, with a short break between blocks. Each trial consisted of 500ms of noise, then two consecutive sentences in noise, then another 500ms of noise, followed by a question. In 25/30 trials, participants were asked 'Wie anstrengend ist es für Sie, die Sprache zu verstehen?' *How strenuous is it for you to understand the speech? w*hich they answered using the ACALES scale. For 5/30 trials, participants were asked 'Bitte wiederholen Sie den soeben gehörten Satz und drücken Sie dann 'OK'.' *Please repeat the sentence you just heard, then press 'OK'* and they then repeated the second sentence of the pair presented in the trial. The experimenter recorded the number of correctly repeated words in the sentence (0-5). These word recognition trials were included to ensure that participants were actively listening to the speech and to provide speech intelligibility data for the stimuli. The listening effort task lasted between 50-60 minutes.

2.3.3 EEG data recording and analysis

EEG recording: EEG data were recorded at 500Hz using a 24-channel cap with Ag/AgCl electrodes arranged according to the 10-20 system (EasyCap) and a Smarting Streamer mobile amplifier (mBrainTrain). Stimuli onsets, behavioural responses, and EEG data recording were synchronised using Lab Streaming Layer (10). Data analysis was performed using EEGLAB (11) in MATLAB.

EEG data pre-processing: channel data were re-referenced offline to the average of the left and right mastoid electrodes and bandpass filtered (1-45Hz). Independent component analysis was run to identify and remove eye blink and muscle artefact components. Continuous data were divided into 5000ms epochs per trial from the onset of the first sentence and baseline corrected (-200-0ms).

Alpha power calculation: for each participant, per condition, a nonequispaced fast Fourier transform was performed on the epoched data and the average spectral power in the alpha band (8-12Hz) was calculated. This calculation was performed for every electrode.

3. RESULTS

3.1 Behavioural Results

Figure 1 shows speech intelligibility rates for all conditions. Percentage words correct significantly increased with increasing SNR in both noise types and in SSN there was a significant interaction between SNR and processing type, with AdaptDRC processed speech being more intelligible at low SNRs (3-way interaction SNR*processing type*noise type; $F(4,1268) = 29.93$, $p < .001$). Figure 2 shows the listening effort ratings for all conditions. There was also a significant three way interaction between SNR, processing type, and noise type ($F(4,1268) = 41.37$, $p < .001$) on listening effort ratings. For both noise types, AdaptDRC significantly decreased listening effort ratings across all SNRs, but more so at lower SNRs. Listening effort ratings were lower for SSN than for cafeteria noise.

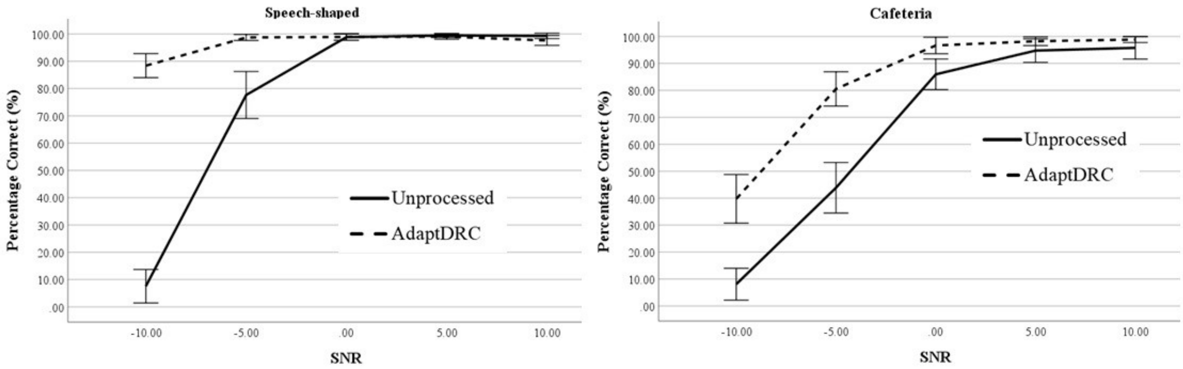

Figure 1 – Mean words correct in percent. Error bars show 95% confidence intervals

3.2 EEG Results

Six participants were excluded from EEG data analysis after being confirmed as statistical outliers (defined as points more than or less than two standard deviations from the mean). Results from each noise type were analysed separately.

SSN: a repeated measures 5x2 ANOVA with the factors SNR and processing type revealed a main effect of SNR on mean spectral alpha power (8-12Hz) at electrode CPz ($F(4,16) = 4.40$, $p = .02$), with spectral alpha increasing with increasing SNR. There was no main effect of processing type.

Cafeteria noise: a repeated measures 5x2 ANOVA with the factors SNR and processing type showed a main effect of processing type on mean spectral alpha power at electrode CPz ($F(1,19) = 7.61$, $p = .01$), showing higher alpha power for AdaptDRC processed speech.

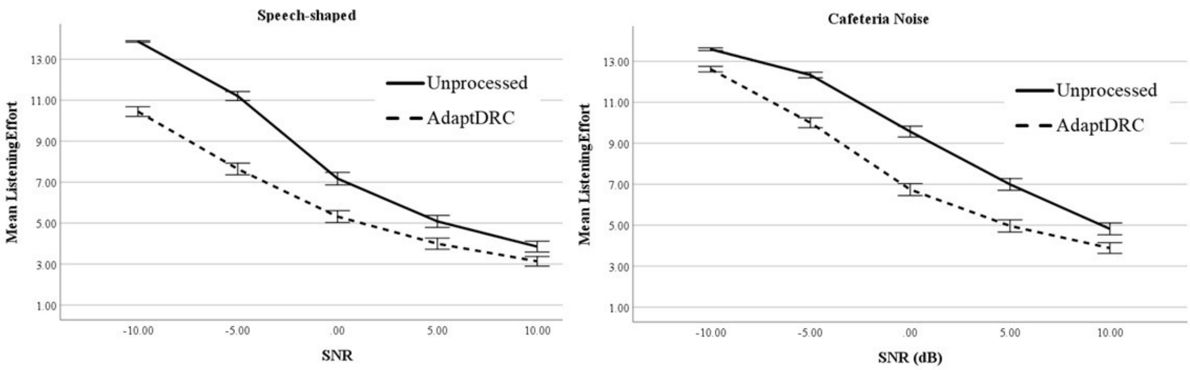

Figure 2 – Mean listening effort in Effort Scaling Categorical Units. Error bars show 95% confidence intervals

4. DISCUSSION

The behavioural data presented here replicate previous findings which show that AdaptDRC improves intelligibility and reduces subjective effort for speech presented in noise. Listening effort was reduced even at SNRs for which speech intelligibility was at 100%, providing further evidence for a compensatory cognitive mechanism that facilitates excellent speech understanding in noise.

AdaptDRC and SNR did not have the same effect on mean spectral alpha power for SSN and cafeteria noise. There was no effect of speech processing for SSN, perhaps due to the close to ceiling performance, even at -10dB SNR. There was an effect of SNR: alpha power increased with increasing SNR. This finding is contrary to previous data which has shown that spectral alpha is lower in easier listening conditions. For cafeteria noise, there is an effect of speech processing, with higher spectral alpha values for AdaptDRC processed speech, and no effect of SNR. This finding does not support the hypothesis that cognitive effort is lower for more easily intelligible speech, and reflected by decreased alpha power. A possible explanation for this finding is that synthetic or artificial sounding speech increases cognitive load, even if it is more intelligible than natural unprocessed speech (12).

5. FINAL REMARKS

There were very large individual differences in baseline alpha power. Thus, further analyses are required to establish the true effects of differential cognitive processing requirements and listening effort on spectral alpha power, by controlling for these individual differences. Future studies will further explore the relationship between the cognitive effort required to understand speech in noise and how it can be measured using EEG and other established methods such as pupillometry.

ACKNOWLEDGEMENTS

ENRICH has received funding from the EU’s H2020 research and innovation programme under the MSCA GA 675324

REFERENCES

1. Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BW, Humes LE, Lemke U, Lunner T, Matthen M, Mackersie CL, Naylor G. Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). Ear and Hearing. 2016 Jul 1;37:5S-27S.
2. Strauß A, Wöstmann M, Obleser J. Cortical alpha oscillations as a tool for auditory selective inhibition. Frontiers in human neuroscience. 2014 May 28;8:350.
3. Winneke A, Meis M, Wellmann J, Bruns T, Rahner S, Rennies J, Walhoff F, Goetze S. Neuroergonomic assessment of listening effort in older call center employees. In Zukunft Lebensräume Kongress 2016 (pp. 20-21).
4. Rennies J, Pusch A, Schepker H, Doclo S. Evaluation of a near-end listening enhancement algorithm by combined speech intelligibility and listening effort measurements. The Journal of the Acoustical Society of America. 2018 Oct 23;144(4):EL315-21.
5. Schepker H, Rennies J, Doclo S. Speech-in-noise enhancement using amplification and dynamic range compression controlled by the speech intelligibility index. The Journal of the Acoustical Society of America. 2015 Nov 3;138(5):2692-706.
6. Cooke M, Mayo C, Valentini-Botinhao C. Intelligibility-enhancing speech modifications: the hurricane challenge. InInterspeech 2013 Aug (pp. 3552-3556).
7. Schepker HF, Rennies J, Doclo S. Improving speech intelligibility in noise by SII-dependent preprocessing using frequency-dependent amplification and dynamic range compression. In INTERSPEECH 2013 Aug (pp. 3577-3581).
8. Krueger M, Schulte M, Brand T, Holube I. Development of an adaptive scaling method for subjective listening effort. The Journal of the Acoustical Society of America. 2017 Jun 22;141(6):4680-93.
9. Kayser H, Ewert SD, Anemüller J, Rohdenburg T, Hohmann V, Kollmeier B. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. EURASIP Journal on Advances in Signal Processing. 2009 Jan 1;2009:6.
10. Kothe C. Lab streaming layer (LSL). https://github.com/sccn/labstreaminglayer. Accessed on October. 2014;26:2015.
11. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods. 2004 Mar 15;134(1):9-21.
12. Govender A, King S. Using pupillometry to measure the cognitive load of synthetic speech. System. 2018;50:100.