Frontal midline θ power as an index of listening effort

Matthew G. Wisniewski[a], Eric R. Thompson[b], Nandini Iyer[a], Justin R. Estepp[a], Max N. Goder-Reiser[a] and Sarah C. Sullivan[c]

Attempts to identify physiological correlates of listening effort have mainly focused on peripheral measures (e.g. pupillometry) and auditory-evoked/event-related potentials. Although nonauditory studies have suggested that sustained time–frequency electroencephalographic (EEG) features in the θ-band (4–7 Hz) are correlated with domain-general mental effort, little work has characterized such features during effortful listening. Here, high-density EEG data was collected while listeners performed a sentence-recognition task in noise, the signal-to-noise ratio (SNR) of which varied across blocks. Frontal midline θ (Fmθ), largely driven by sources localized in or near the medial frontal cortex, showed greater power with decreasing SNR and was positively correlated with self-reports of effort. Increased Fmθ was present before speech onset and during speech presentation. Fmθ power also differed across SNRs when including only trials in which all words were recognized, suggesting that the effects were unrelated to performance differences. Results suggest that frontal cortical networks play a larger role in listening as acoustic signals are increasingly masked. Further, sustained time–frequency EEG features may usefully supplement previously used peripheral and event-related potential measures in psychophysiological investigations of effortful listening. *NeuroReport* 26:94–99 Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

NeuroReport 2015, **26**:94–99

Keywords: anterior cingulate cortex, audiological rehabilitation, auditory, event-related spectral perturbation, independent components analysis

[a]711th Human Performance Wing, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH, USA, [b]Ball Aerospace Technologies Ball Aerospace and Technologies Corp., Fairborn, OH, USA and [c]DOD Hearing Center of Excellence, Lackland Air Force Base, TX, USA

Correspondence to Matthew G. Wisniewski, PhD, Building 441, Area B, 711th Human Performance Wing, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH 45433, USA
Tel: + 937 255 3169; fax: + 937 656 7680; e-mail: matt.g.wisniewski@gmail.com

Received 3 November 2014 **accepted** 27 November 2014

Introduction

Listening to degraded target acoustic signals generally requires greater effort than listening to signals that are clear. Evidence stems from self-reports of the hearing-impaired [1] and experiment participants performing listening tasks at varying levels of signal degradation [2]. In addition, dual-task studies have found increased secondary task detriment when a primary listening task is made increasingly difficult [3,4]. This detriment may quantify the cognitive resources needed for listening, assuming that the tasks compete for resources [1,3,4]. Similar claims have been made using pupillometry, skin conductance, and auditory-evoked/event-related potentials (AEP/ERPs), with larger responses associated with greater effort [2,5–7].

Secondary task detriment occurs for nonauditory secondary tasks (e.g. tactile pattern recognition [3]), pointing to the use of resources at a higher than sensory level. In addition, the features of psychophysiological measures associated with listening effort are correlated with demands placed on attention and working memory [8,9]. Presumably, these processes rely heavily, although not exclusively, on frontal and parietal cortical networks [10], which modulate and/or generate such responses [8,11,12]. Greater use of these networks may be important for active 'listening', which involves intention and the expenditure of mental/cognitive effort, in contrast to 'hearing', which can be thought of as a passive process [1]. Little work has explored direct measures of activity in these networks as they relate to degrees of effortful listening.

Here, we examined the dynamics of the 4–7 Hz frontal midline θ (Fmθ) rhythm in the electroencephalogram (EEG), which is largely driven by medial frontal sources and is well demonstrated to vary in power as a function of attention, memory load, and task difficulty [13–15]. On the basis of this research and PET and functional MRI studies associating activation of medial frontal areas with attentive listening under adverse listening conditions [16], it was hypothesized that Fmθ power would increase with decreasing SNR and would correlate positively with listeners' reports of effort in a sentence-recognition task. Our goal was to determine whether analyses of Fmθ could usefully supplement other psychophysiological measures of listening effort to further our understanding of the brain dynamics that support successful listening. In addition, if Fmθ power is a reliable index of listening effort, it could potentially be useful in quantifying the benefits of audiological rehabilitation techniques [1–6].

Methods

Listeners

Fourteen adults (five male; age: M = 23.5, SD = 4.0) were paid for participation. All had normal hearing and

---

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (*www.neuroreport.com*).

DOI: 10.1097/WNR.0000000000000306

provided written informed consent. One participant was dropped for excessive movement artifacts. The US Air Force Research Laboratory Institutional Review Board approved the study.

Stimuli

A 40-word corpus containing five categories of words (names, verbs, numbers, adjectives, and nouns) and 18 talkers (nine male) was used [17]. Any selection of words in name, verb, number, adjective, and noun order yields a syntactically correct, yet unpredictable, sentence (e.g. 'Lynn took five blue shoes'; see Fig. 1 for all words). A word from each category and a single talker were randomly selected on each trial. Speech-shaped noise masked speech at −12, −6, 0, 6, and 12 dB SNRs by modifying the noise level.

Apparatus

Sounds were presented diotically through earphones (ER-2; Etymotic Research, Elk Grove Village, Illinois, USA). Procedures were executed in MATLAB (Mathworks, Natick, Massachusetts, USA). Listeners sat in a sound booth.

Procedures

Trials included five intervals (Fig. 1): pretrial (1.5 s), warning (mask only, 3.5 s), stimulus (speech and mask, ~3 s), retention (3 s), and response (≤9 s). During the response interval, all corpus items appeared on an 8×5 grid (word×category). Listeners were given 9 s after the onset of the response interval to click on words matching those that they had heard. They were told not to move the mouse until words appeared on the screen. SNRs were presented in separate randomly ordered blocks of 35 trials (5 blocks, 175 total trials). After a block, listeners were asked 'How hard did you have to listen to accomplish your level of performance in that block?' Response options ranged from 1 (not very hard) to 5 (very hard).

Electrophysiology

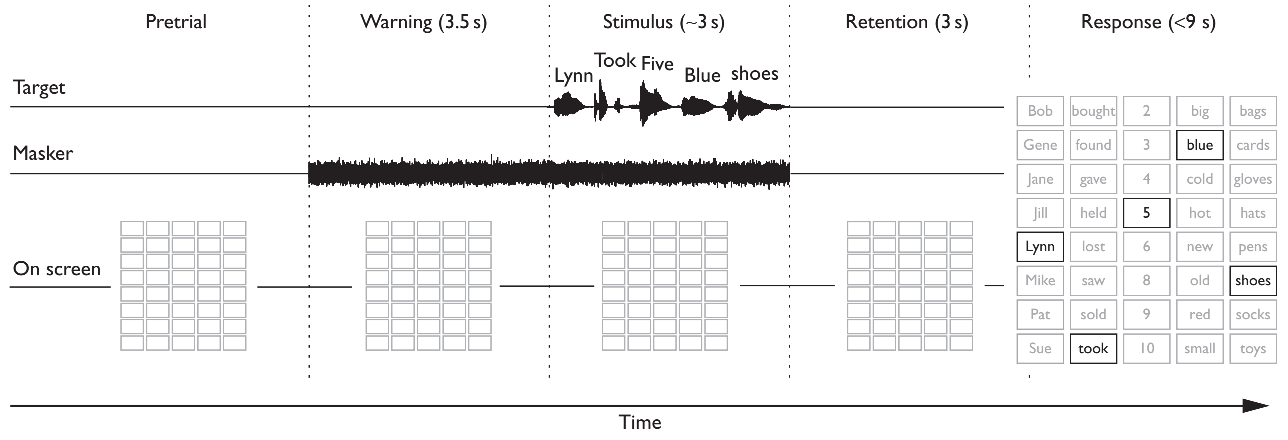
Data (132 channels) were collected at a 2048 Hz sample rate, 24 bit A/D resolution, and were referenced to CMS-DRL of a BioSemi Active II system (BioSemi, Amsterdam, the Netherlands). Scalp electrodes (128) were arranged in a BioSemi equiradial layout. Four electrodes were placed at the outer canthi and inferior orbits.

Offline analyses were carried out using EEGLAB ([18]; *http://sccn.ucsd.edu/eeglab*) and custom MATLAB scripts. Data were resampled at 256 Hz, band-pass filtered between 0.5 and 100 Hz, then re-referenced using an average reference. Portions of data and channels contaminated by artifacts (e.g. muscle noise) were rejected.

The remaining data were submitted to Infomax independent components analysis (ICA [18,19]) using the binica() function in EEGLAB. ICA extracts a set of components accounting for channel data that are spatially fixed and temporally independent. ICA was used to remove artifactual source contributions to channel data (e.g. eye movements) and to analyze Fmθ-related brain processes separately from others. The latter method is particularly useful for addressing our hypotheses as ICA algorithms consistently extract independent components (ICs) that capture a large amount of Fmθ variability [14,18,19]. Standard extended mode binica() parameters were used.

IC processes identified as artifacts by spectra and scalp maps were removed from channel data by backprojecting nonartifactual source activations to channels [18]. The

Fig. 1



Bob	bought	2	big	bags
Gene	found	3	blue	cards
Jane	gave	4	cold	gloves
Jill	held	5	hot	hats
Lynn	lost	6	new	pens
Mike	saw	8	old	shoes
Pat	sold	9	red	socks
Sue	took	10	small	toys

Depiction of a typical trial. The response grid did not become larger in the response interval, but is enlarged for readability in the figure. All 40 items of the corpus are shown on the grid.

remaining ICs were fit with single equivalent current dipoles using the dipfit() function in EEGLAB [18].

Independent component clustering/selection

IC processes fit with dipoles in the brain that explained at least 85% of the variance in scalp projections were clustered into 10 groups on the basis of dipole locations, scalp maps, and spectra, using an automated K-means procedure. The IC process cluster showing the clearest θ activity was selected for analysis [14].

Event-related spectral perturbations

Epochs were extracted from − 5 to 10 s relative to stimulus interval onset. A trial's event-related spectrum was computed using Morlet wavelets, returning power at 193 linearly spaced frequencies from 2 to 50 Hz, centered at 200 time points. Single trials were linearly time-warped to produce equal numbers of data points between the onset and offset of the stimulus interval. Warping assured that the offset of short speech sequence durations did not induce spectral perturbations that could be interpreted as occurring within the stimulus interval. The spectrum from the prestimulus period (mean of all trials) was used as a common divisive baseline [14,18,19].

Results

Behavior

All analyses of variance (ANOVAs) in which Mauchly's test of sphericity was significant were interpreted with Hyunh–Feldt corrections (uncorrected *d.f.*s reported in the text). The mean proportions of correct words reported were as follows: 0.49 (− 12 dB), 0.88 (− 6 dB), 0.96 (0 dB), 0.97 (6 dB), and 0.97 (12 dB). A repeated measures ANOVA found a significant main effect of SNR [$F(4,48) = 323.26$, $P < 0.001$, $\eta_p^2 = 0.96$], supporting accuracy differences. Accuracy was high and relatively flat between 0 and 12 dB. Clear drops in accuracy did not appear until − 6 dB, with a sharp drop at − 12 dB. Generally, reports of effort decreased as SNR increased: 4.77 (− 12 dB), 3.39 (− 6 dB), 2.54 (0 dB), 1.15 (6 dB), and 1.61 (12 dB). Supported by a main effect of SNR [$F(4,48) = 83.76$, $P < 0.001$, $\eta_p^2 = 0.88$]. Less effort was reported at a 6 dB SNR compared with a 12 dB SNR. However, both were rated lower than other SNRs.

Electrophysiology

The grand-average event-related spectral perturbation (ERSP; mean across listeners and SNRs) shows several time–frequency features at electrode Fz (Fig. 2a). A short burst of low-frequency power follows the warning interval onset, likely associated with the ERP to masker onset. Power decreases in the α (8–12 Hz) and β (12–30 Hz) bands start during the retention interval and proceed through the trial. This likely reflects μ rhythm desynchronization associated with motor aspects of the task [19]. There may also be α-related power decreases reflecting other attentional and/or memory-related processes [21]. Our intended focus, a sustained increase in Fmθ power, appears to be present at Fz, possibly accompanied by increases in low-β power.

IC processes grouped into several prototypical clusters (e.g. μ rhythm, occipital α, parietal α) including an Fmθ cluster. Clusters of IC processes can, and did, contain different numbers of ICs per individual. To avoid disproportionate listener contribution, only one IC process in the Fmθ cluster per listener was analyzed (IC explaining the most variance in channel data).

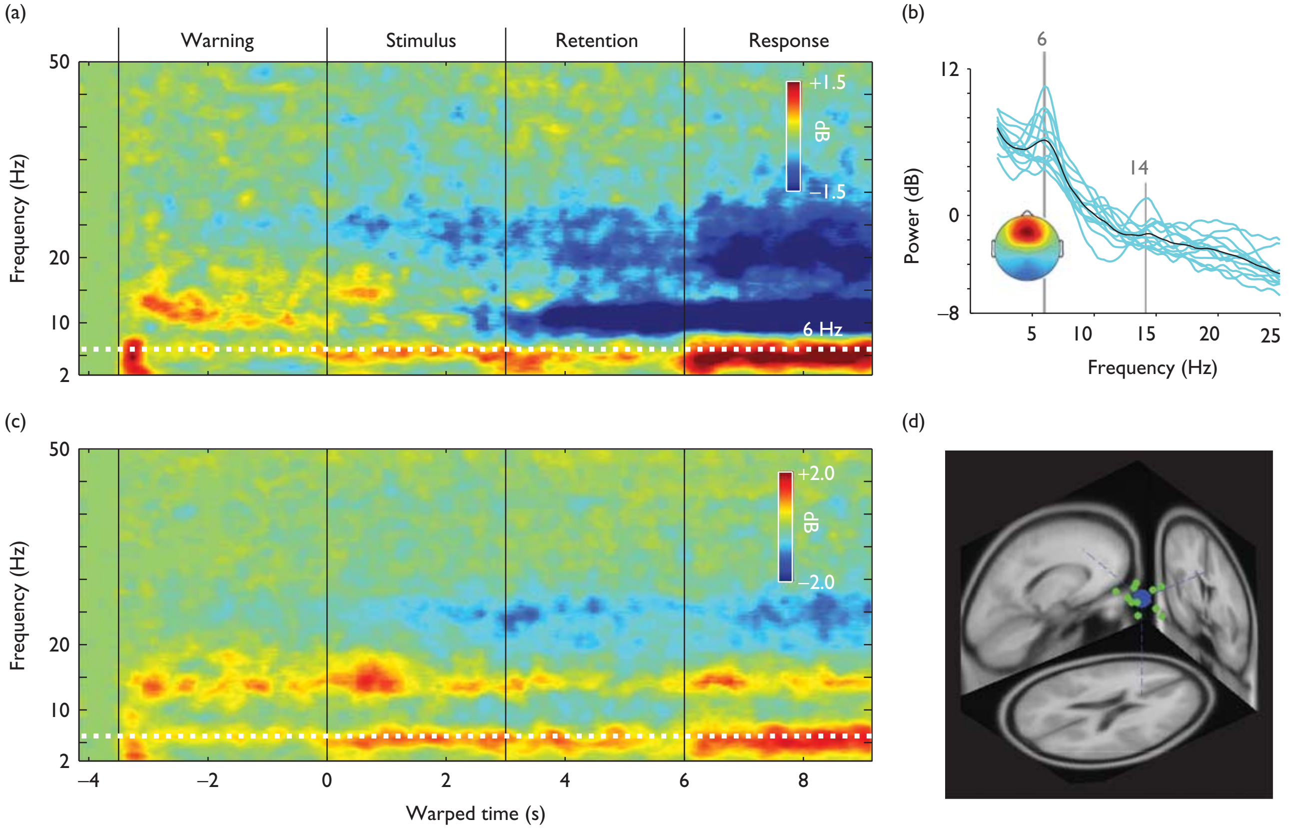
The individual (thin cyan lines) and mean (thick black line) spectra of ICs in the Fmθ cluster (Fig. 2b) show clear θ peaks, with additional peaks around 14 Hz (low-β). The mean scalp map for this cluster indicates strong projection at frontal midline electrodes. Estimated equivalent current dipoles for individual IC processes within the cluster fall mostly in the medial frontal cortex (centroid located at Talairach coordinates: $X = 0$, $Y = 21$, $Z = 33$; Fig. 2d). The grand-average ERSP (Fig. 2c) shows significant increases in both θ (cf. Fig. 2a) and low-β power relative to the pretrial period. All characteristics are consistent with previous frequency/time–frequency dynamics of Fmθ-related ICs and source estimations of Fmθ [13,14,19].

To analyze the potential differences across SNRs and trial intervals, the mean relative θ power for ICs during the warning (− 2.5 to 0 s), stimulus (0.5–3 s), retention (3.5–6 s), and response (6.5–9 s) intervals was entered into a 4 (interval) × 5 (SNR) repeated measures ANOVA. Figure 3 shows the relative θ power across SNRs for each interval. A significant main effect of interval was found [$F(3,36) = 4.98$, $P = 0.005$, $\eta_p^2 = 0.29$], likely related to high relative θ power during the response interval. A significant main effect of SNR [$F(4,48) = 3.45$, $P = 0.017$, $\eta_p^2 = 0.22$] and a marginally significant interval × SNR interaction were also found [$F(12,144) = 1.92$, $P = 0.089$, $\eta_p^2 = 0.14$], suggesting that the effects of SNR on Fmθ differed across intervals.

Separate repeated measures ANOVAs run for each interval found significant main effects of SNR for the warning [$F(4,48) = 4.69$, $P = 0.003$, $\eta_p^2 = 0.28$] and stimulus [$F(4,48) = 4.92$, $P = 0.002$, $\eta_p^2 = 0.29$] intervals. The mean values for the warning [$F(1,12) = 16.24$, $P = 0.002$, $\eta_p^2 = 0.58$] and stimulus intervals [$F(1,12) = 10.60$, $P = 0.007$, $\eta_p^2 = 0.47$] were significantly fit to linear trends, suggesting that Fmθ increased as SNR decreased. No effect of SNR was found for the other intervals ($F < 2$). The lack of an effect in the response interval suggests that θ effects are not attributable to movement-related EEG dynamics.

The lower SNRs that show the greatest θ also show the lowest accuracy. To address a possible θ power correlate with accuracy rather than effort, the 12 and − 6 dB conditions were compared after removing epochs for which

Fig. 2



(a) Grand-average time-warped ERSP at Fz. (b) The mean (black line) and individual (cyan lines) spectra for IC processes within the Fmθ cluster. (c) The mean grand-average time-warped ERSP for IC processes in the Fmθ cluster (one per listener). (d) Locations of estimated equivalent current dipoles for each IC process (small green spheres) plotted on a standard MRI image (Montreal Neurological Institute). The center of the cluster (large blue sphere) is also plotted. In (a) and (c), individual ERSPs were masked for significance ($P < 0.01$; bootstrapping procedure [19]). The mean ERSPs shown were created by averaging the significance-masked ERSPs. Statistics reported in the text, however, were run on relative power measures extracted from unmasked ERSPs. ERSP, event-related spectral perturbation; Fmθ, frontal midline θ; IC, independent component.

listeners failed to report all words correctly. The number of 12 and –6 dB epochs were equated for each individual by randomly dropping epochs from the condition with a greater amount of epochs. ERSPs were recomputed. A 2 (interval) × 2 (SNR) repeated measures ANOVA found a main effect of SNR [$F(1,12) = 12.11$, $P = 0.005$, $\eta_p^2 = 0.50$], showing that even on trials in which recognition accuracy was perfect, Fmθ was greater under the –6 dB, relative to the 12 dB, condition. The main effect of interval [$F(1,12) = 3.40$, $P = 0.09$, $\eta_p^2 = 0.22$] was marginally significant, likely related to greater power in the stimulus interval. The interaction did not reach significance ($F < 2$).

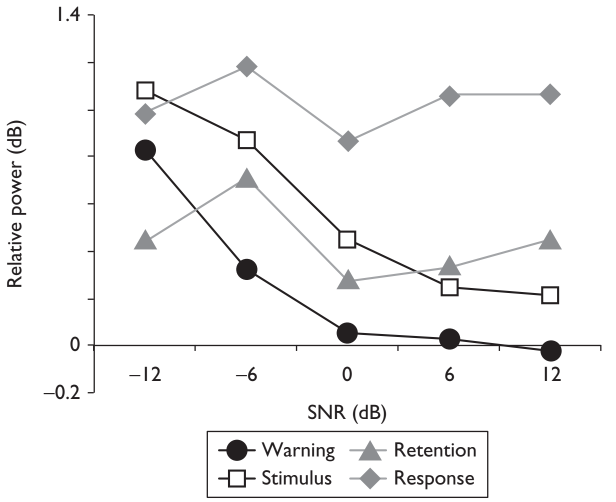
A Pearson correlation between the mean relative θ power at each SNR (averaged across warning and stimulus intervals) and the mean self-reports of listening effort was calculated to assess the relationship between electrophysiological measures and behavior. The relationship was positive and significant [$r(3) = 0.97$, $P = 0.006$], suggesting that Fmθ parallels subjective reports [2].

Discussion

A medial frontal cluster of IC processes extracted from EEG with ICA showed greatest θ power when SNRs were low in a sentence-recognition task. This was true even when recognition accuracy was equated across SNRs at 12 and −6 dB. Fmθ power was also positively correlated with the listeners' reports of effort. The localization of ICs within the Fmθ cluster to the medial frontal cortex is consistent with previous source estimations of Fmθ and corroborates blood-flow-based neuroimaging studies showing increased activation of medial frontal regions under attentive listening conditions.

Most EEG/ERP studies of listening effort have associated greater amplitude transient phase-locked responses (e.g. P3 [6]; intertrial phase coherence < 200 ms post stimulus [5]) with greater effort. Generally, it has been concluded that enhanced amplitudes reflect greater use of attentional resources. However, some difficult listening conditions (e.g. low SNRs) reduce the amplitude of

Fig. 3



The time windows analyzed omit at least the first 0.5 s of the intervals to limit the contributions of ERPs to relative power. The relative power measures analyzed were extracted from ERSPs before creating the significance-masked ERSPs ($P < 0.01$, bootstrapping) shown in Fig. 2. Relative θ power across SNRs during each interval. Error bars are omitted from the plot for clarity. Repeated measures SEs ($\sqrt{MSe/n}$) [20] were as follows: warning, 0.16; stimulus, 0.17; retention, 0.17; response, 0.20. SNR, signal-to-noise ratio. (warning:–2.5 to 0 s, stimulus: 0.5–3 s, retention: 3.5–6 s, and response: 6.5–9 s).

these transient responses [22], making it difficult to tease apart stimulus-related and effort-related effects. We ran a second experiment in which listeners were instructed to ignore sounds and watch a captioned movie (for data see Supplemental digital content 1, *http://links.lww.com/WNR/A313*). Relative Fmθ was low and flat across SNRs, suggesting that the effects reported here are more sensitive to task engagement than to stimulus features. Further, listeners are presumably exerting effort during time periods exceeding 500 ms after sound onset and when to-be-heard sounds are absent (i.e. windows of time not well captured by ERPs). We see greater Fmθ at low SNRs even before speech is presented. Similarly, others have reported sustained spectral perturbations in the α-band, which appear to be related to anticipatory listening [23] and listening difficulty [21]. Future work will need to explore the full gamut of features contained in EEG to understand how listeners accomplish successful listening under difficult conditions.

Given the connection between Fmθ and working memory [14], it could be that greater Fmθ reflects greater use of working memory processes for comparing degraded input to long-term memory representations for words [4]. It could also be that θ serves to establish long-range communication in cortical networks through oscillatory synchrony [24], plays a role in attention-related gain modulation of obligatory sensory responses [25], or in encoding of sounds into memory [24]. Future studies will need to manipulate attention and memory demands, perhaps with nonspeech stimuli, to assess these hypotheses. Nevertheless, frontal cortical networks appear to be increasingly utilized as listening becomes increasingly effortful [16].

Conclusion

Fmθ can be used to index listening effort. Analysis of Fmθ dynamics and other time–frequency EEG features (for review see Weisz *et al.* [21]) may complement studies on listening effort that focus on transient EEG/ERP features and peripheral psychophysiological measures, giving a better picture of how human brain dynamics support successful listening. Fmθ power could also potentially be used to quantify the benefits of treatments for the hearing-impaired with regard to listening effort, especially when benefits indexed by behavior are minimal or absent.

Acknowledgements

This research was performed while M.G.W. held a National Research Council Associateship Award at the US Air Force Research Laboratory. M.N.G.-R. was supported in part by an appointment to the Student Research Participant Program at the US Air Force Research Laboratory administered by the Oak Ridge Institute for Science and Education.

Conflicts of interest

There are no conflicts of interest.

References

1 Pichora-Fuller MK. Perceptual effort and apparent cognitive decline: Implications for audiological rehabilitation. *Semin Hear* 2006; **27**:284–293.
2 Mackersie CL, Cones H. Subjective and psychophysiological indexes of listening effort in a competing-talker task. *J Am Acad Audiol* 2011; **22**:113–122.
3 Gosselin PA, Gagné JP. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int J Audiol* 2011; **50**:786–792.
4 Rönnberg J, Rudner M, Foo C, Lunner T. Cognition counts: a working memory system for ease of language understanding (ELU). *Int J Audiol* 2008; **47 (Suppl 2)**:S99–S105.
5 Bernarding C, Strauss DJ, Hannemann R, Seidler H, Corona-Strauss FI. Neural correlates of listening effort related factors: influence of age and hearing impairment. *Brain Res Bull* 2013; **91**:21–30.
6 Bertoli S, Bodmer D. Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. *Clin Neurophysiol* 2014; **125**:1030–1041.
7 Zekveld AA, Kramer SE, Festen JM. Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear Hear* 2010; **31**:480–490.
8 Garcia A, Uribe CE, Tavares MC, Tomaz C. EEG and autonomic responses during performance of matching and non-matching to sample working memory tasks with emotional content. *Front Behav Neurosci* 2011; **5**:82.
9 Tursky B, Shapiro D, Crider A, Kahneman D. Pupillary, heart rate, and kin resistance changes during a mental task. *J Exp Psychol* 1969; **79**:164–167.
10 Posner MI, Raichle ME. *Images of mind*. New York: Scientific American Library; 1994.
11 Dawson ME, Schell AM, Filion DL. The electrodermal system. In: Cacioppo JT, Tassinary LG, Berntson GG, editors. *Handbook of psychophysiology*. New York: Cambridge University Press; 2000. pp. 159–181.

12 Debener S, Makeig S, Delorme A, Engel AK. What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Brain Res Cogn Brain Res* 2005; **22**:309–321.

13 Ishii R, Shinosaki K, Ukai S, Inouye T, Ishihara T, Yoshimine T, *et al*. Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport* 1999; **10**:675–679.

14 Onton J, Delorme A, Makeig S. Frontal midline EEG dynamics during working memory. *Neuroimage* 2005; **27**:341–356.

15 Paus T, Zatorre RJ, Hofle N, Caramanos Z, Gotman J, Petrides M, Evans AC. Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. *J Cogn Neurosci* 1997; **9**:392–408.

16 Erb J, Henry MJ, Eisner F, Obleser J. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J Neurosci* 2013; **33**:10688–10697.

17 Kidd G Jr, Best V, Mason CR. Listening to every other word: examining the strength of linkage variables in forming streams of speech. *J Acoust Soc Am* 2008; **124**:3793–3802.

18 Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 2004; **134**:9–21.

19 Makeig S, Onton J. ERP features and EEG dynamics: an ICA perspective. In: Luck S, Kappenman E, editors. *Oxford handbook of event-related potential components*. New York: Oxford University Press; 2009. pp. 51–86.

20 Loftus GR, Masson ME. Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1994; **1**:476–490.

21 Weisz N, Hartmann T, Müller N, Lorenz I, Obleser J. Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol* 2011; **2**:73.

22 Bennett KO, Billings CJ, Molis MR, Leek MR. Neural encoding and perception of speech signals in informational masking. *Ear Hear* 2012; **33**:231–238.

23 Bastiaansen MC, Brunia CH. Anticipatory attention: an event-related desynchronization approach. *Int J Psychophysiol* 2001; **43**:91–107.

24 Buzsáki G. *Rhythms of the brain*. New York: Oxford University Press; 2006.

25 Bruneau N, Roux S, Guérin P, Garreau B, Lelord G. Auditory stimulus intensity responses and frontal midline theta rhythm. *Electroencephalogr Clin Neurophysiol* 1993; **86**:213–216.