Indices of Effortful Listening Can Be Mined from Existing Electroencephalographic Data

Matthew G. Wisniewski

Objectives: Studies suggest that theta (~4 to 7 Hz), alpha (~8 to 12 Hz), and stimulus-evoked dynamics of the electroencephalogram index effortful listening. Numerous auditory event-related potential datasets exist, without thorough examination of these features. The feasibility of mining those datasets for such features is assessed here.

Design: In a standard auditory-oddball paradigm, 12 listeners heard deviant high-frequency tones (10%) interspersed among low-frequency tones (90%) “near” or “far” separated in frequency.

Results: During active listening (deviance detection; experiment 1), sustained frontal midline theta power, and gamma-band inter-trial phase coherence, were greater for the near condition. No significant “near”/“far” differences were observable during passive exposure to the same sounds (experiment 2).

Conclusions: Increased theta power likely reflects increased utilization of cognitive-control processes (e.g., working memory) that rely on frontal cortical networks. Inter-trial phase coherence differences may reflect differences in attention-modulated stimulus encoding. Reanalysis of existing datasets can usefully inform future work on listening effort.

Key words: Attention, Auditory-evoked potential, Cognitive hearing, Data mining, Event-related spectral perturbation, Working memory.

(Ear & Hearing 2017;38;e69–e73)

INTRODUCTION

Difficult listening scenarios often feel as though they require extensive mental effort (McGarrigle et al. 2014). An increasing amount of studies are starting to explore psychophysiological correlates of this *listening effort* in the electroencephalogram (EEG). Recently, frontal midline theta (4 to 7 Hz) power was found to increase as signal to noise ration (SNR) decreased in a sentence-recognition task, paralleling self-reports of effort (Wisniewski et al. 2015). Effortful listening has also been associated with sustained increases in alpha (8 to 12 Hz) power (for review, see Strauß et al. 2014), and inter-trial phase coherence (ITPC) of the EEG shortly following sound presentation (Mulert et al. 2007; Bernarding et al. 2013).

Potentially, these features could be mined from existing event-related potential (ERP) datasets to address a growing number of effort-related questions in cognitive-hearing science (e.g., How does aging affect effort? How does effort vary in nonspeech tasks?). This prospect’s feasibility is assessed here. An EEG auditory-oddball task was used wherein standard and deviant tones were “near” or “far” separated in frequency. This paradigm was chosen because (1) it is popularly used in ERP work (for review, see Näätänen et al. 2007); (2) “near” and “far” conditions differ in task difficulty which is often used to study mental effort (Ishii et al. 1999; McGarrigle et al. 2014); and (3) task-difficulty manipulations, intended or not, are frequently present in auditory ERP studies.

In experiment 1, participants were asked to detect deviant tones (active listening). In experiment 2, participants were presented the same sounds without a task objective (passive exposure). Based on prior research (Mulert et al. 2007; Bernarding et al. 2013; Strauß et al. 2014; Wisniewski et al. 2015), it was hypothesized that frontal midline theta, posterior alpha, and ITPCs would be larger in amplitude for the “near” condition, but only during active listening. If these features are readily observable, map onto task difficulty, and are dependent on active listening, this would suggest that mining former ERP datasets could usefully supplement the study of listening effort.

MATERIALS AND METHODS

Participants

Thirteen adults (ages 19 to 34; 5 male; self-reported normal-hearing) participated for pay or volunteered. All signed a U.S. Air Force Institutional Review Board approved informed consent document. One participant was dropped for excessively noisy EEG data.

Stimuli & Procedures

Sounds were presented diotically at ~81 dB SPL through earphones (ER-2; Etymotic Research, Elk Grove Village, IL). Event-code and stimulus presentation timing were controlled via a TDT system 3 real-time processor (RP2.1; Tucker-Davis Technologies, Alachua, FL). Procedures were executed in MATLAB R2013a (Mathworks, Natick, MA). Participants sat in a sound-attenuated booth.

Experiment 1

Instructions were to button press upon hearing “high frequency,” 60 msec sinusoidal tones at either 515 Hz (near) or 1200 Hz (far). High-frequency tones occurred at a rate of 10%, interspersed between “low-frequency” tones at 500 Hz, which occurred the other 90% of the time. All tones had amplitude rise and fall times of 10 msec. Intertone intervals were randomized between 1000 and 2000 msec. Near and far conditions were blocked, with 3 blocks per condition, and 70 tone presentations per block. Deviant and standard tones were pseudorandomized so that there were at least two standards in between each deviant tone. Block order was pseudorandomized so that no more than two blocks of the same condition occurred back-to-back.

Experiment 2

Stimuli, block structure, and conditions were the same as experiment 1. Instead of performing deviance detection, participants read material of their choice during sound presentation. They were asked to ignore sounds. Half of participants completed experiment 2 after experiment 1. The order was reversed for the other half.

0196/0202/2017/381-0e69/0 • Ear & Hearing • Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

EEG Acquisition & Analysis

Data were collected at a 2048 Hz sample rate, 24 bit A/D resolution, and referenced to the Common Mode Sense/Driven Right Leg (CMS/DRL) reference of the BioSemi Active II system (BioSemi, Amsterdam, The Netherlands). Thirty-two scalp electrodes were positioned at 10-20 system locations. Four additional electrodes were placed adjacent and underneath the eyes to monitor ocular artifacts.

Analyses were performed with EEGLAB (Delorme & Makeig 2004) and custom MATLAB functions/scripts. Data were downsampled to 256 Hz, band-pass filtered (0.5 to 100 Hz), and re-referenced using an average reference. Continuous data and channels contaminated by excessive high-frequency activity were rejected. Remaining data were submitted to independent components analysis. Independent components identified as artifacts (e.g., eye movements) by time courses, spectra, and scalp projections were removed. Missing channels were then interpolated (spherical interpolation). Epochs were extracted from −3 to 3 sec surrounding each tone onset in which the time-locked tone, previous tone, and following tone were not deviants. The elimination of epochs containing deviants served to minimize effects of deviant-related ERP responses (e.g., P300), and potential effects specific to deviant tone-frequencies, which differed between conditions. Single-epoch time-frequency representations were then computed for 95 logarithmically spaced frequencies between 3 and 50 Hz, centered at 800 time points using the *newtimef()* function in EEGLAB (5 cycle complex Morlet wavelets).

Theta and Alpha Power Dynamics

Single-epoch time-frequency representations were linearly time warped to produce equal numbers of data points between the time-locked tone onset and the next tone onset for each epoch such that the next tone for each epoch was aligned to 1500 msec (Delorme & Makeig 2004). Median event-related spectral power matrices were then obtained by taking the median power (across epochs) at each time-frequency point. Separate theta (4 to 7 Hz) and alpha (8 to 12 Hz) power traces were extracted by taking the mean across frequencies (within band) at each time point. These traces were then divided by the mean within-band power across all time points within the "near" and "far" traces (i.e., a common divisive baseline), and converted to decibels.

ITPCs

ITPC indices of listening effort have been examined in the theta (Bernarding et al. 2013) and gamma-bands (Mulert et al. 2007). Both are analyzed here. Taking the single-epoch (r) phase angles (k) at each time-frequency point (tf), and the total number of epochs (n), the formula:

$$\mathrm{ITPC}_{\mathrm{tf}} = \left| 1/n \sum_{r=1}^{n} e^{ik_{tfr}} \right|$$

produces an ITPC value between 0 and 1 for each time-frequency point. Since differences in epoch numbers can bias ITPCs (Cohen 2014), and because some individuals by chance showed a greater than 30 epoch count difference between conditions, a normalized ITPC_Z (Rayleigh's Z) was calculated using the formula:

$$\mathrm{ITPC}_z = n \times \mathrm{ITPC}^2$$

ITPC_Z traces were then created by averaging ITPC_Z across the 4 to 7 Hz (theta) and 30 to 50 Hz (gamma) bands for each time point.

Statistics

Analyses were conducted on data averaged across electrodes where effects were expected to manifest a priori. Theta power dynamics were analyzed at electrode groups AF3, AF4, F3, F4, and Fz, alpha power dynamics at P7, P8, P3, P4, and Pz, and ITPCs at Fz, FC1, FC2, and Cz.

A nonparametric permutation method was used to compare theta, alpha, and ITPC traces obtained in the "near" and "far" conditions. First a null-hypothesis distribution of t values was generated by randomly assigning condition labels for each individual's traces for 1000 iterations. A t statistic was calculated for each iteration, creating a distribution of t values expected under the null hypothesis at each time point. The p values for real data were considered to be the proportion of null-hypothesis distribution iterations having more extreme t values than the real data (Cohen 2014). At this step, significant time points were identified for each electrode group at an alpha level of 0.05.

To control for multiple comparisons, an additional step was taken to identify clusters of contiguously significant time points that were larger than expected by chance. For another 1000 iterations (shuffling condition labels), the largest cluster of contiguously significant time points was identified. A second null-hypothesis distribution was then obtained by taking the sum of the squared t values within the identified largest clusters. Any cluster in the real data in which the sum of squared t values was larger than 95% of the null-hypothesis distribution was considered significant (for review, see Cohen 2014).

RESULTS

Experiment 1 Behavior

Button presses given within 2 sec following a deviant tone's onset were considered "hits," while presses given within 2 sec after a standard sound (excluding those already marked as hits) were considered "false alarms." The *A'* signal detection measure was then calculated. Participants showed an *A'* of 0.83 (SD = 0.18) in the "near" condition and an *A'* of 0.99 (SD = 0.03) in the "far" condition. A Wilcoxon Signed Ranks test found this difference to be significant, $Z = 2.80$, interquartile range = 0.20, $p = 0.005$.

Theta and Alpha Dynamics

Theta traces from experiment 1 are shown in Figure 1A for the "near" (solid blue lines) and "far" (dashed red lines) conditions. There were transient increases in theta power shortly following sound onsets in both conditions, as is typical (e.g., Onton et al. 2005; Wisniewski et al. 2015). More importantly, there was a sustained significant difference in theta from ~400 to ~1400 msec post tone onset (highlighted in yellow). The scalp map of this difference shows that the effect was strongest at the expected frontal midline electrodes (cf. Onton et al. 2005;

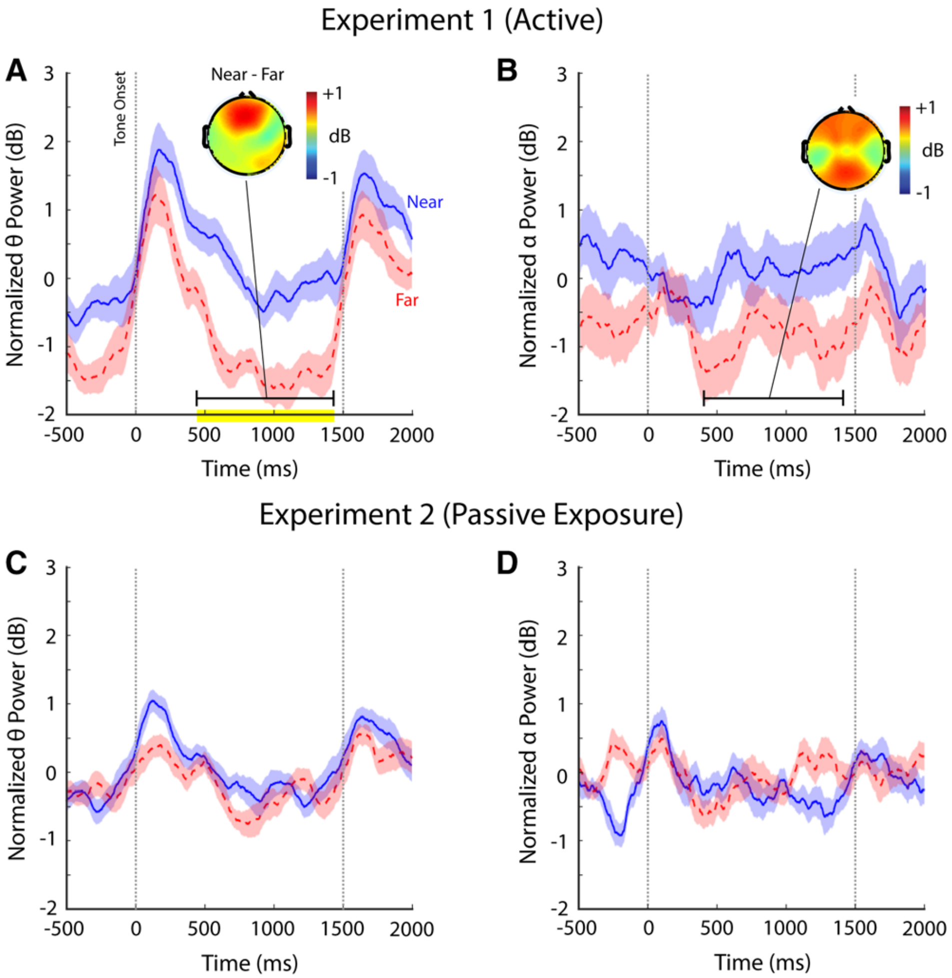

Fig. 1. Theta (A, C) and alpha (B, D) power dynamics in experiment 1 (A, B) and experiment 2 (C, D). Blue solid lines depict power traces in the "near" conditions. Red dashed lines depict power traces in the "far" conditions. Shading above and below lines represent within-subject standard errors (Cousineau 2005). Regions of significant difference between conditions are highlighted in yellow on the abscissa (cluster corrected, $p < 0.05$). For experiment 1, scalp maps of the condition differences over marked time windows are shown ("near" minus "far").

Wisniewski et al. 2015).* In contrast, there did not seem to be any differences in frontal midline theta during passive exposure (experiment 2; Fig. 1C).

There was a similar trend for greater alpha power in the "near" condition relative to the "far" condition during the same time window in experiment 1 (Fig. 1B). This trend did not reach significance. No such trend was observable in experiment 2 (Fig. 1D).

ITPCs

Figure 2A, B shows theta- and gamma-band ITPCs for experiment 1. Although there is a trend for greater ITPC of the theta band when listening for "near" deviants (Fig. 2A), this did not reach significance following cluster correction for multiple comparisons.† A similar trend was observable for gamma. This trend did reach significance between ~35 and ~140 msec post tone onset. Scalp maps of "near"/"far" ITPC differences for both theta and gamma show the largest effects around fronto-central electrodes, with a slight tendency for gamma to show a stronger effect at right-side electrodes.

Theta and gamma ITPCs in "near" and "far" conditions were basically indistinguishable in experiment 2 (Fig. 2C, D). No significant differences were detected. Also notable, is that ITPCs were generally lower in passive exposure (experiment 2) compared with active listening (experiment 1).

DISCUSSION

This study investigated whether or not recently reported indices of effortful listening are observable in a commonly used auditory-oddball paradigm. In experiment 1, listening for relatively small frequency differences ("near" versus "far"

---

*The possibility that the "near"/"far" difference in theta was attributable to a suppression of theta in the "far" condition rather than enhancement of theta in the "near" condition was also assessed. For the 10 participants who had data with large enough clean interblock intervals, powers within these inactive periods were used as an alternative normalization for theta traces (i.e., substituting for mean power across all epoch time points). Both the "near" and "far" conditions showed positive relative powers consistent throughout epochs, suggesting that effects are related to differences in theta enhancement, rather than suppression.

†A less stringent analysis wherein the mean theta ITPC_Z was averaged from between 100 and 300 msec post stimulus onset for the same electrode group only reached marginal significance, $t(11) = 1.94$, $p = 0.078$.

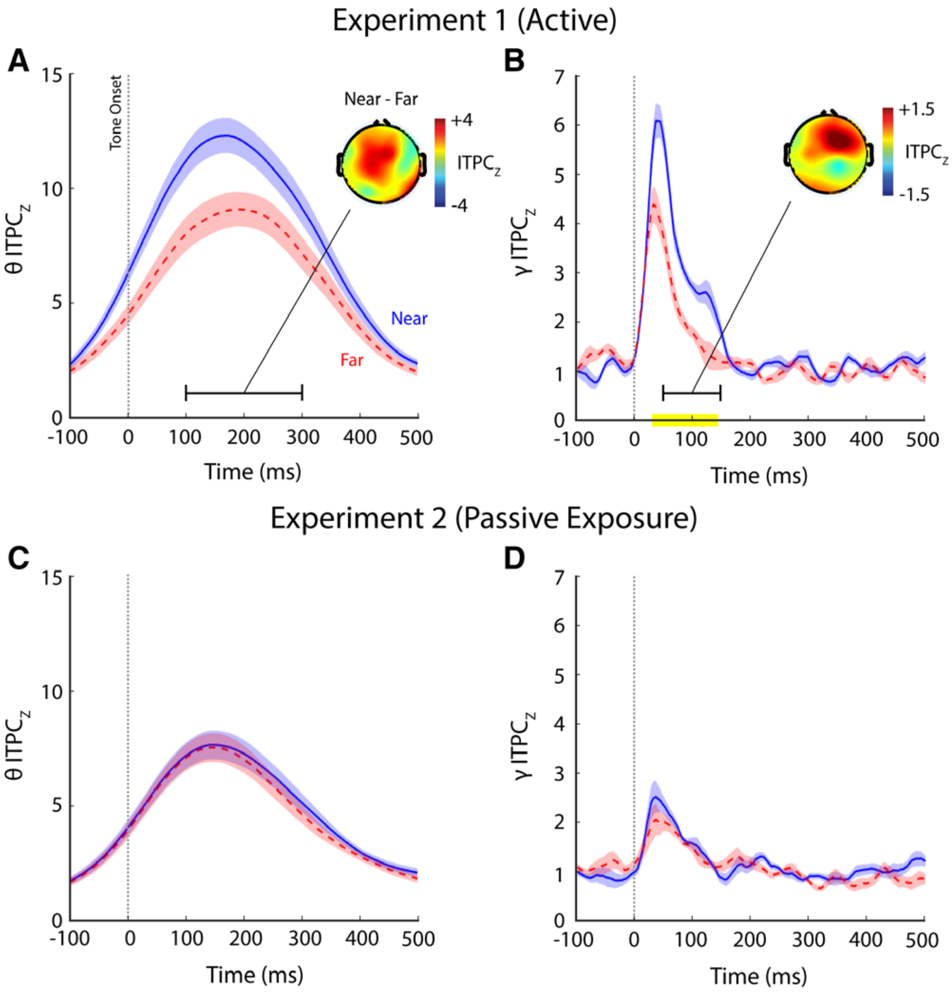

Fig. 2. Theta (A, C) and gamma (B, D) ITPCs in experiment 1 (A, B) and experiment 2 (C, D). Blue solid lines depict ITPC traces in the "near" conditions. Red dashed lines depict ITPC traces in the "far" conditions. Shading above and below lines represent within-subject standard errors (Cousineau 2005). Regions of significant difference between conditions are highlighted in yellow on the abscissa (cluster corrected, $p < 0.05$). For experiment 1, scalp maps of the condition differences over marked time windows are shown ("near" minus "far").

standard/deviant frequency separation) was associated with greater frontal midline theta power (cf. Wisniewski et al. 2015). It was also found that gamma-band ITPC following standard 500 Hz tones was significantly enhanced in the "near" condition compared with the "far" condition (cf. Mulert et al. 2007). Experiment 1 "near"/"far" differences in alpha power at parietal electrode sites (Strauß et al. 2014), and ITPCs of the theta band (Bernarding et al. 2013), showed trends consistent with the literature, but failed to reach significance. During passive exposure to the same sounds (experiment 2), there were no significant or apparent "near"/"far" differences. This suggests that the effects observed in experiment 1 are dependent upon a listener's engagement.

The lack of significance for alpha power and theta-band ITPC measures may reflect low power, or perhaps the use of a task (nonspeech frequency discrimination) that fails to produce effects comparable with previously employed speech perception paradigms. However that differences were observable in frontal midline theta and ITPC of gamma, suggests that listening-effort features are minable from previously collected ERP datasets. There exist auditory ERP investigations on various populations (e.g., hearing impaired and older age groups), various types of masking (e.g., informational and energetic), treatment regimens (e.g., auditory training), and stimulus parameters (e.g., varying SNRs). Most have ignored the time-frequency features researchers have associated with effort (Mulert et al. 2007; Bernarding et al. 2013; Strauß et al. 2014; Wisniewski et al. 2015). Dataset re-examination may be useful for assessing a growing number of questions raised in regards to listening effort. For instance, ERP data that exists on individuals that vary in age could be used to initially investigate relationships between cognitive decline and effort, before a more extensive, expensive, and time-consuming investigation is initiated. A mining approach is post hoc and will therefore require cautious interpretation. Furthermore, it will be necessary to examine the data in a manner that addresses possible impacts of nonlistening-related factors that may differ between conditions (e.g., varying SNRs; Bennett et al. 2012). Nevertheless, careful mining can play a useful role by supplementing, motivating, and deriving testable hypotheses for planned studies.

The functional roles that frontal midline theta, posterior alpha, and ITPC reflect during listening are debated. Nonauditory EEG studies have repeatedly associated frontal midline theta with cognitive-control processes (Ishii et al. 1999; Onton et al.

2005), while blood flow-based neuroimaging studies have implicated its suspected sources (anterior cingulate cortex; Ishii et al. 1999; Onton et al. 2005) in the same processes during listening (e.g., Erb & Obleser 2013). One possibility for the observed theta effect is that in the "near" condition, listeners' strategy was to hold the previously presented tone in working memory for later comparison. Such a strategy would be similar to that used in nonauditory tasks that have repeatedly shown increases in frontal midline theta power as the number of to-be-retained items increases (Onton et al. 2005). In the "far" condition, listeners can easily detect changes in frequency; hence, such a strategy may not be needed. It could also be that theta reflects increased error monitoring (Botvinick et al. 2004) or inter-regional communication between frontal networks and long-term memory systems in the hippocampus (Buzsáki 2006). Increased alpha in difficult listening conditions may also reflect inter-regional communication, or possibly the inhibition of task-irrelevant brain areas (Strauß et al. 2014). In regards to enhancement of ITPC, several have claimed that increased ITPCs reflect enhanced stimulus encoding, affected by attention (e.g., Mulert et al. 2007; Bernarding et al. 2013). The current work was not designed to disentangle hypotheses regarding the functional roles of these features. Future experiments that examine functional roles, perhaps aided by data-mining work, should move the study of listening effort toward a deeper understanding of the cognitive processes that underlie it. Such work may also reveal other EEG features (e.g., spectral dynamics of other EEG bands, functional connectivity, etc.) that characterize cognitive processes during listening.

CONCLUSION

The majority of auditory ERP work has ignored, or underanalyzed, the time-frequency dynamics that seem to relate to listening effort. Although such ERP work may not have been explicitly designed to reveal these dynamics, these dynamics are likely observable and quantifiable. Former datasets can be used as a resource to grow our understanding of the processes that underlie listening effort, and the effects of varying conditions on it.

U.S. Air Force Research Laboratory, Wright Patterson Air Force Base, Ohio, USA.

ACKNOWLEDGMENTS

Thank you to Max N. Goder-Reiser and Stephanie Kenzig for assistance in data collection and members of the Battlespace Acoustics branch of the 711th Human Performance Wing at the U.S. Air Force Research Laboratory for helpful discussions on this data and topic. Jonas Obleser is also thanked for reading and commenting on a previous version of this manuscript.

This research was performed while M.G.W. held a National Research Council Associateship Award and an Oak Ridge Institute for Science and Education Postdoctoral Fellowship at the U.S. Air Force Research Laboratory.

The author has no conflicts of interest to disclose.

Address for correspondence: U.S. Air Force Research Laboratory, Wright Patterson Air Force Base, OH, USA. E-mail: matt.g.wisniewski@gmail.com

Received December 1, 2014; accepted June 5, 2016.

REFERENCES

Bennett, K. O., Billings, C. J., Molis, M. R., et al. (2012). Neural encoding and perception of speech signals in informational masking. *Ear Hear*, *33*, 231–238.

Bernarding, C., Strauss, D. J., Hannemann, R., et al. (2013). Neural correlates of listening effort related factors: influence of age and hearing impairment. *Brain Res Bull*, *91*, 21–30.

Botvinick, M. M., Cohen, J. D., Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn Sci*, *8*, 539–546.

Buzsáki, G. (2006). *Rhythms of the Brain*. New York, NY: Oxford University Press.

Cohen, M.X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, MA: MIT Press.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutor Quant Methods Psychol*, *1*, 42–45.

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

Erb, J., & Obleser, J. (2013). Upregulation of cognitive control networks in older adults' speech comprehension. *Front Syst Neurosci*, *7*, 116.

Ishii, R., Shinosaki, K., Ukai, S., et al. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport*, *10*, 675–679.

McGarrigle, R., Munro, K. J., Dawes, P., et al. (2014). Listening effort and fatigue: what exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. *Int J Audiol*, *53*, 433–440.

Mulert, C., Leicht, G., Pogarell, O., Mergl, R., Karch, S., Juckel, G., et al. (2007). Auditory cortex and anterior cingulate cortex sources of the early evoked gamma-band response: relationship to task difficulty and mental effort. *Neuropsychologia*, *45*, 2294–2306.

Näätänen, R., Paavilainen, P., Rinne, T., et al. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol*, *118*, 2544–2590.

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

Strauß, A., Wöstmann, M., Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci*, *8*, 350.

Wisniewski, M. G., Thompson, E. R., Iyer, N., et al. (2015). Frontal midline θ power as an index of listening effort. *Neuroreport*, *26*, 94–99.