Hearing Research 358 (2018) 37–41

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Short communication

Sustained frontal midline theta enhancements during effortful listening track working memory demands

Matthew G. Wisniewski*, Nandini Iyer, Eric R. Thompson, Brian D. Simpson

711th Human Performance Wing, U.S. Air Force Research Laboratory, United States

ARTICLE INFO

Article history:
Received 9 August 2017
Received in revised form
16 October 2017
Accepted 24 November 2017
Available online 27 November 2017

Keywords:
Event-related spectral perturbation (ERSP)
Listening effort
Perceptual anchor
Event-related synchronization

ABSTRACT

Recent studies demonstrate that frontal midline theta power (4–8 Hz) enhancements in the electroencephalogram (EEG) relate to effortful listening. It has been proposed that these enhancements reflect working memory demands. Here, the need to retain auditory information in working memory was manipulated in a 2-interval 2-alternative forced-choice delayed pitch discrimination task ("Which interval contained the higher pitch?"). On each trial, two square wave stimuli differing in pitch at an individual's ~70.7% correct threshold were separated by a 3-second ISI. In a 'Roving' condition, the lowest pitch stimulus was randomly selected on each trial (uniform distribution from 840 to 1160 Hz). In a 'Fixed' condition, the lowest pitch was always 979 Hz. Critically, the 'Fixed' condition allowed one to know the correct response immediately following the first stimulus (e.g., if the first stimulus is 979 Hz, the second must be higher). In contrast, the 'Roving' condition required retention of the first tone for comparison to the second. Frontal midline theta enhancements during the ISI were only observed for the 'Roving' condition. Alpha (8–13 Hz) enhancements were apparent during the ISI, but did not differ significantly between conditions. Since conditions were matched for accuracy at threshold, results suggest that frontal midline theta enhancements will not always accompany difficult listening. Mixed results in the literature regarding frontal midline theta enhancements may be related to differences between tasks in regards to working memory demands. Alpha enhancements may reflect task general effortful listening processes.

Excessive listening effort can discourage socializing (Kiessling et al., 2003), impact performance in concurrently performed tasks (Rabbitt, 1991), and lead to fatigue (Hornsby, 2013). Psychophysiological correlates of listening demands exist in skin conductance (Mackersie and Cones, 2011), pupil dilation (Koelewijn et al., 2015), and various M/EEG features (Bernarding et al., 2013; Weisz and Obleser, 2014; Wisniewski, 2017; Wisniewski et al., 2015). These objective measures have allowed scientists to examine how different conditions affect effort, and may eventually serve in the development of procedures for reducing it (Bertoli and Bodmer, 2014; McGarrigle et al., 2014).

Of these approaches, EEG provides an especially rich set of features for indexing the complex set of cognitive processes that underlie listening (for review, see Rönnberg et al., 2008; Strauss and Francis, 2017; Wisniewski, 2017). Some researchers have reported effort-related modulations to the power of alpha (~8–13 Hz) oscillations. Alpha has been related to attentional processes based on comparisons of active versus passive listening (e.g., Dimitrijevic et al., 2017) and different selective attention conditions (e.g., Wöstmann et al., 2016). We have reported that enhancements to frontal midline theta-band (~4–8 Hz) oscillations are affected by task difficulty and parallel self-reports of increased effort (Wisniewski, 2017; Wisniewski et al., 2015, 2017). Given non-auditory work repeatedly relating the frontal midline theta rhythm to working memory demands (e.g., Onton et al., 2005), we proposed that enhancements could reflect a working memory component of effortful listening (cf. Pesonen et al., 2006).

Though this hypothesis is viable, several reasons leave it weak. First, working memory demands have not been explicitly manipulated in our previous studies. Effects of such manipulations would provide stronger support than similarities to results from non-auditory work. Second, several studies have failed to find significant frontal-midline theta enhancements. In some of these cases, tasks did not entail a strong working memory component (e.g.,

* Corresponding author. 7108 Creek Water Dr., Centerville, OH 45459, United States.
E-mail address: matt.g.wisniewski@gmail.com (M.G. Wisniewski).

https://doi.org/10.1016/j.heares.2017.11.009

Marsella et al., 2017). However, other studies have failed to find effects while using auditory working memory paradigms. For example, Wöstmann et al. (2015) had listeners compare spoken integers under varying levels of signal degradation. Listeners were asked to indicate whether an integer presented after a retention interval was larger or smaller than a prior integer. Significant relationships with signal degradation in the alpha-band were found, with no apparent effects in the theta-band. Similarly, Dimitrijevic et al. (2017) employed a task that required the retention of three consecutively presented spoken digits masked by noise. Though they did observe greater theta power in active compared to passive listening conditions, theta failed to parallel self-reports of effort in the active condition. It could be argued that, similar to non-auditory studies (e.g., Onton et al., 2005), frontal midline theta enhancements are only apparent under demanding working memory conditions. Retention of a single integer (cf. Wöstmann et al., 2015) may not be sufficiently taxing. Nevertheless, while alpha enhancements have been widely observed (for review, see Weisz and Obleser, 2014), analyses of frontal midline theta enhancements during effortful listening have yielded mixed results.

Here, we tested the hypothesis that frontal midline theta enhancements reflect working memory demands in a delayed pitch discrimination task. Listeners heard two square wave stimuli separated by a 3-s ISI at their ~70.7% correct pitch discrimination threshold. The task was to indicate which interval contained the higher pitch stimulus. In a 'Fixed' condition the low-frequency stimulus on each trial was fixed at 979 Hz. In a 'Roving' condition, the low-frequency stimulus was selected randomly on each trial. The important distinction is that in the 'Fixed' condition, listeners can accomplish the task in a single-interval manner. For instance, if a 979 Hz stimulus is heard first, the first stimulus can be designated the lowest without ever hearing the second. In contrast, the 'Roving' condition forces listeners to memorize the first stimulus long enough to compare it to the second. The 'Fixed' and 'Roving' conditions are indistinguishable on the single trial level, and have comparable accuracies, but differ in their working memory demands (Ahissar et al., 2006). We hypothesized that frontal midline theta enhancements would be stronger in the 'Roving' compared to the 'Fixed' condition. We also analyzed potential differences between conditions in the alpha band. Given the variety of tasks that have related alpha enhancements to listening difficulty, we expected to see alpha enhancements in both conditions.

1. Methods

1.1. Participants

Ten listeners (5 female, ages 19–32) were compensated for participation. Normal hearing was confirmed through audiometric testing (<20 dB HL, 0.25–8 kHz).

1.2. Stimuli and apparatus

Stimuli were square waves (250-ms, 10-ms on ramps, 240-ms linear decay) of varying fundamental frequencies ($f0$s). White noise was presented concurrently with tonal stimuli at a signal-to-noise ratio of +4 dB. Sounds were generated digitally via MATLAB 2014a and were presented over Etymotic ER-2 earphones (<81 dB SPL, fixed across listeners).

1.3. Procedures

Procedures were executed in MATLAB 2014a. A 2-interval, 2-alternative forced-choice delayed pitch discrimination task was used (see Fig. 1). On each trial an initial square wave stimulus (stimulus 1) was presented, followed by a 3-second ISI, and a second square wave stimulus (stimulus 2) with a different $f0$. White noise was gated on with stimulus 1 and off with the offset of stimulus 2. The continuation of white noise through the ISI served to mask any unintended background sounds. Listeners' task was to indicate which interval contained the higher pitched stimulus using a computer keyboard. Instructions were to withhold responding until sounds finished playing. No feedback was given. After the response, the next trial commenced after a variable ITI (uniform distribution 3.5–3.7-s).

In a 'Fixed' condition, each trial contained a 979-Hz stimulus in one interval and a higher pitch stimulus in the other. Pitch difference was adapted to track 70.7% correct accuracy. After incorrect responses, the percent frequency difference ($100 \times (f_{high}\text{-}f_{low})/f_{low}$) was doubled. After two consecutive correct responses, the frequency difference was halved. In a 'Roving' condition, the difference between the pitches of the two stimuli was also adjusted adaptively; however the low-pitched stimulus was drawn at random from trial to trial (840–1160-Hz; uniform distribution). Prior to the experiment, each listener was given a short (5 trial) practice block under each condition at a frequency separation of 25%. This served to inform participants as to the differences between conditions.

For each condition, an initial block (60 trials) was run, starting at a frequency difference of 4% to adapt an individual to his or her 70.7% correct threshold. These blocks were not analyzed. Four experimental blocks were then completed (2 'Roving' blocks, 2 'Fixed' blocks; 240 trials total). The first trial of a block started at the last set frequency difference for that condition. Blocks were pseudorandomly ordered such that the 1st and 2nd blocks were forced to be different conditions.

1.4. EEG acquisition and processing

A BioSemi Active II system (BioSemi, Amsterdam, Netherlands), recording at a 2048-Hz sampling rate, and 24-bit A/D resolution was used. Sixty-four electrodes were fixed within a cap and arranged according to the international 10–20 system. Six additional electrodes were placed at the mastoids, and on lateral sides and below each eye. Data were referenced online to the Common-Mode-Sense/Driven-Right-Leg (CMS/DRL) reference of the BioSemi system. Electrode offsets relative to CMS/DRL were brought within 25 μV or else were rejected from analysis.

Offline processing was performed using EEGLAB (Delorme and Makeig, 2004) and custom MATLAB scripts/functions. Data were referenced offline using an average reference, resampled at 256 Hz (after applying a zero-phase antialiasing filter), and then bandpass filtered between 0.5 and 100 Hz (FIR, order 1536). Channels and portions of continuous data contaminated by excessive noise or movement artifacts were removed based on visual inspection.

Full-rank extended infomax independent components analysis (ICA) was applied to each individual's data using the *binica()* function in EEGLAB. Independent components (ICs) were selected for rejection based on visual inspection of their activities and spectrum, then subsequently removed (for review and guidelines, see Makeig and Onton, 2009).

1.5. EEG analysis

Based on prior research (e.g., Wisniewski et al., 2015), a fronto-central group of channels was selected for analysis. These channels were: AFz, Fz, FCz, F1, and F2. Similarly, a group of occipital channels were selected for alpha enhancements: Oz, O1, O2, PO7, PO8 (cf. Wisniewski et al., 2017). Epochs of 7-s (from 2-s before stimulus 1 onset to 5-s after) were extracted. Each channel's event-related spectrum was computed using 7 cycle complex Morlet wavelets

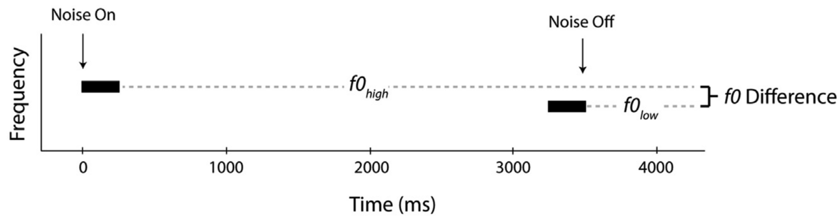

Fig. 1. A depiction of the employed delayed pitch discrimination task. On each trial, two stimuli of different *f0*s were presented, separated by a 3 s ISI. White noise (not shown) was presented at a signal-to-noise ratio of +4 dB and was gated on with the onset of the first stimulus and off with the offset of the second stimulus. The order of 'High' and 'Low' stimuli was randomized across trials. Listeners' task was to indicate which interval contained the higher pitch.

(3–50 Hz; ~0.35-Hz steps) centered at 400 time points (~12-ms intervals). The mean power spectrum from −1500-ms to 0-ms was used as a baseline for computing relative power in percentage (see Cohen, 2014).

1.6. Statistics

Conditions were compared on behavioral measures by means of paired-sample *t*-tests. For EEG data, a nonparametric permutation-based procedure was used. Mean relative theta and alpha power were examined in an *a priori* selected time-frequency window corresponding to 1250-ms post stimulus 1 onset to stimulus 2 onset from 4 to 8-Hz and 8–13-Hz.[1] For 1000 iterations, condition labels were shuffled and the means for conditions were recomputed. This process created a distribution of means expected under the null hypothesis. A *p*-value was considered the proportion of iterations having a more extreme difference between means than the actual data. Note that time-frequency windows begin well beyond the onset of the first stimulus, and thus minimize the influence of transient enhancements (e.g., Wisniewski et al., 2014). The design was not intended to yield unambiguous interpretation of transient stimulus-induced or evoked activity. For example, greater transient theta in the 'Roving' condition could be related to increased focus of attention (cf., Bernarding et al., 2013) or adaptation of transient responses in the fixed condition. We did not formally analyze these features.

2. Results & discussion

Accuracies were consistent with the 'Roving' and 'Fixed' conditions tracking ~70.7% correct performance (M = 68.80%, SE = 0.34). Mean frequency differences across trials in the 'Roving' and 'Fixed' conditions were $M = 3.43\%$ ($SE = 0.96$), and $M = 1.88\%$ ($SE = 0.81$) respectively. The 'Roving' condition showed a significantly greater mean frequency difference than the 'Fixed' condition, $t\,(9) = 3.87$, $p = 0.004$, *Cohen's* $d = 1.32$. This is consistent with previous research (for review, see Ahissar et al., 2006; Mathias et al., 2010). Median response times (relative to stimulus 2 onset) were 1.31-s (SE = 0.09) and 1.32-s (SE = 0.10) for 'Roving' and 'Fixed' conditions respectively. Response times were not significantly different, $p > 0.9$.

Event-related spectral perturbations (ERSPs) for 'Roving' and 'Fixed' conditions at the frontal midline group of electrodes are shown in the top panels of Fig. 2. The 'Roving' condition showed a clear band of theta enhancement during the ISI, while the 'Fixed' condition showed no such enhancement. Additionally, scalp maps of relative theta power demonstrate consistency with previous research on the topography of frontal midline theta (i.e., enhancements were strongest at frontal midline electrodes). Mean relative powers in the designated time-frequency windows (dashed rectangles) were M = 17.12% (SE = 9.08) for 'Roving', and M = −3.91% (SE = 3.93) for the 'Fixed' condition. Relative theta power was significantly greater in the 'Roving' than the 'Fixed' condition, $p = 0.018$. In addition, the difference ERSP ('Fixed' minus 'Roving') shows a clear band of increased theta power in the 'Roving' compared to the 'Fixed' condition.

Alpha enhancements were apparent during the ISI for both 'Roving' and 'Fixed' conditions (Fig. 2; bottom panels). The 'Roving' condition (M = 7.74%, SE = 5.54) showed slightly stronger alpha enhancement than the 'Fixed' condition (M = 2.61%, SE = 2.59) in the analysis window. However, no significant difference between conditions was found, $p > 0.30$.

Why have some works revealed frontal midline theta enhancements during listening, while others have not? We suspect that mixed findings in the literature can be explained by differences in task. For instance, Marsella et al. (2017) presented single words to children (8–10 years) with hearing aids in quiet, and in several masking conditions. While there was significantly greater alpha power when a masker was present, theta showed no significant differences between masker and silent conditions. Note that Marsella et al.'s paradigm requires identification of words, with no necessary retention across an interval longer than it takes to make a response. Also, the analyzed time-window was just prior to sound onset. If frontal midline theta enhancements are related to demands placed on working memory it should not be expected that theta relates to task difficulty in tasks that are not working memory based (cf., the 'Fixed' condition in the current study).

Still, other studies have explicitly used paradigms that entail short-term retention of auditory information, yet have failed to find theta-enhancements. Recently, we reported a study in which listeners performed an auditory delayed match-to-sample task with sounds varying in frequency modulation (FM) rate (Wisniewski et al., 2017). Trials in which the task was very easy showed very weak frontal-midline theta enhancements even though it was necessary to retain information in a retention interval. In contrast, difficult trials (at listeners' 70.7% correct thresholds) showed clear frontal-midline theta enhancement. Retaining a single number (Wöstmann et al., 2015), a series of digits (which could be chunked into a single number; Dimitrijevic et al., 2017), or order information (e.g., slow FM, then fast FM; Wisniewski et al., 2017) may not be as taxing on working memory as retaining the level of acoustic detail required to perform a difficult discrimination. Similar to non-auditory work on frontal midline theta enhancements and working memory, observing enhancements may necessitate that working memory is taxed to a sufficient degree (cf. Onton et al., 2005).

The current data fits well with a recently proposed model of effortful listening processes. Strauss and Francis (2017) proposed that two independent dimensions of attention contribute to effortful listening. "External attention" refers to endogenously

[1] Read-only files dated prior to data collection that designate *a priori* analysis parameters and hypotheses are available upon request from M.G.W.

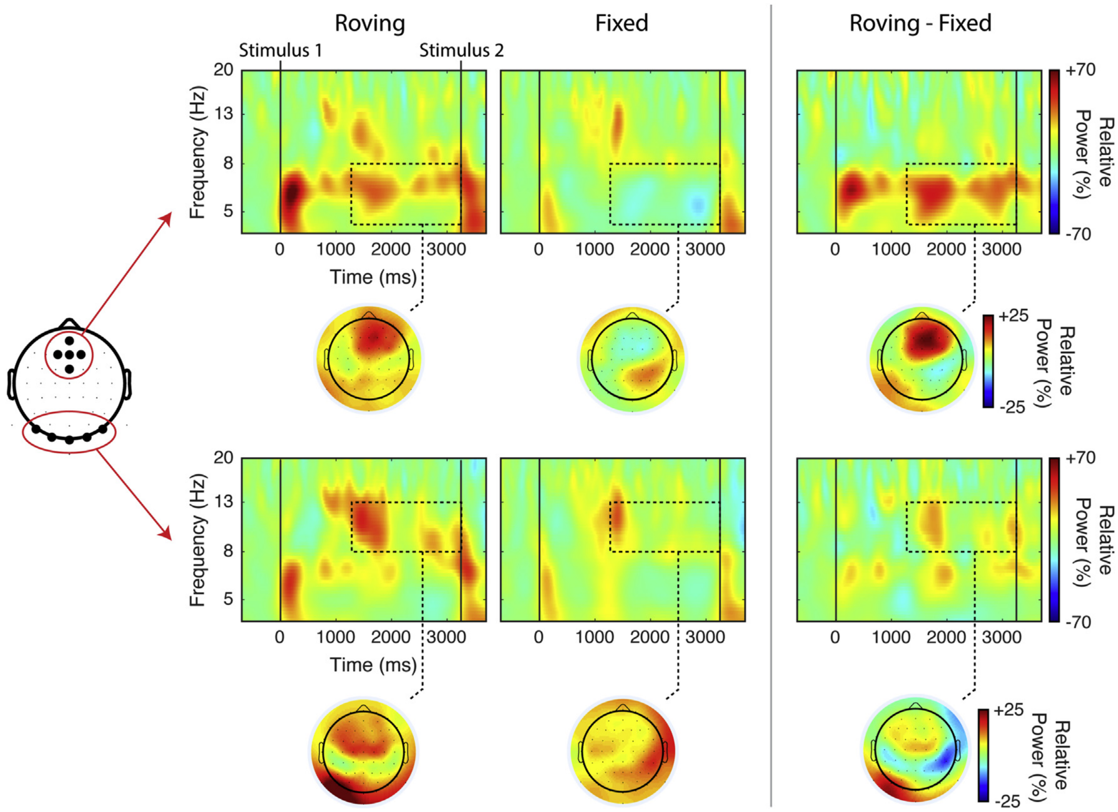

Fig. 2. Mean ERSPs in the 'Roving' and 'Fixed' conditions at the frontal midline (top) and occipital (bottom) electrode groups. The mean difference ERSPs, subtracting 'Fixed' ERSPs from 'Roving' ERSPs are also shown. Time-frequency windows used for analyses are depicted by the dashed rectangles. Scalp maps of relative power within these windows are shown.

controlled attention directed toward processing of externally generated object representations (e.g., competing speech streams). "Internal attention" refers to endogenously controlled attention operating on internally generated objects (e.g., representations in working memory). Two different listening tasks can be equally effortful, yet require distinct cognitive processes that rely on distinct neural substrates and/or dynamics. Some measures may reflect overall exerted effort, while others may be more sensitive to the different underlying processes. Frontal midline theta may be the latter, reflecting an internal attentional process. In contrast, alpha enhancements may be observed across a wider range of tasks because they reflect "external attention", both types of attention, or exerted effort in general.

3. Conclusions

Listening is inherently complex, and the processes involved are likely to vary depending upon the nature of the task. The current data suggest that frontal midline theta enhancements are related to working memory demands during listening. Specifically, a condition that was sufficiently difficult (at threshold performance), but did not require retention of a stimulus in working memory showed no enhancement. In contrast, a condition that required retention did show significant enhancement. Alpha enhancements, though present, did not differ between conditions. These results demonstrate that not all difficult listening tasks will be accompanied by the same EEG indices of listening effort.

Acknowledgments

This research was performed while M.G.W. held a National Research Council Associateship Award and an Oak Ridge Institute for Science and Education fellowship at the U.S. Air Force Research Laboratory. Justin Estepp is thanked for loaning of equipment and materials for EEG data collection.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.heares.2017.11.009.

References

Ahissar, M., Lubin, Y., Putter-Katz, H., Banai, K., 2006. Dyslexia and the failure to form a perceptual anchor. Nat. Neurosci. 9, 1558–1564.

Bernarding, C., Strauss, D.J., Hannemann, R., Seidler, H., Corona-Strauss, F.I., 2013. Neural correlates of listening effort related factors: influence of age and hearing impairment. Brain Res. Bull. 91, 21–30.

Bertoli, S., Bodmer, D., 2014. Novel sounds as a psychophysiological measure of listening effort in older listeners with and without hearing loss. Clin. Neurophysiol. 125, 1030–1041.

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

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.

Dimitrijevic, A., Smith, M.L., Kadis, D.S., Moore, D.R., 2017. Cortical alpha oscillations predict speech intelligibility. Front. Hum. Neurosci. 11, 88.

Hornsby, B.W., 2013. The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear Hear. 34, 523–534.

Kiessling, J., Pichora-Fuller, M.K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., et al., 2003. Candidature for and delivery of audiological services: special needs of older people. Int. J. Audiol. 42 (Suppl. 2), S92–S101.
Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B.G., Zekveld, A.A., Kramer, S.E., 2015. The pupil response reveals increased listening effort when it is difficult to focus attention. Hear. Res. 323, 81–90.
Mackersie, C.L., Cones, H., 2011. Subjective and psychophysiological indices of listening effort in a competing-talker task. J. Am. Acad. Audiol. 22, 113–122.
Makeig, S., Onton, J., 2009. ERP features and EEG dynamics: an ICA perspective. In: Luck, S., Kappenman, E. (Eds.), Oxford Handbook of Event-related Potentials. Oxford University Press, New York.
Marsella, P., Scorpecci, A., Cartocci, G., Giannantonio, S., Maglione, A.G., et al., 2017. EEG activity as an objective measure of cognitive load during effortful listening: a study on pediatric subjects with bilateral, asymmetric sensorineural hearing loss. Int. J. Pediatr. Otorhinolaryngol. 99, 1–7.
Mathias, S.R., Micheyl, C., Bailey, P.J., 2010. Stimulus uncertainty and insensitivity to pitch-change direction. J. Acoust. Soc. Am. 127, 3026–3037.
McGarrigle, R., Munro, K.J., Dawes, P., Stewart, A.J., Moore, D.R., 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.
Onton, J., Delorme, A., Makeig, S., 2005. Frontal midline EEG dynamics during working memory. Neuroimage 15, 341–356.
Pesonen, M., Björnberg, C.H., Hämäläinen, H., Krause, C.M., 2006. Brain oscillatory 1-30 Hz EEG ERD/ERS responses during the different stages of an auditory memory search task. Neurosci. Lett. 399, 45–50.
Rabbitt, P.M.A., 1991. Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. Acta Otolaryngol. 476 (Suppl. l.), 167–176.
Rönnberg, J., Rudner, M., Foo, C., Lunner, T., 2008. Cognition counts: a working memory system for ease of language understanding (ELU). Int. J. Audiol. 47 (Suppl. 2), S99–S105.
Strauss, D.J., Francis, A.L., 2017. Toward a taxonomic model of attention in effortful listening. Cognit. Affect. Behav. Neurosci. 17, 809–825.
Weisz, N., Obleser, J., 2014. Synchronisation signatures in the listening brain: a perspective from non-invasive neuroelectrophysiology. Hear. Res. 307, 16–28.
Wisniewski, M.G., 2017. Indices of effortful listening can be mined from existing electroencephalographic data. Ear Hear. 38, e69–e73.
Wisniewski, M.G., Mercado III, E., Church, B.A., Gramann, K., Makeig, S., 2014. Brain dynamics that correlate with effects of learning on auditory distance perception. Front. Neurosci. 8, 396.
Wisniewski, M.G., Thompson, E.R., Iyer, N., 2017. Theta- and alpha-power enhancements in the electroencephalogram as an auditory delayed match-to-sample task becomes impossibly difficult. Psychophysiology 54, 1916–1928.
Wisniewski, M.G., Thompson, E.R., Iyer, N., Estepp, J.R., Goder-Reiser, M.N., Sullivan, S.C., 2015. Frontal midline θ power as an index of listening effort. Neuroreport 26, 94–99.
Wöstmann, M., Herrmann, B., Maess, B., Obleser, J., 2016. Spatiotemporal dynamics of auditory attention synchronize with speech. Proc. Natl Acad. Sci. U. S. A. 113, 3873–3878.
Wöstmann, M., Herrmann, B., Wilsch, A., Obleser, J., 2015. Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. J. Neurosci. 35, 1457–1458.