

Three New Outcome Measures That Tap Into Cognitive Processes Required for Real-Life Communication

Thomas Lunner,^{1,2,3,4} Emina Alickovic,^{1,3} Carina Graversen,¹ Elaine Hoi Ning Ng,^{2,5}
Dorothea Wendt,^{1,4} and Gitte Keidser^{1,2}

To increase the ecological validity of outcomes from laboratory evaluations of hearing and hearing devices, it is desirable to introduce more realistic outcome measures in the laboratory. This article presents and discusses three outcome measures that have been designed to go beyond traditional speech-in-noise measures to better reflect realistic everyday challenges. The outcome measures reviewed are: the Sentence-final Word Identification and Recall (SWIR) test that measures working memory performance while listening to speech in noise at ceiling performance; a neural tracking method that produces a quantitative measure of selective speech attention in noise; and pupillometry that measures changes in pupil dilation to assess listening effort while listening to speech in noise. According to evaluation data, the SWIR test provides a sensitive measure in situations where speech perception performance might be unaffected. Similarly, pupil dilation has also shown sensitivity in situations where traditional speech-in-noise measures are insensitive. Changes in working memory capacity and effort mobilization were found at positive signal-to-noise ratios (SNR), that is, at SNRs that might reflect everyday situations. Using stimulus reconstruction, it has been demonstrated that neural tracking is a robust method at determining to what degree a listener is attending to a specific talker in a typical cocktail party situation. Using both established and commercially available noise reduction schemes, data have further shown that all three measures are sensitive to variation in SNR. In summary, the new outcome measures seem suitable for testing hearing and hearing devices under more realistic and demanding everyday conditions than traditional speech-in-noise tests.

Key word: Ecological validity.

Abbreviations: SNR, signal-to-noise ratio; SWIR, Sentence-final word Identification and Recall; HA, hearing aid; MEG, magnetoencephalography; EEG, electroencephalography; NR, noise reduction; FUEL, Framework for Understanding Effortful Listening; ELU, Ease of Language Understanding

(*Ear & Hearing* 2020;41:39S–47S)

INTRODUCTION

In the laboratory, hearing and hearing devices are traditionally evaluated using outcome measures that have been designed to maximize control of the independent variables and

sensitivity. These measures often have little resemblance to demanding real-life listening situations. For example, popular speech perception tests, which are frequently used in hearing research, are usually configured to measure performance at the steepest (most sensitive) part of the psychometric function, which is around the 50% correct point. This point, together with the use of idealized stimuli and a simplified acoustic reproduction of the listening environment, tends to push testing to occur at signal-to-noise ratios (SNR) much poorer; that is, lower, than those found for real-world listening and communication situations (Smeds et al. 2015) and those that hearing devices have been designed to function optimally in (Naylor, 2016). Administering such tests in the laboratory at higher and more realistic SNRs tend to push performances to ceiling, that is, 100% correct. To increase the ecological validity of outcomes from laboratory evaluations of hearing and hearing devices, it is necessary to partly obtain a better understanding of the real-life situations that are relevant for people with hearing problems and hearing device users—the outside-in approach—and partly introduce more realistic listening environments and outcome measures in the laboratory—the inside-out approach. Progress in both directions have been made in recent years and are reviewed in several publications in this issue (Carlile & Keidser 2020; Hohmann et al. 2020; Smeds et al. 2020).

Of the steps required to increase ecological validity of laboratory outcome measures, the development of new outcome measures that both possess good psychometric characteristics and provide meaningful information about real-life function is particularly challenging. As mentioned earlier, frequently used laboratory tests designed to measure speech perception; in particular, the ability to perceive and recognize phonemes, single words, or simple sentences, have very desirable psychometric characteristics. However, correct speech perception depends almost entirely on the auditory processing skills of a person, whereas a more relevant real-life situation would require listening to ongoing discourse that additionally involves different cognitive processes, such as interpretation of information, decision making, turn-taking, and retrieving events from memory. With reference to the World Health Organization's International Classification of Functioning, Disability, and Health (WHO-ICF) framework (WHO 2001), Kiessling and colleagues (Kiessling et al. 2003) outlined four processes involved in the auditory domain. These are (1) hearing, the detection of sound; (2) listening, the process of intentional and attentional hearing; (3) comprehending, the extraction of meaning and information that follows listening; and (4) communicating, which refers to an interactive and bidirectional way of exchanging meaning and information. While auditory processing is fundamental to all these functions, listening, comprehending, and communicating depend to a great extent on cognitive functions.

¹Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark; ²Department of Behavioural Sciences and Learning, Linnaeus Centre HEAD, Linköping University, Linköping, Sweden; ³Department of Electrical Engineering, Division Automatic Control, Linköping University, Linköping, Sweden; ⁴Department of Health Technology, Hearing Systems, Technical University of Denmark, Lyngby, Denmark; ⁵Oticon A/S, Kongebakken, Denmark.

Copyright © 2020 The Authors. *Ear & Hearing* is published on behalf of the American Auditory Society, by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

It is well established that there is a link between cognitive abilities and the listening process as measured with sentence recognition tests (Dryden et al. 2017; Gatehouse & Gordon 1990; Lunner 2003). For example, in one of the classic studies on hearing and cognition, Lunner (2003) reported that, for first-time hearing aid (HA) users, good cognitive abilities, measured using the reading span (Daneman & Carpenter, 1980; Rönnberg et al. 1989) and rhyme judgment (Lyxell 1994) tests, positively correlated with speech recognition performance obtained in both aided and unaided conditions. In other words, listeners with better working memory and phonological processing skills had better speech recognition scores. Recent data (Ng & Rönnberg 2020) have demonstrated that such a relationship also exists in long-term HA users; albeit only when speech was presented in a changing-state background noise, which better resembled daily listening environments than more conventional stationary test background noises. Other cognitive abilities, including speed of information processing and lexical access have also been found to be associated with speech recognition (Hällgren et al. 2001; Lyxell et al. 2003; Rönnberg et al. 2008). In addition, executive functions, in particular inhibition, which is the process involved in preventing irrelevant information from entering working memory, as well as attention, the process of selecting and limiting the amount of information entering or remaining in working memory, play important roles in higher-level speech processing such as comprehension and communication (Atkinson & Shiffrin 1968).

Two listeners may obtain similar scores on traditional speech perception tests, but with different mental effort applied, meaning that one listener can sustain a conversation in an everyday listening situation for longer than the other. Different techniques/measures, including functional brain imaging, physiological responses, or behavioral responses, have been introduced to better understand cognitive processes exerted when listening to speech in noise (Peelle 2018; Strauss & Francis 2017; Pichora-Fuller et al. 2016). In particular, there is sufficient evidence that listening effort, as measured by pupillometry, can reveal speech processing difficulties that are not necessarily reflected by traditional speech perception measures (Ohlenforst et al. 2018; Wendt et al. 2017). Listening effort was recently defined as “the deliberate allocation of resources to overcome obstacles in goal pursuit when carrying out a listening task” (Pichora-Fuller et al. 2016). According to the Framework for Understanding Effortful Listening (FUEL), effort mobilization consumes cognitive resources and is dependent on the task demands, but can be further affected by motivation, fatigue, or psychosocial aspects (Hopstaken et al. 2015; Peelle & Wingfield 2016; Richter et al. 2016).

In summary, cognitive ability has strong links to processes related to real-life listening and conversation situations, and hence the ecological validity of outcome measures is likely to be very much affected by the degree to which they engage cognitive abilities in realistic ways (Keidser et al. 2020). Therefore, new outcome measures aiming to capture a person’s ability to participate in everyday communications should ideally tap into cognitive functions applied during real-life listening and conversation situations. To be useful for hearing-related research, such measures should be sensitive to acoustic contrasts, that is different background noises and different device signal processing algorithms that can vary the SNR. In this article, we review some newer outcome measures, the Sentence-final

Word Identification and Recall (SWIR) test, neural tracking and pupillometry, that aim to better measure real-world listening demands by tapping into working memory, selective attention, and listening effort, respectively. The sensitivity of each measure to acoustic contrasts has been evaluated, using various noise reduction (NR) algorithms. The measures are all developed to better understand how the interaction between hearing and cognition affects a person’s listening and communication skills; an understanding that potentially will lead to the development of improved hearing interventions and candidacy criteria. Thus, such measures support Purpose A (Understanding) directed toward better understanding of the role of hearing in everyday life, Purpose B (Development) directed toward supporting the development of improved hearing-related procedures and interventions, and Purpose C (Assessment) directed toward facilitating improved methods for assessing and predicting the ability of people and systems to accomplish specific real-world hearing-related tasks for increasing the ecological validity of hearing-related research (Keidser et al. 2020).

THE SWIR TEST—A WORKING MEMORY PARADIGM

The SWIR test (Ng et al. 2013) was developed specifically to investigate the cognitive benefit of NR schemes during the simultaneous perception and processing of speech, predicted by the Ease of Language Understanding (ELU) framework (Rönnberg 2003; Rönnberg et al. 2008, 2013). The ELU framework conceptualizes the interplay between individual differences in speech recognition and working memory. Working memory is the limited capacity for the simultaneous storage and online processing of information (Baddeley & Hitch 1974). The ELU framework predicts that in an easy listening condition where the input signal is clear, implicit and effortless processing occurs. In this case, the input signal matches readily with the phonological representation in long-term memory. When the listening condition is challenging and suboptimal, for instance when the input signal is distorted as a perceptual consequence of cochlear damage or masked by noise, mismatch occurs because the input signal cannot be readily matched with the phonological representation in the lexicon. This requires top-down remedial processing, which is then explicit and effortful and loads the capacity-limited working memory. Advanced signal processing in hearing devices, such as NR schemes, is designed to improve speech intelligibility by minimizing the impact of noise. Based on the ELU framework, it can be predicted that NR schemes reduce the engagement of explicit processing, leaving more cognitive resources for higher-level processing of auditory input, such as speech recall.

In the SWIR test, listeners are presented with a number of short sentences. There are two tasks performed in sequence: an *identification task*, in which listeners repeat the final word immediately after listening to each sentence in the list, followed by a *free recall task*, in which listeners at the end of the list recall, in any order, the final words that have been successfully repeated in the identification task. Relative to the traditional speech recognition tests, where listeners are asked to listen to and repeat speech-in-noise stimuli in mostly negative SNRs, the SWIR test has several characteristics that are assumed will increase the ecological validity of its outcome. First, because the recall task of the SWIR test is only meaningful if the words

in the identification task were correctly identified, or heard, the speech and noise stimuli are presented at positive and individualized SNRs targeting 95% speech intelligibility [average SNRs were +4.2 and +7.5 dB in Ng et al. (2013) and Ng et al. (2015), respectively]. These SNRs are representative of typical real-life listening conditions (Smeds et al. 2015). Second, the SWIR test simultaneously assesses perception and higher-level processing (memory recall) of speech, which taps into working memory. Additionally, target speech and background noise are spatially separated, with the background noise presented from multiple spatially separated locations, which better parallels realistic acoustic environments.

Testing the Effect of Noise Reduction Algorithms

Using the SWIR test, the effect of active NR processing was evaluated in listeners with hearing impairment in a series of studies. In Ng et al. (2013) and Ng et al. (2015), the test stimuli were preprocessed using binary masking NR schemes (Boldt et al. 2008), whereas a NR scheme implemented in a commercially available HA was used in Ng et al. (2016). In agreement with the ELU framework prediction, results showed that active NR processing alleviated the negative impact of noise on working memory to improve recall for highly intelligible speech by freeing up cognitive resources. This finding was successfully replicated in Lunner et al. (2016), using the same test paradigm but in a different language (see Fig. 1). Of particular interest, the positive effects of NR processing on recall performance were modulated by individual differences in working memory capacity, as measured with the reading span test in Ng et al. (2013), with only listeners with good working memory capacity showing benefit from the NR schemes. In the subsequent studies (Lunner et al. 2016; Ng et al. 2015, 2016), the SWIR test was modified by reducing the list length from eight to seven (i.e., the number of to-be-recalled words), such that the test became less cognitively demanding. These studies showed that when presenting shorter lists, the use of NR processing improved recall performance, regardless of individual

differences in working memory capacity, when speech intelligibility was approaching 100%.

Discussion and Future Directions

The SWIR test builds on traditional speech recognition tests by tapping into listeners working memory capacity while they listen to and process simple sentences and has been found to be sensitive to NR schemes. Based on the results reported by Ng et al. (2013, 2015, 2016) and Lunner et al. (2016), the positive effects of NR processing may not be captured if task difficulty, determined by the number of to-be-recalled words or length of the sentence lists, exceeds individual working memory capacity. Ongoing investigations focus on whether the SWIR test can be further modified by varying the task difficulty, which is the working memory demands, so that it is adaptive to individual cognitive capacity. As a first step, Micula et al. (2020) examined whether the positive benefit of NR processing on recall performance is dependent on task difficulty. This was done by introducing varying list lengths. Results indicated that the SWIR test with varying list lengths can reliably detect the positive effect of NR processing on recall. Future investigations will focus on the implementation of an adaptive procedure so that the SWIR test will be adaptive to individual cognitive capacity. All previous and current investigations concerning the SWIR test have tested people with presumably normal cognitive functions. Future investigations could also focus on the implications for testing older HA users with particular pathologies that affect cognition.

NEURAL TRACKING AS A TOOL TO STUDY SELECTIVE AUDITORY ATTENTION

The assessment of selective auditory attention (i.e., listener's ability to attend selectively to a specific talker in cocktail party-like situations with multiple competing speakers) is gaining increased interest in the field of Hearing Science. Much of the progress in the area is closely tied to progress in developing computational models that can describe how the auditory system encodes the incoming speech streams. The speech sounds evoke neural responses that are heavily modulated by attention, showing selective neural tracking of attended speech, that is, selective enhancement of neural responses to attended versus ignored sounds (Mesgarani & Chang 2012; Ding & Simon 2012; O'Sullivan et al. 2015). Measures of neural tracking not only make it possible to infer what the listener is attending to but also to infer how well the signal is encoded (e.g., Das et al. 2018). In Choi et al. (2014), a relationship between neural amplification and performance was found across several auditory listening tasks; listeners with stronger neural amplification were better performers. Therefore, combined with listening tasks designed to resemble everyday listening situations, the use of neural tracking could provide objective and more ecologically valid information about how well a listener is attending to and understanding a particular talker.

Literature provides three powerful interrelated computational models from linear system identification theory (Ljung 1999) to investigate neural tracking (Alickovic et al. 2019). Briefly, the overall goal of these computational models is to identify a linear kernel (or response function) that best describes how the neural responses captured using magnetoencephalography (MEG) or electroencephalography (EEG) are correlated to the envelopes

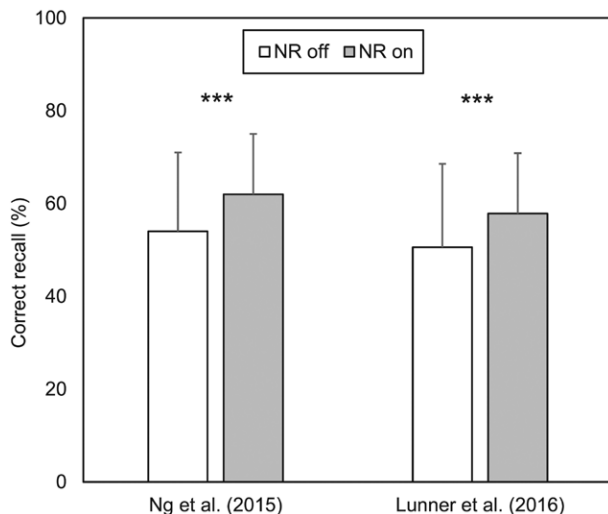


Fig. 1. The effect of active noise reduction processing on correct recall of words as measured with the SWIR test implemented in Swedish (Ng et al. 2015) and Danish (Lunner et al. 2016). Asterisks indicate the differences were highly significant ($p < 0.001$). Error bars represent standard deviations.

of the incoming speech signals (Lalor et al. 2009; Dmochowski et al. 2018; de Cheveigné et al. 2018; Crosse et al. 2016). First, the decoding model [or stimulus reconstruction (SR) model] that links brain and speech in a backward direction, attempts to learn the incoming speech envelope from the speech-induced neural responses by modeling the speech envelope as a linear convolution of the EEG or MEG responses and estimated linear kernel (or decoder) (Ding & Simon 2012; Mesgarani & Chang, 2012; O'Sullivan et al. 2015; Das et al. 2018; Presacco et al. 2019). Second, the encoding model that links brain and speech in forward direction, attempts to predict the EEG or MEG responses from the speech envelope(s), by modeling the neural responses as a linear convolution of the speech envelope and a linear kernel (or temporal response function) (Lalor et al. 2009; Ding & Simon 2012; Alickovic et al. 2016). Finally, more recently, a hybrid model that combines the strengths of the encoding and decoding models and that is based on a canonical correlation analysis has also been proposed (Dmochowski et al. 2018; de Cheveigné et al. 2018; Iotzov & Parra 2019).

Of the three models, the SR model has received the greatest attention in the literature and has inspired the development of different measures of assessing different aspects of the listener's ability to attend selectively to a specific talker in a cocktail party environment. The different aspects include: impact of focused attention (O'Sullivan et al. 2015; Fiedler et al. 2017); impact of hearing loss (Presacco et al. 2019); impact of listener's age (Brodbeck et al. 2018a; Decruy et al. 2019); impact of acoustic environment (Fuglsang et al. 2017; Das et al. 2018; Khalighinejad et al. 2019); and impact of speech intelligibility (Ding & Simon 2012; Vanthornhout et al. 2018; Lesenfants et al. 2019).

One key finding derived using the SR model is that the strength of EEG-based neural tracking is indicative of whether or not the listener is attending to a specific talker in multitalker situations (O'Sullivan et al. 2015; Mirkovic et al. 2015; Akram et al. 2016). However, neural tracking from single-trial EEG and MEG has been mainly investigated in tightly controlled listening conditions with two spatially separated competing talkers without accompanying background noise or at favorable SNRs. With their simplified listening conditions that lack the dynamics and multiple competing talkers of real-world cocktail party environments, these controlled experiments may reduce the level of ecological validity of outcomes and hence limit the progress in truly understanding the ability to attend selectively to a specific talker in a cocktail party environment.

Sensitivity to Acoustic Contrast and Hearing Loss

Only recently have data been collected to investigate the effect of noise level and speaker position on neural tracking of speech in normal-hearing listeners. Das et al. (2018) demonstrated that the speech-related brain activity, and thus the neural tracking of speech, was very sensitive to background noise level and angular speaker separation. Specifically, it was found that increasing noise levels degrades the quality of neural tracking of attended speech, with the degrading effect of the noise being stronger when the competing talkers were located closer to each other. More specifically, it was shown that the quality of neural tracking was significantly degraded at negative SNRs, that is, with increasing background noise levels. On the other hand, it was reported that the quality of neural tracking improves with angular speaker separation, showing the benefit of spatial release from masking.

Recent publications have further reported that hearing loss plays an important role in altering the neural tracking of speech (Petersen et al. 2016; Presacco et al. 2019), with the brain's ability to perform selective attention tasks especially challenged when background noise is present. Specifically, Petersen et al. (2016) reported that hearing loss reduced the neural tracking of speech with poorer SNR, with the differential neural tracking of attended versus ignored speech further decreasing with poorer hearing, while Presacco et al. (2019) reported that hearing loss reduced differential midbrain and cortical responses. The implications of these results are that measures of changes in neural tracking of speech attention in noise could be sufficiently sensitive to changes in the acoustic input; for example, resulting from HA signal processing designed to reduce interference from competing talkers under everyday conditions. That is, neural tracking can be used as an objective measure to reveal if such acoustic changes are effective in improving (restoring to normal patterns of neural activity) speech attention measures for listeners with hearing loss.

Testing the Effect of Noise Reduction Algorithms

Modern HAs can alleviate the impact of background noise by applying advanced NR signal processing algorithms while providing adequate access to the information of the foreground speech (Chung 2004; Dillon 2012). However, state-of-the-art HAs do not know which speech sources (talkers) the HA user's attention is focused on, leaving it to the brain to enhance the talker of interest and suppress unwanted sounds. For HA users to succeed in complex listening situations, it is therefore of interest to develop HA compensation strategies that are supportive of the brain's selective auditory attention processing. For this purpose, neural tracking of speech may be used to objectively measure how well different HA settings achieve enhancement of the talkers of interest in the foreground and suppression of the noise in the background.

In a recent study (Alickovic, et al., Reference Note 1), the role of a commercially available NR scheme on selective auditory attention using neural tracking of speech was examined. Twenty-two experienced HA users with mild-to-moderate sensorineural hearing loss were presented with two talkers in a background of a four-talker babble noise presented uncorrelated from four loudspeakers in the rear hemisphere (see upper right corner in Fig. 2). Listeners were instructed to focus on one of the talkers and ignore the competing talker and background noise. The angular locations of attended and ignored talkers in the foreground were $\pm 22^\circ$, aimed to simulate a realistic listening situation. The attended and ignored talkers were each presented at 62 dB SPL (decibels sound pressure level) and had an SNR of +3 dB or +8 dB relative to the background noise, simulating a perceptually difficult and easy listening condition, respectively. EEG activity was recorded via 64 scalp electrodes. Neural tracking of speech was estimated using the SR method and quantified by the correlation between actual speech envelope (at the input of the hearing aid) and reconstructed speech envelope. Neural tracking of speech was compared for active versus inactive NR processing. The commercially available NR processing was set to improve SNR of speech coming from the two frontal loudspeakers (i.e., both the attended and ignored talkers) by predominantly attenuating the background noise coming from the four background loudspeakers. Technical measurements were performed in both +3 and +8 dB SNR conditions to verify the

SNR improvement in the HA, which is defined as the difference between the input level of the environment (two frontal talkers and background noise from the four loudspeakers) and output responses from the HA. On average, the articulation-index weighted SNR improvements were between 5 and 6 dB for active NR processing compared with when it was off.

Our results showed that under more difficult listening (i.e., at +3 dB SNR) the neural tracking of the attended and ignored talkers was significantly enhanced by the activation of NR, while the neural tracking of background noise was significantly suppressed (see Fig. 2). Overall, these results demonstrate, that active NR processing reduces background noise sufficiently to enable the HA user to better follow the talker of interest (attended talker), and switch their attention to another (i.e., ignored) talker when necessary. Altogether, these results provide evidence that NR activation in HAs can enhance the neural tracking of speech in complex listening situations.

Discussion and Future Directions

Popular neural tracking measures, such as SR, have been found to reflect selective attention in listeners and to be sensitive to hearing loss and changes in SNR. As noted by Petersen et al. (2016), the differential neural tracking of attended versus ignored speech further decreases with poorer hearing. Thus, it seems as if listeners with hearing loss have more difficulty segregating between attended and ignored speech, which may impair their ability to attend to the desired source, or vice versa. Understanding selective auditory attention in listeners with hearing loss is a step toward better understanding the effects of hearing impairment on selective listening in a cocktail party situation and

how HAs, including NR strategies, can provide benefits. Here, we have shown that a HA signal processing that improves SNR may mitigate the attention loss caused by hearing loss.

Neural tracking measures hold a potential to delineate how higher-order acoustic, phonetic, and linguistic speech features are represented and integrated at different levels of the auditory system. For example, Brodbeck et al. (2018a) showed that, for normal-hearing listeners, at the linguistic level the neural tracking was completely dominated by the attended stream. A research priority for the future is to investigate if that is also the case for listeners with hearing loss under more realistic test conditions.

A potential issue with the neural tracking paradigm is a lack of control over to what extent test participants just focus on a particular voice without listening with intent to understand what has been said. This poses a challenge for the implementation of a speech tracking task that can also be used to verify to what extent the participant could follow what was being said. Recent results by Brodbeck et al. (2018b) indicates that, while responses to acoustic features reflect attention through selective amplification of attended speech as discussed above, responses consistent with a lexical processing model reveal categorically selective processing. Therefore, by analyzing neural tracking at both the acoustic feature level and lexical level, it might be possible to distinguish between just surfacing a voice or if the speech attended to is actually understood.

PUPILLOMETRY AS A TOOL TO MEASURE LISTENING EFFORT

Measurements of listening effort have been realized using different approaches including subjective, behavioral, and

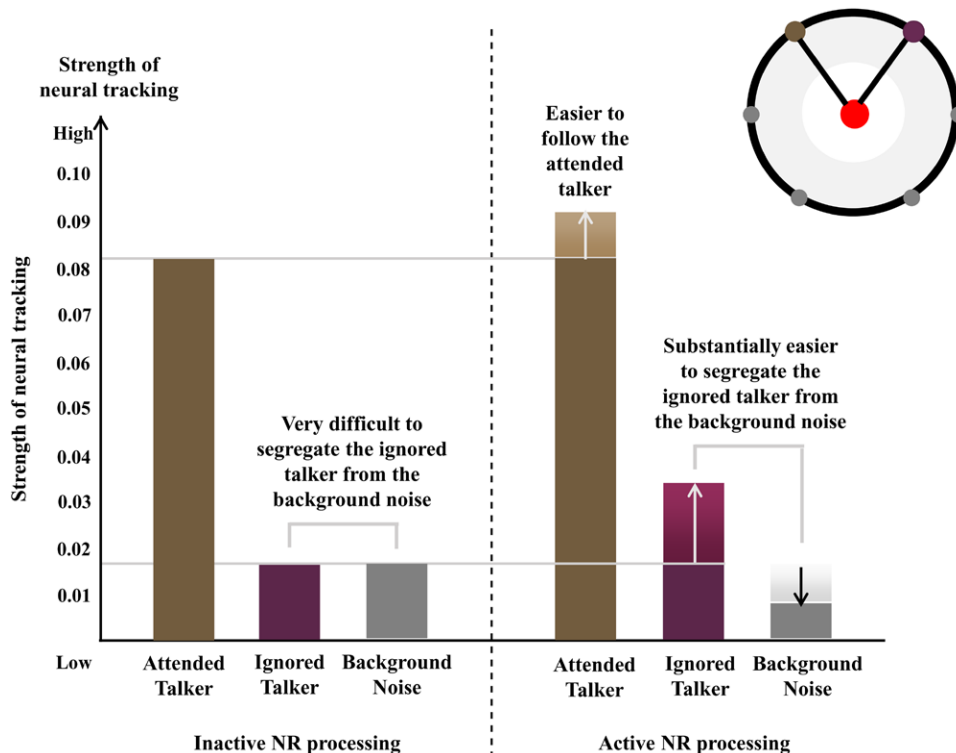


Fig. 2. The strength of neural tracking of an attended talker (brown bar), an ignored talker (purple bar), and the background noise (gray bar), when NR processing is inactive (left) and active (right). Data showed for the +3 dB SNR condition. The diagram in the upper right corner shows the spatial configuration of attended (brown dot) and ignored (purple dot) talkers and background noise (pale blue dots), relative to the listener (center), in the circular test setup.

physiological methods (see Ohlenforst et al. 2017a). The latter includes pupillometry, which has been applied to illustrate changes in autonomic nervous system activity during task performance (McGarrigle et al. 2014; Winn et al. 2018), with a combination of sympathetic and parasympathetic nervous system contributions determining the pupil size (Loewenfeld & Lowenstein 1993). Changes in the pupil size have been associated with locus coeruleus activity (Koss 1986; Aston-Jones & Cohen 2005). More particularly, the task-evoked pupil response has been suggested to index phasic activity reflecting the momentary processing of an attended task-relevant event (Aston-Jones & Cohen 2005), whereas the resting state or baseline pupil size might index the baseline arousal level (Jepma & Nieuwenhuis 2011; Murphy et al. 2014).

Sensitivity to Acoustic Contrast, Including Noise Reduction Algorithms

Pupil dilation can signal changes in mental task load and effort (Kahneman & Beatty 1966; Beatty 1982). In recent years, a considerable number of studies have measured people's pupil dilation while they were listening to speech and demonstrated that pupil dilation changed with, for example, changing SNR (Ohlenforst et al. 2017b), intelligibility (Zekveld et al. 2010), masker type (Koelewijn et al. 2012; Wendt et al. 2018) or linguistic complexity (Wendt et al. 2016). Of interest is that pupil dilation has been found to be sensitive to changes in acoustic degradation of the listening situation whether speech perception performance is changing or not (Kramer et al. 1997; Wendt et al. 2018). This means, that pupillometry can be used as an outcome measure at SNRs that better resemble those experienced in real life and where the traditional speech perception tests become highly insensitive to different backgrounds and HA signal processing.

Furthermore, research has demonstrated reduced listening effort with an active NR scheme at speech intelligibility levels above 50% speech recognition (Ohlenforst et al. 2018; Wendt et al. 2017). For example, in Wendt et al. (2017), pupil dilation was measured on 24 aided listeners with hearing loss while they completed the Danish HINT test presented in a four-talker babble noise. In a first experiment, it was found that pupil dilation was significantly reduced when a commercially available NR scheme was activated compared with when it was off. In a second experiment, two different NR schemes were compared that are both implemented in modern HAs. One uses a traditional single microphone noise estimate and consists of an adaptive first-order directionality system combined with slow-acting NR processing, whereas the other uses a faster NR scheme (see Wendt et al. 2017 for details). Pupil dilation was significantly reduced when listening via the faster NR scheme. The effect of active NR processing on pupil dilation was further demonstrated over a wide range of SNRs (Ohlenforst et al. 2018).

Discussion and Future Directions

Recent hearing-related research has demonstrated that pupillometry is adding another dimension when evaluating speech perception of people with and without hearing impairment. Typically, literature on pupillometry and listening effort, using traditional speech-in-noise tests, apply trial-based designs by presenting single sentences or words in background noise (Winn et al. 2018). It has been argued that more natural speech

stimuli such as conversations, news clips, or monologues would enable studying cognitive processes involved in natural speech processing that are more relevant for everyday listening scenarios (Alexandrou et al. 2018; Hamilton & Huth, 2018), and newer studies have started to examine this using pupillometry. Interestingly, studies using longer (speech) stimuli often report a different temporal characteristic of the pupil response curve compared with the studies examining task-evoked (phasic) response of the pupil (e.g., Zhao et al. 2019; Hjortkjær et al. 2018). More research is currently needed to determine how and if, the characteristics of the pupil response curve obtained for longer (speech) stimuli correlate to the task-evoked pupil response as it is observed in literature applying single sentences, or provide a different outcome that better represents cognitive processes engaged during real-world listening.

According to the FUEL framework (Pichora-Fuller et al. 2016), listening effort increases with increasing task demands. However, this is only valid if the individual stays engaged in the task (Granholtm et al. 1996; Pichora-Fuller, et al. 2016). With increasing task demands and when available resources are exceeded, effort allocation declines due to disengagement. Thus, a decline in pupil dilation might not always refer to a reduced effort allocation but may be interpreted as signs of disengagement or giving-up (Granholtm et al. 1996; Ohlenforst et al. 2017a; Wendt et al. 2018). This might limit the application of pupillometry as a measure of listening effort in situations with high task demands, for example, in situations with low intelligibility levels, since reduced dilation could suggest disengagement of the listener. Alternatively, pupillometry might also provide a tool to detect signs of disengagement and “giving-up” in people with hearing impairment. It has furthermore been argued that motivation can affect effort allocation as indicated by changes in pupil size. In a pupillometry study, Koelewijn et al. (2018) examined whether motivation (manipulated by monetary reward) had a mediating effect on listening effort as reflected by the pupil dilation. Their results indicated that reward impacted effort allocation even when speech recognition stayed unchanged. Those findings were interpreted in line with the motivational intensity theory, which states that effort mobilization can be affected by motivational arousal (Brehm & Self 1989).

Further research toward a better understanding of the effects of effort mobilization under more realistic conditions should focus on the interplay of task demands, motivation, fatigue, and social aspects of communication. On the one hand, consequences of social stress on effort mobilization might be relevant for everyday communication as people with hearing impairment may become stressed when experiencing communication difficulties (Pichora-Fuller 2016). On the other hand, being able to follow a conversation as well as staying engaged in a conversation might be relevant and highly motivating for people with hearing impairment. Thus, social motivation as well as identifying signs of disengagement/giving-up (and how hearing devices might help the user to stay engaged) using pupillometry will help to better understand resource allocation and effort mobilization in people with hearing impairment in everyday communication situations. One known downside of the popular pupillometry method, when it comes to sourcing more realistic testing scenarios, is probably its sensitivity to luminance, which is difficult to control outside the typical laboratory situation.

SUMMARY AND CONCLUSION

Three newer outcome measures (the SWIR test, neural tracking, and pupillometry) that aim at tapping into cognitive processes associated with real-life listening and conversation situations were presented and discussed, with the three measures capturing working memory capacity, selective attention, and listening effort, respectively. Several studies revealed that both the SWIR test, which asks listeners to recall items they have just been listening to, and listening effort, as indicated by changes in pupil dilation while listening to speech, can add another dimension to more traditional speech perception testing, since they provide a sensitive measure in situations where speech perception performance might be unaffected. As such, changes in working memory capacity and effort mobilization can be found at positive SNRs, that is, in an SNR range that might better reflect everyday situations. As for neural tracking, several studies, using SR, have demonstrated that this method is robust at determining whether a listener is attending to a specific talker in a typical cocktail party situation. Therefore, all three measures appear to be useful for testing specific cognitive processes under more realistic communication scenarios. Using both established and commercially available NR schemes, data further revealed that all three measures are sensitive to variation in SNR, and therefore, can be used to evaluate HA signal processing designed to improve this variable.

As a caution, it should be noted that none of the three measures are standardized and some potential caveats have been identified for each measure that should be considered if implemented and used in future hearing-related research studies. For the SWIR test, this means being mindful of the list length, that is, the number of sentences presented and thus number of words to recall, as the test becomes insensitive if the number of words to recall exceeds individual working memory capacity. In terms of recording neural responses while listening to speech, the speech tracking task should ideally be selected so that it can be verified to what extent the speech attended to has been understood. Finally, when recording pupil dilation, considerations should be made to the demand of the test situation, as very demanding listening situations may show reduced listening effort as a result of the listener giving-up on the task. Generally, HA signal processing parameters have been optimized to traditional speech-in-noise measures, which is insensitive to many everyday conditions, such as typical real-life SNRs, and the cognitive processes involved in understanding speech in challenging listening situations. The new outcome measures of real-life listening demands presented and discussed here can be particularly useful for better understanding the everyday challenges that listeners with hearing loss encounter and for optimizing HA signal processing targeting the SNR under everyday conditions.

ACKNOWLEDGMENTS

Work presented in this article was supported by a research grant from the Oticon Foundation; the EU H2020-ICT COCOHA (Cognitive Control of a Hearing Aid) grant agreement no: 644732; the Swedish Research Council (Vetenskapsrådet, VR 2017-06092 Mekanismer och behandling vid ålderrelaterad hörselnedsättning); and the Linnaeus Centre HEAD excellence center grant (349-2007-8654)

All authors were employees of Oticon at the time of manuscript preparation. Address for correspondence: Thomas Lunner, 16525 NE 92 PL Redmond, WA 98052, USA. E-mail: thomas.lunner@hotmail.com

Received May 15, 2020; accepted July 11, 2020.

REFERENCES

- Akram, S., Presacco, A., Simon, J. Z., Shamma, S. A., & Babadi, B. (2016). Robust decoding of selective auditory attention from MEG in a competing-speaker environment via state-space modeling. *Neuroimage*, *124*(Pt A), 906–917.
- Alexandrou, A. M., Saarinen, T., Jan, K., & Salmelin, R. (2018). Cortical entrainment: What we can learn from studying naturalistic speech perception. *Lang Cogn Neurosci*, *35*, 681–693.
- Alickovic, E., Lunner, T., & Gustafsson, F. (2016). A system identification approach to determining listening attention from EEG signals. In 2016 24th European Signal Processing Conference (EUSIPCO). Budapest, Hungary.
- Alickovic, E., Lunner, T., Gustafsson, F., & Ljung, L. (2019). A tutorial on auditory attention identification methods. *Front Neurosci*, *13*, 153.
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu Rev Neurosci*, *28*, 403–450.
- Atkinson, R. C., & Shiffrin, R. M. (1968). *Human Memory: A Proposed System and Its Control Processes*. Academic Press.
- Baddeley, A., & Hitch, G. (1974). Working memory. *Psychol Learn Motiv*, *8*, 47–89.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol Bull*, *91*, 276–292.
- Boldt, J. B., Kjems, U., Pedersen, M. S., Lunner, T., & Wang, D. (2008). Estimation of the ideal binary mask using directional systems. In *Proceedings of the 11th International Workshop on Acoustic Echo and Noise Control*. Seattle, Washington.
- Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annu Rev Psychol*, *40*, 109–131.
- Brodbeck, C., Hong, L. E., Simon, J. Z. (2018). Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Curr Biol*, *28*, 3976–3983.
- Brodbeck, C., Presacco, A., Anderson, S., Simon, J. Z. (2018). Overrepresentation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust United Acust*, *104*, 774–777.
- Brungart, D., Caduff, A., Campos, J., Carlile, S., Carpenter, M., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., Smeds K. (2020). The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it. *Ear Hear*, *41*(Suppl 1), 5S–19S.
- Carlile, S., & Keidser, G. (2020). Conversational interaction is the brain in action: Implications for the evaluation of hearing and hearing interventions. *Ear Hear*, *41*(Suppl 1), 56S–67S.
- Choi, I., Wang, L., Bharadwaj, H., & Shinn-Cunningham, B. (2014). Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hear Res*, *314*, 10–19.
- Chung, K. (2004). Challenges and recent developments in hearing aids. Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. *Trends Amplif*, *8*, 83–124.
- Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Front Hum Neurosci*, *10*, 604.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *J Verb Learn Verb Behav*, *19*, 450–466.
- Das, N., Bertrand, A., Francart, T. (2018). EEG-based auditory attention detection: boundary conditions for background noise and speaker positions. *J Neural Eng*, *15*, 066017.
- de Cheveigné, A., Wong, D. D. E., Di Liberto, G. M., Hjortkjær, J., Slaney, M., Lalor, E. (2018). Decoding the auditory brain with canonical component analysis. *Neuroimage*, *172*, 206–216.
- Decruy, L., Vanthornhout, J., Francart, T. (2019). Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *J Neurophysiol*, *122*, 601–615.
- Dillon, H. (2012). *Hearing Aids*. Hodder Arnold.
- Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc Natl Acad Sci U S A*, *109*, 11854–11859.
- Dmochowski, J. P., Ki, J. J., DeGuzman, P., Sajda, P., & Parra, L. C. (2018). Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity. *NeuroImage*, *180*, 134–146.
- Dryden, A., Allen, H. A., Henshaw, H., Heinrich, A. (2017). The Association Between Cognitive Performance and Speech-in-Noise Perception

- for Adult Listeners: A Systematic Literature Review and Meta-Analysis. *Trends Hear*, 21, 2331216517744675.
- Fiedler, L., Wöstmann, M., Graversen, C., Brandmeyer, A., Lunner, T., Obleser, J. (2017). Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. *J Neural Eng*, 14, 036020.
- Fuglsang, S. A., Dau, T., Hjortkjær, J. (2017). Noise-robust cortical tracking of attended speech in real-world acoustic scenes. *Neuroimage*, 156, 435–444.
- Gatehouse, S., & Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *Br J Audiol*, 24, 63–68.
- Granhölm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33, 457–461.
- Hällgren, M., Larsby, B., Lyxell, B., Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *J Am Acad Audiol*, 12, 357–370.
- Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: natural stimuli in speech neuroscience. *Lang Cogn Neurosci*, 35, 573–582.
- Hjortkjær, J., Märcher-Rørsted, J., Fuglsang, S. A., & Dau, T. (2018). Cortical oscillations and entrainment in speech processing during working memory load. *Eur J Neurosci*, 51, 1279–1289.
- Hohmann, V., Paluch, R., Krueger, M., Meis, M., & Grimm, G. (2020). The Virtual Reality Lab: Realization and application of virtual sound environments. *Ear Hear*, 41(Suppl 1), 31S–38S.
- Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A. (2015). The window of wemy eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biol Psychol*, 110, 100–106.
- Iotzov, I., & Parra, L. C. (2019). EEG can predict speech intelligibility. *J Neural Eng*, 16, 036008.
- Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: Evidence for the adaptive gain theory. *J Cogn Neurosci*, 23, 1587–1596.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.
- Keidser, G., Naylor, G., Brungart, D., Caduff, A., Campos, J., Carlile, S., Carpenter, M., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., Smeds, K. (2020). The quest for ecological validity in hearing science: What it is, why it matters, and how to advance it. *Ear Hear*, 41(Suppl 1), 5S–19S.
- Khalighinejad, B., Herrero, J. L., Mehta, A. D., Mesgarani, N. (2019). Adaptation of the human auditory cortex to changing background noise. *Nat Commun*, 10, 2509.
- Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., . . . others. (2003). Candidature for and delivery of audiological services: Special needs of older people. *Int J Audiol*, 42, 92–101.
- Koelwijn, T., Zekveld, A. A., Lunner, T., Kramer, S. E. (2018). The effect of reward on listening effort as reflected by the pupil dilation response. *Hear Res*, 367, 106–112.
- Koelwijn, T., Zekveld, A. A., Festen, J. M., Rönnerberg, J., Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *Int J Otolaryngol*, 2012, 865731.
- Koss, M. C. (1986). Pupillary dilation as an index of central nervous system alpha 2-adrenoceptor activation. *J Pharmacol Methods*, 15, 1–19.
- Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *Audiology*, 36, 155–164.
- Lalor, E. C., Power, A. J., Reilly, R. B., Foxe, J. J. (2009). Resolving precise temporal processing properties of the auditory system using continuous stimuli. *J Neurophysiol*, 102, 349–359.
- Lesenfans, D., Vanthornhout, J., Verschueren, E., Decruy, L., Francart, T. (2019). Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. *Hear Res*, 380, 1–9.
- Ljung, L. (1999). *System Identification: Theory for the User*. PTR Prentice Hall.
- Loewenfeld, I. E., & Lowenstein, O. (1993). *The Pupil: Anatomy, Physiology, and Clinical Applications*. Wiley-Blackwell.
- Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int J Audiol*, 42 Suppl 1, S49–S58.
- Lunner, T., Rudner, M., Rosenbom, T., Ågren, J., Ng, E. H. (2016). Using speech recall in hearing aid fitting and outcome evaluation under ecological test conditions. *Ear Hear*, 37, 145S–154S.
- Lyxell, B. (1994). Skilled speechreading: A single-case study. *Scand J Psychol*, 35, 212–219.
- Lyxell, B., Andersson, U., Borg, E., & Ohlsson, I.-S. (2003). Working-memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *Int J Audiol*, 42, S86–S89.
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., Amitay, S. (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.
- Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485, 233–236.
- Micula, A., Ng, E. H., El-Azm, F., & Rönnerberg, J. (2020). The effects of background noise, noise reduction and task difficulty on recall. *Int J Audiol*, 1–9.
- Mirkovic, B., Debener, S., Jaeger, M., De Vos, M. (2015). Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. *J Neural Eng*, 12, 046007.
- Murphy, P. R., O’Connell, R. G., O’Sullivan, M., Robertson, I. H., Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Hum Brain Mapp*, 35, 4140–4154.
- Naylor, G. (2016). Theoretical issues of validity in the measurement of aided speech reception threshold in noise for comparing nonlinear hearing aid systems. *J Am Acad Audiol*, 27, 504–514.
- Ng, E. H., Rudner, M., Lunner, T., Rönnerberg, J. (2015). Noise reduction improves memory for target language speech in competing native but not foreign language speech. *Ear Hear*, 36, 82–91.
- Ng, E. H., Rudner, M., Lunner, T., Pedersen, M. S., Rönnerberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int J Audiol*, 52, 433–441.
- Ng, E. N., Lunner, T., & Rönnerberg, J. (2016). *Real-Time Noise Reduction System Enhances Auditory Word Recall Performance in Noise*. International Hearing Aid Research Conference.
- Ng, E. H. N., & Rönnerberg, J. (2020). Hearing aid experience and background noise affect the robust relationship between working memory and speech recognition in noise. *Int J Audiol*, 59, 208–218.
- Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., Versfeld, N. J., Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hear Res*, 351, 68–79.
- Ohlenforst, B., Wendt, D., Kramer, S. E., Naylor, G., Zekveld, A. A., Lunner, T. (2018). Impact of SNR, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response. *Hear Res*, 365, 90–99.
- Ohlenforst, B., Zekveld, A. A., Jansma, E. P., Wang, Y., Naylor, G., Lorens, A., Lunner, T., Kramer, S. E. (2017a). Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: A Systematic Review. *Ear Hear*, 38, 267–281.
- O’Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cereb Cortex*, 25, 1697–1706.
- Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear Hear*, 39, 204–214.
- Peelle, J. E., & Wingfield, A. (2016). Listening effort in age-related hearing loss. *Hear J*, 69, 10–12.
- Petersen, E. B., Wöstmann, M., Obleser, J., Lunner, T. (2017). Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J Neurophysiol*, 117, 18–27.
- Pichora-Fuller, M. K. (2016). How social psychological factors may modulate auditory and cognitive functioning during listening. *Ear Hear*, 37, 92S–100S.
- Pichora-Fuller, M. K. (2016). How social psychological factors may modulate auditory and cognitive functioning during listening. *Ear Hear*, 37, 92S–100S.
- Presacco, A., Simon, J. Z., & Anderson, S. (2019). Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. *PLoS One*, 14, e0213899.
- Richter, M., Gendolla, G. H., & Wright, R. A. (2016). Three decades of research on motivational intensity theory: What we have learned about effort and what we still don’t know. *Adv Motiv Sci*, 3, 149–186.
- Rönnerberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int J Audiol*, 42, S68–S76.
- Rönnerberg, J., Arlinger, S., Lyxell, B., Kinnefors, C. (1989). Visual evoked potentials: Relation to adult speechreading and cognitive function. *J Speech Hear Res*, 32, 725–735.

- Rönnerberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front Syst Neurosci*, 7, 31.
- Rönnerberg, 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, S99–105.
- Smeds, K., Gotowiec, S., Wolters, F., Herrlin, P., Larsson, J., & Dahlquist, M. (2020). Selecting scenarios for hearing-related laboratory testing. *Ear Hear*, 41(Suppl 1), 20S–30S.
- Smeds, K., Wolters, F., Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *J Am Acad Audiol*, 26, 183–196.
- Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cogn Affect Behav Neurosci*, 17, 809–825.
- Vanhornhout, J., Decruy, L., Wouters, J., Simon, J. Z., Francart, T. (2018). Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. *J Assoc Res Otolaryngol*, 19, 181–191.
- Wendt, D., Dau, T., Hjortkjær, J. (2016). Impact of background noise and sentence complexity on processing demands during sentence comprehension. *Front Psychol*, 7, 345.
- Wendt, D., Hietkamp, R. K., & Lunner, T. (2017). Impact of noise and noise reduction on processing effort: A pupillometry study. *Ear and hearing*, 38, 690–700.
- Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hear Res*, 369, 67–78.
- WHO. (2001). *International Classification of Functioning, Disability and Health (ICF)*. Organisation mondiale de la santé, World Health Organization, and World Health Organization Staff.
- 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 Hear*, 22.
- Zekveld, A. A., Kramer, S. E., Festen, J. M. (2010). Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear Hear*, 31, 480–490.
- Zhao, S., Bury, G., Milne, A., Chait, M. (2019). Pupillometry as an objective measure of sustained attention in young and older listeners. *Trends Hear*, 23, 2331216519887815.

REFERENCE NOTE

1. Alickovic, E., Lunner, T., Wendt, D., Fiedler, L., Hietkamp, R., Ng, E. H., & Graversen, C. (in press). Noise Reduction Processing in Hearing Aids Enhances Selective Auditory Attention. *Frontiers in Neuroscience*.