

Towards Objective Measures of Functional Hearing Abilities

Hamish Innes-Brown, Renee Tsongas, Jeremy Marozeau and Colette McKay

Abstract *Aims* People with impaired hearing often have difficulties in hearing sounds in a noisy background. This problem is partially a result of the auditory systems reduced capacity to process temporal information in the sound signal. In this study we examined the relationships between perceptual sensitivity to temporal fine structure (TFS) cues, brainstem encoding of complex harmonic and amplitude modulated sounds, and the ability to understand speech in noise. Understanding these links will allow the development of an objective measure that could be used to detect changes in functional hearing before the onset of permanent threshold shifts.

Methods We measured TFS sensitivity and speech in noise performance (Quick-SIN) behaviourally in 34 normally hearing adults with ages ranging from 18 to 63 years. We recorded brainstem responses to complex harmonic sounds and a 4000 Hz carrier signal modulated at 110 Hz. We performed cross correlations between the stimulus waveforms and scalp-recorded brainstem responses to generate a simple measure of stimulus encoding accuracy, and correlated these measures with age, TFS sensitivity and speech-in-noise performance.

Results Speech-in-noise performance was positively correlated with TFS sensitivity, and negatively correlated with age. TFS sensitivity was also positively correlated with stimulus encoding accuracy for the complex harmonic stimulus, while

H. Innes-Brown (✉) · R. Tsongas · J. Marozeau · C. McKay
Bionics Institute, Melbourne, Australia
e-mail: hinnes-brown@bionicsinstitute.org

R. Tsongas
e-mail: r.tsongas@student.unimelb.edu.au

J. Marozeau
Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark
e-mail: jemarozeau@elektro.dtu.dk

C. McKay
e-mail: cmckay@bionicsinstitute.org

H. Innes-Brown
Research Group Experimental Oto-laryngology, KU Leuven, Leuven, Belgium

© The Author(s) 2016

P. van Dijk et al. (eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, Advances in Experimental Medicine and Biology 894, DOI 10.1007/978-3-319-25474-6_33

increasing age was associated with lower stimulus encoding accuracy for the modulated tone stimulus.

Conclusions The results show that even in a group of people with normal hearing, increasing age was associated with reduced speech understanding, reduced TFS sensitivity, and reduced stimulus encoding accuracy (for the modulated tone stimulus). People with good TFS sensitivity also generally had less faithful brainstem encoding of a complex harmonic tone.

Keywords Speech in noise · Objective measures · Temporal fine structure · Aging · Electrophysiology · Envelope-following response · Frequency-following response

1 Introduction

The benefits of amplification are greatest when hearing interventions are made as early as possible. There is therefore great interest in the clinical audiology community in the development of objective techniques to measure various hearing abilities that do not require behavioural responses from the patient and are able to determine fitting parameters for hearing aids and cochlear implants. While the use of cortical responses is in development (Billings et al. 2007, 2011; Carter et al. 2010; Billings et al. 2012; Chang et al. 2012), transient brainstem responses and steady state responses (Luts et al. 2004, 2006; Alaerts et al. 2010) are already used in clinical practice to objectively assess hearing thresholds in young children. However, while these measures provide an estimate of *audibility* with the prescribed hearing aid gain, they do not provide any indication of the expected *hearing ability* of the patient.

In order to develop an objective measure of hearing ability, it is necessary to establish the links between speech perception, psychophysical measures of perceptual sensitivity to the acoustic cues that underlie effective speech perception, and the proposed objective measures. This paper describes our initial investigations in this direction.

The overall aim of this study was to examine the relationships between perceptual sensitivity to temporal fine-structure cues, brainstem encoding of complex harmonic and amplitude-modulated sounds, and the ability to understand speech in noise. Understanding these links will allow the development of an objective measure that could be used to detect changes in functional hearing before the onset of permanent threshold shifts.

2 Methods

2.1 Participants

Thirty-four participants (14 men and 20 women) aged between 18–63 years took part in the experiment. All participants had normal hearing bilaterally, defined by four frequency average hearing loss thresholds (500 Hz, 1 kHz, 2 kHz and 4 kHz) of less than 25 dB HL. Pure tone hearing thresholds for the 34 participants are shown in Fig. 1. Thresholds were measured using an Otometrics MADSEN Itera II audiometer with TDH-39 Stereo Headphones. The study was approved by Royal Victorian Eye & Ear Hospital Human Research Ethics Committee. Consent was written and informed.

2.2 Temporal Fine Structure Sensitivity

The TFS1 test (Moore and Sek 2009, 2012) was used to measure participants' temporal fine structure (TFS) sensitivity. The test was carried out using the TFS1 test software (<http://hearing.psychol.cam.ac.uk/>) and was based on the standard TFS1 test protocol. The task was performed on the participant's better ear as determined by the audiogram.

One practice run was given prior to the test. If participants could perform the task and attain a threshold they were given three real runs. For some participants the staircase procedure saturated (reached the 'easy' limit). Instead of receiving a threshold, these participants received a percent-correct score obtained from 40 trials of the 2AFC task at the easiest level of the staircase. Both the threshold and the

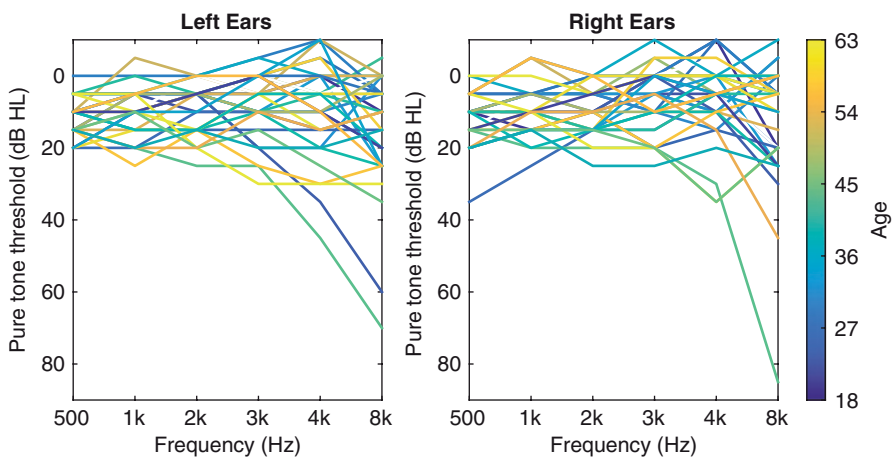


Fig. 1 Pure tone hearing thresholds for each ear

percent-correct scores were converted to a d' sensitivity measure using the method outlined by Hopkins and Moore (2007).

2.3 *Speech in Noise Tests (QuickSIN)*

Speech-in-noise performance was assessed behaviourally using the QuickSIN test (Killion et al. 2004). Six sets of six sentences with five key words per sentence were presented to both ears in four-talker babble noise. The sentences were presented with decreasing signal-to-noise ratios (SNRs) from +25 dB SNR to 0 dB SNR in 5 dB steps (+25, +20, +15, +10, +5 and 0 dB). The average SNR loss was calculated across the six lists. This score indicates the increase in SNR required for the participant to understand the sentence compared with a normal hearing person. Any score less than 4 dB is considered normal, and a lower SNR loss score reflects better speech-in-noise performance.

2.4 *Electrophysiology*

2.4.1 *Stimuli*

Envelope-following responses (EFRs) from the brainstem were elicited in response to two stimulus types—a complex harmonic tone, and a 4 KHz sinusoidal carrier tone modulated at 110 Hz. Both sounds were 100 ms in duration with 5 ms linear onset and offset ramps. The complex harmonic tone had an F0 of 180 Hz and 20 harmonics of equal amplitude and random phase. The modulated tone had a modulation depth of 50%. Each stimulus was presented with alternating polarities. The first 20 ms of each stimulus are shown in Fig. 2.

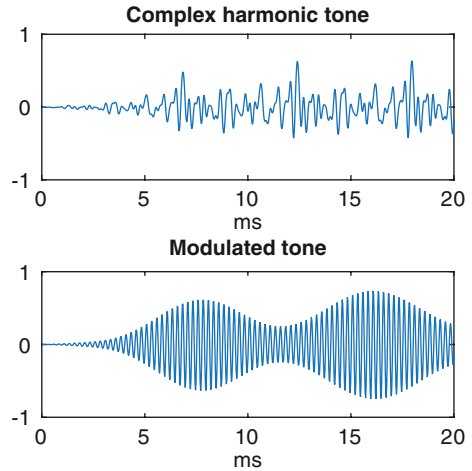
Stimuli were controlled via custom software in MAX/MSP (Cycling '74), played through an RME Fireface 400 audio interface and Etymotic ER3-A insert-phones.

2.4.2 *EEG Recordings and Pre-processing*

EEG data were recorded from the scalp via Ag-AgCl electrodes, using a BioSemi ActiveTwo EEG System. Electrode offsets were ± 40 mV. The EEG data were collected in continuous mode at a sampling rate of 16.384 kHz. The -3 dB point of the antialiasing lowpass filter was 3276 Hz.

The EEG recordings were segmented into epochs of -50 to 150 ms, separately for each stimulus type. The epochs were artefact rejected using the `pop_autorej` function from the EEGLAB toolbox (Delorme and Makeig 2004) using MATLAB software. EFRs were computed by adding responses to the positive and negative stimulus polarities ($\text{EFR} = (\text{Pos} + \text{Neg})/2$). All subsequent correlational analyses were conducted on the averaged EFR waveforms.

Fig. 2 The first 20 ms of the complex harmonic tone (*top*) and modulated tone (*bottom*)



The signal-to-noise ratio (SNR) of the EFR responses was calculated as $20\log_{10}(\text{RMS}_{\text{post}}/\text{RMS}_{\text{pre}})$, where RMS_{post} and RMS_{pre} were the root mean square of the amplitude in the ‘response’ period (defined as 25–75 ms post-stimulus) and the root mean square of the pre-stimulus period (defined as 50 ms pre-stimulus until stimulus onset at 0 ms) respectively. Three participants were removed from the EEG analysis as they had SNRs < 1.5 dB due to movement and other muscle artefact.

Stimulus-to-response cross-correlations (r -values) were generated using the Brainstem Toolbox 2013 (Skoe and Kraus 2010). The maximum cross-correlation values were chosen irrespective of the lag. Cross-correlations were performed against the Hilbert envelope of the stimulus, as recorded through the transducer and an artificial ear (GRAS Type 43-AG). All data transformation and statistical tests (correlational analyses) were conducted using Matlab and Minitab®. Spearman’s rank correlations were performed where data were skewed.

3 Results

3.1 Hearing Thresholds and Age

Hearing sensitivity generally declined slightly with age. The correlation between age and the pure-tone average hearing loss (at 500 Hz, 1 kHz, 2 kHz and 4 kHz) for the best ear was significant, $r=0.43$, $p=0.001$.

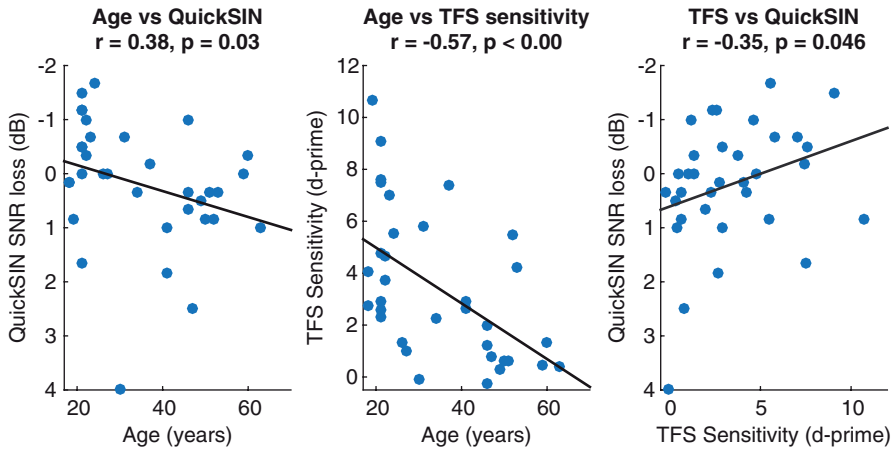


Fig. 3 *Left panel:* Speech in noise performance (QuickSIN SNR loss) as a function of age (*left*). Note that the direction of the y-axis has been reversed so that better performance is up. *Middle panel:* Sensitivity to TFS as a function of age. *Right panel:* Speech in noise performance (QuickSIN) as a function of TFS sensitivity. TFS sensitivity is expressed using a d' measure. Note that the direction of the y-axis has been reversed so that better performance is up. The *grey line* in all three panels indicates a least-squares linear regression

3.2 Correlations: QuickSIN, TFS Sensitivity, Age and Speech in Noise

Figure 3 shows the relationships between age, speech in noise perception, and TFS sensitivity. Both speech in noise scores ($r = -0.38, p < 0.03$) and TFS sensitivity ($r = -0.57, p < 0.001$) were significantly negatively correlated with age. Older participants generally had worse speech scores, and worse TFS sensitivity, with a stronger relationship in the case of TFS sensitivity. Speech in noise performance was moderately and significantly related to TFS sensitivity ($r = -0.34, p = 0.046$). Participants who had good TFS sensitivity generally had good speech in noise scores (good scores are negative).

3.3 Electrophysiology

Grand average responses for the complex harmonic and modulated tones are shown in Fig. 4. The responses show clear phase-locking to periodicity in the stimuli.

A measure of stimulus encoding strength was generated by calculating the cross-correlation between the stimulus envelope (as measured through the transducer and artificial ear) and the brainstem response. Figure 5 shows this process for one listener.

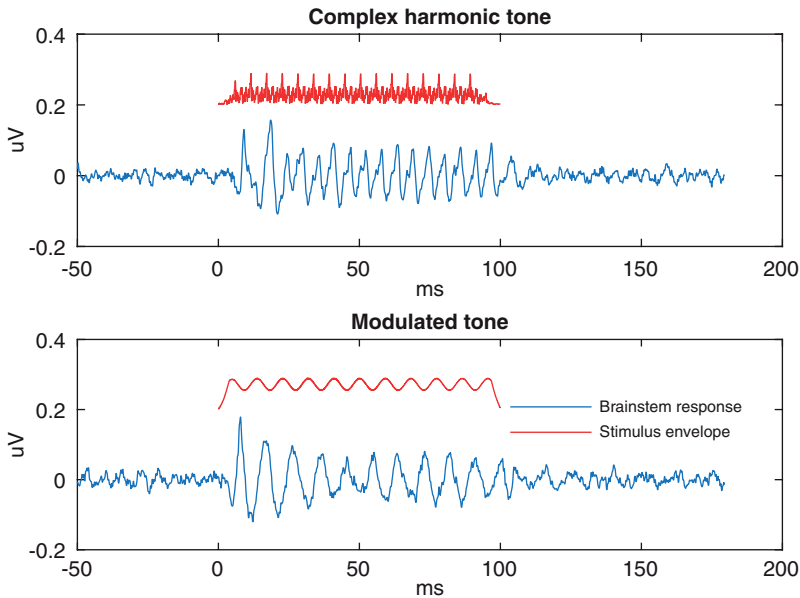


Fig. 4 Grand average (across all participants) brainstem responses for the complex harmonic tone (*top panel*) and modulated tone (*bottom panel*). The Hilbert envelope of the stimulus waveform is shown in *red*

3.4 Correlations: Stimulus Encoding Accuracy with Age, TFS Sensitivity and QuickSIN

The maximum cross-correlation value obtained from each participant and stimulus type was correlated with age, the TFS sensitivity and speech in noise scores. The top row of Fig. 6 shows that increasing age was associated with decreasing stimulus encoding accuracy, but only for the modulated tone stimulus ($r=-0.53$, $p<0.001$). The middle row of Fig. 6 also shows a striking relationship between increased TFS sensitivity and reduced stimulus encoding for the complex harmonic tone ($r=-0.56$, $p=0.004$). Interestingly, there was no such relationship for the modulated tone. There were also no significant relationships between stimulus encoding accuracy and the speech in noise scores (bottom row).

3.5 Regression Analysis

In order to determine which psychophysical and/or EEG measure best predicted the speech scores, best subsets regression was performed. The QuickSIN scores were entered as the response variable. Age, four-frequency pure-tone hearing thresholds (in the better ear), TFS sensitivity, and stimulus encoding accuracy for both the

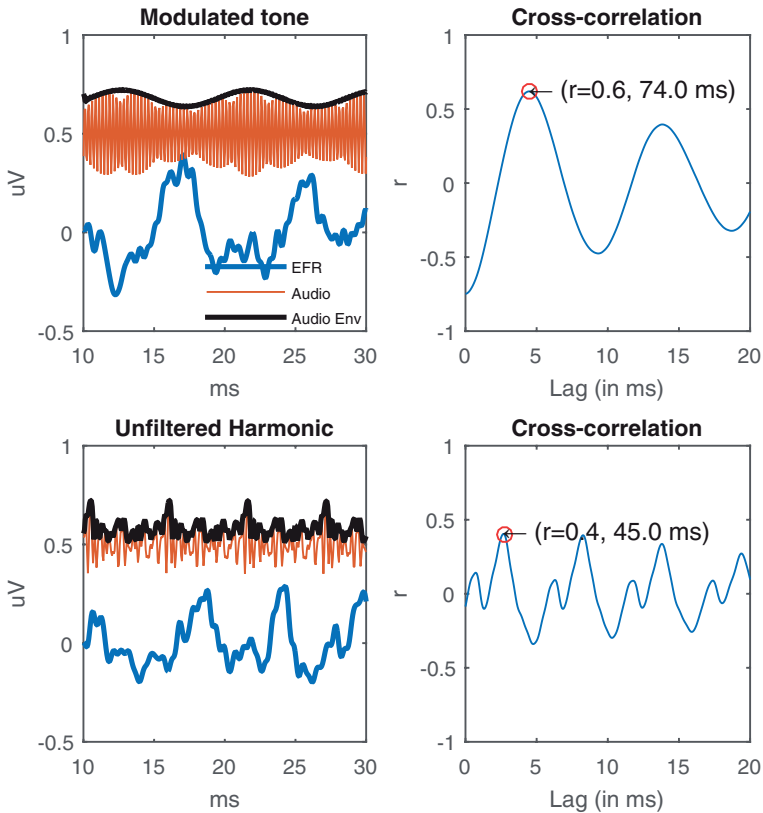


Fig. 5 An illustration for a single listener of the calculation of the cross-correlation values for the modulated tone (*top panels*) and complex harmonic tone (*bottom panels*). The *left panels* show the brainstem responses (in *blue*) and the stimulus waveform (*red*) and its Hilbert envelope (*black*). Note the stimulus waveform here is not the electrical signal: rather it has been recorded through the transducer and an artificial ear. The *right panels* show the cross-correlation values between the stimulus envelope and the brainstem response as a function of the lag between them. The maximum of the cross-correlation function was determined for further analysis

complex harmonic and modulated tones were entered as predictors. The regression indicated that QuickSIN scores were best predicted by a combination of all variables except the hearing thresholds ($R^2=64.4$, $F_{(4,26)}=6.32$, $p=0.004$). A follow-up standard linear regression using the four predictors identified by the best-subsets procedure echoed these results, although TFS sensitivity was not a significant factor. The main model was significant ($R^2=41.8$, $F_{(4,26)}=4.49$, $p=0.007$), with age accounting for most of the variance in the model (18.1%), followed by the stimulus encoding accuracy variables contributing 13.7 and 10.0% for the modulated and complex harmonic tones respectively. When the pure-tone average hearing thresholds were added to the model, they contributed only 2% variance.

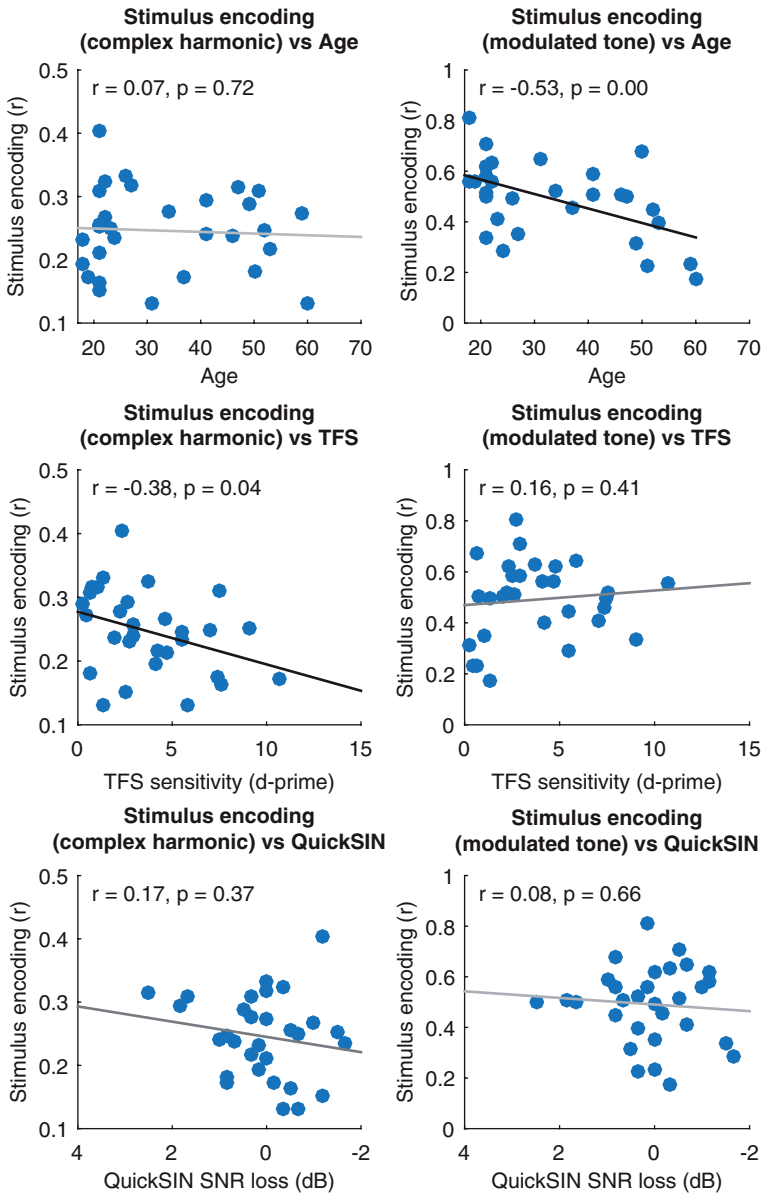


Fig. 6 Stimulus-to-response cross-correlation values between the complex harmonic (*left*) and modulated (*right*) stimuli and brainstem response as a function of age (*top row*), TFS sensitivity (*middle row*) and speech in noise scores (*bottom row*)

4 Discussion

In summary, we found that in our group of normally-hearing adults (with a variety of ages), speech-in-noise performance was negatively correlated with TFS sensitivity and age. TFS sensitivity was also positively correlated with stimulus encoding accuracy for the complex harmonic stimulus, while increasing age was associated with lower stimulus encoding accuracy for the modulated tone stimulus. Surprisingly, we found that better speech in noise understanding was associated with worse stimulus encoding accuracy. Despite this unexpected direction of correlation the measures did contribute modestly to a regression model predicting the speech in noise scores. A regression analysis found that age and the combination of the two stimulus encoding accuracy measures had roughly equal contributions to the model.

Further work in this area should consider other psychophysical predictors that are known to be associated with speech understanding, such as measures of temporal modulation sensitivity, and EEG measures that more closely match the stimuli used in the psychophysics.

Acknowledgments Supported provided by the National Health and Medical Research Council of Australia (Peter Doherty Early Career Research Fellowship #1069999 to HIB). The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program. HIB is currently supported by a FWO Pegasus Marie Curie Fellowship at KU Leuven, Belgium.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

- Alaerts J, Luts H, Van Dun B, Desloovere C, Wouters J (2010) Latencies of auditory steady-state responses recorded in early infancy. *Audiol Neuro-Otol* 15(2):116–127
- Billings CJ, Papesh MA, Penman TM, Baltzell LS, Gallun FJ (2012) Clinical use of aided cortical auditory evoked potentials as a measure of physiological detection or physiological discrimination. *Int J Otolaryngol* 14
- Billings CJ, Tremblay KL, Souza PE, Binns MA (2007) Effects of hearing aid amplification and stimulus intensity on cortical auditory evoked potentials. *Audiol Neuro-Otol* 12(4):234–246. doi:10.1159/000101331
- Billings CJ, Tremblay KL, Miller CW (2011) Aided cortical auditory evoked potentials in response to changes in hearing aid gain. *Int J Audiol* 50(7):459–467. doi:10.3109/14992027.2011.568011
- Carter L, Golding M, Dillon H, Seymour J (2010) The detection of infant cortical auditory evoked potentials (CAEPs) using statistical and visual detection techniques. *J Am Acad Audiol* 21(5):347–356. doi:10.3766/jaaa.21.5.6

- Chang H-W, Dillon H, Carter L, Van Dun B, Young S-T (2012) The relationship between cortical auditory evoked potential (CAEP) detection and estimated audibility in infants with sensorineural hearing loss. *Int J Audiol* 51:663–670
- 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(1):9–21. [10.1016/j.jneumeth.2003.10.009]
- Hopkins K, Moore BC (2007) Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J Acoust Soc Am* 122(2):1055–1068. doi:10.1121/1.2749457
- Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116(4 Pt 1), 2395–2405
- Luts H, Desloovere C, Kumar A, Vandermeersch E, Wouters J (2004) Objective assessment of frequency-specific hearing thresholds in babies. *Int J Pediatr Otorhinolaryngol* 68(7):915–926
- Luts H, Desloovere C, Wouters J (2006) Clinical application of dichotic multiple-stimulus auditory steady-state responses in high-risk newborns and young children. *Audiol Neuro-Otol* 11(1):24–37
- Moore BCJ, Sek A (2009) Development of a fast method for determining sensitivity to temporal fine structure. *Int J Audiol* 48:161–171
- Moore BCJ, Sek A (2012) Implementation of two tests for measuring sensitivity to temporal fine structure. *Int J Audiol* 51:58–63
- Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31(3):302–324