Brainstem correlates of cochlear nonlinearity measured via the scalp-recorded frequency-following response

Gavin M. Bidelman[a,b,c] and Shaum Bhagat[d,e]

The frequency-following response (FFR) is an EEG-based potential used to characterize the brainstem encoding of complex sounds. Adopting techniques from auditory signal processing, we assessed the degree to which FFRs encode important properties of cochlear processing (e.g. nonlinearities) and their relation to speech-in-noise (SIN) listening skills. Based on the premise that normal cochlear transduction is characterized by rectification and compression, we reasoned these nonlinearities would create measurable harmonic distortion in FFRs in response to even pure tone input. We recorded FFRs to nonspeech (pure- and amplitude-modulated-tones) stimuli in normal-hearing individuals. We then compared conventional indices of cochlear nonlinearity, via distortion product otoacoustic emission (DPOAE) I/O functions, to total harmonic distortion measured from neural FFRs (FFR_{THD}). Analysis of DPOAE growth and the FFR_{THD} revealed listeners with higher cochlear compression thresholds had lower neural FFR_{THD} distortion (i.e. more linear FFRs), thus linking cochlear and brainstem correlates of auditory nonlinearity. Importantly, FFR_{THD} was also negatively correlated with SIN perception whereby listeners with higher FFR_{THD} (i.e. more nonlinear responses) showed better performance on the QuickSIN. We infer individual differences in SIN perception and FFR nonlinearity even in normal-hearing individuals may reflect subtle differences in auditory health and suprathreshold hearing skills not captured by normal audiometric evaluation. Future studies in hearing-impaired individuals and animal models are necessary to confirm the diagnostic utility of FFR_{THD} and its relation to cochlear hearing loss or peripheral neurodegeneration in humans.

NeuroReport 2020, **31**:702–707

Keywords: auditory brainstem response, cochlear nonlinearity, distortion product otoacoustic emission, hidden hearing loss, synaptopathy

[a]Institute for Intelligent Systems, University of Memphis, [b]School of Communication Sciences & Disorders, University of Memphis, [c]Department of Anatomy and Neurobiology, University of Tennessee Health Sciences Center, Memphis, Tennessee [d]Department of Communicative Disorders and Sciences and [e]Department of Audiology, San Jose State University, San Jose, California, USA

Correspondence to Gavin M. Bidelman, PhD, School of Communication Sciences & Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA
Tel: +901 678 5826; fax: +901 525 1282; e-mail: gmbdlman@memphis.edu

Received 19 November 2019 **Accepted** 18 March 2020

Introduction

The cochlea shows several nonlinearities including a compressive input-output function [1,2] and harmonic distortion that are robust indicators of cochlear health because they diminish with sensorineural hearing loss (SNHL) [3] as basilar membrane responses are linearized [3]. Monitoring these mechanisms might aid detection of subtle changes in auditory function with nascent (preclinical) hearing loss. In this vein, cochlear nonlinearities have been estimated physiologically via distortion product otoacoustic emissions (DPOAEs) [4,5]. However, DPOAEs are difficult to measure in ears with more than mild losses [4] and are weak, back-propagated acoustic signals that are susceptible to environmental noise [6]. DPOAEs are also 'pre-neural' and thus may miss subtle auditory deficits and potential 'hidden' losses related to early neuropathy [7].

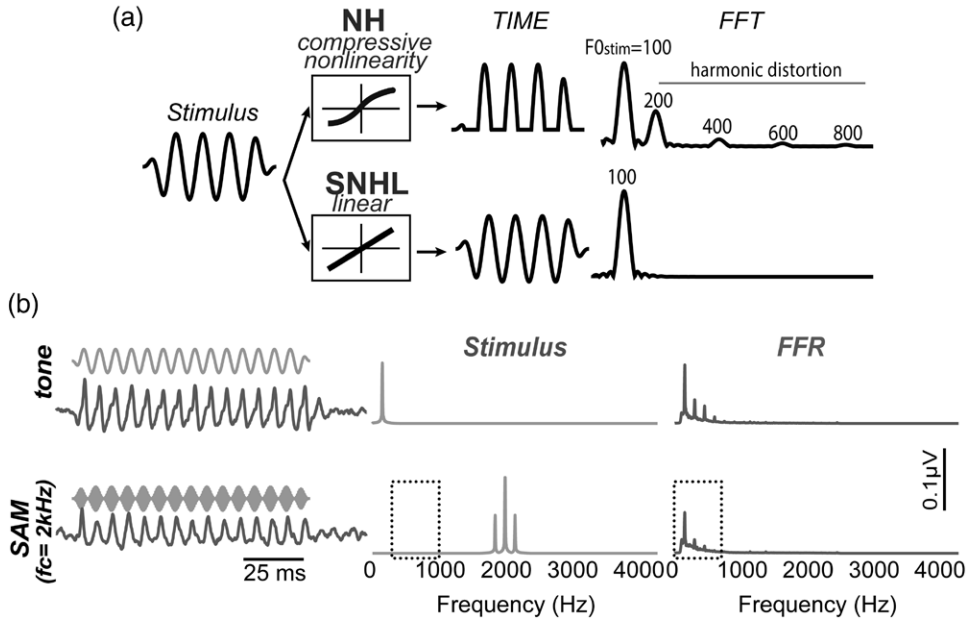
Theoretically, the cochlea's compressive nonlinearity [1] coupled with half-wave rectification of inner hair cell transduction [2,8,9] will distort a pure-tone input, generating harmonic components in the neural output of the cochlea at multiples of the signal frequency (e.g. *2f*, *3f*,…,*nf*) (Fig. 1a). Harmonic distortion does not exist in the stimulus and must therefore reflect physiological processing. Germane to our study, cochlear distortion products are forward propagated to the VIIIth nerve and are detectable in midbrain inferior colliculus neurons [10,11]. In humans, such distortion products have been documented in the scalp-recorded frequency-following response (FFR) [6,10,11], a sustained potential that reflects phase-locked ensemble activity dominantly from the midbrain [12].

Here, we evaluated (1) whether FFRs carry a neural manifestation of cochlear nonlinearities (e.g. framework in Fig. 1a) and (2) whether such nonlinearities relate to complex listening abilities. Recent studies suggest subtle neurodegeneration in the cochlea might account for difficulties in speech-in-noise (SIN) understanding and other hearing problems in people with clinically normal hearing (i.e. 'hidden hearing loss') [7,13,14]. We posited that reduced distortion broadcast into FFRs might reflect a neural proxy of nonlinear cochlear processing that is also

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.neuroreport.com.

DOI: 10.1097/WNR.0000000000001452

Fig. 1



Cochlear distortion measured in FFRs to complex sounds. (a) Theoretical framework for detecting changes in cochlear (nonlinear) processing via neural FFR scalp potentials. Toy examples of the signal chain for cochleae with normal hearing (NH) and sensorineural hearing loss (SNHL). NH is characterized by a compressive nonlinearity (I/O response) and rectification produced by haircell physiology [2,8,9]. In response to a pure tone input ($F0_{stim}$), this yields harmonic distortion in the FFR spectrum. In cases of SNHL (or preclinical declines in cochlear processing), the system is linearized and only a response at the stimulus $F0_{stim}$ is present. (b) Neural FFR and acoustic stimulus time waveforms and spectra. Each stimulus has a common F0 = 150 Hz. Higher amplitude harmonics in the region of the F0 (dotted box) indicate distortion produced by nonlinear cochlear processing that is propagated from the cochlea to the midbrain source of the FFR. To a pure tone, the FFR spectrum is complex and contains harmonics of F0. Similarly, the SAM tone contains no energy in its spectrum at 150 Hz but the F0 component is introduced into the FFR from phase-locking to the cochlear-generated F0. FFR, frequency-following response; SAM, sinusoidally amplitude modulated.

sensitive to variations in these suprathreshold listening skills.

Materials and methods

Participants

Fourteen young adults (mean age: 23.6 ± 1.0 years) participated in the study. All had normal hearing thresholds (i.e. <20 dB HL; 250–8000 Hz) and middle ear function (i.e. type-A tympanograms, acoustic reflex thresholds). Thresholds were within ±2 dB between ears (puretone average: left = 5 ± 4.9 dB HL; right = 7.2 ± 5.0 dB HL). Each listener was right-handed and college-educated (18.0 ± 1.4 years). Subjects had minimal amounts of music training (mean: 3.5±3.8 years). All gave informed consent in compliance with a protocol approved by the University of Memphis IRB.

Stimuli

FFRs were recorded in response to nonspeech stimuli with a common low pitch [i.e. fundamental frequency (F0) = 150 Hz]: (1) 150 Hz pure tone and (2) sinusoidally amplitude modulated (SAM) tone with a carrier frequency (f_c) of 2000 Hz and modulation frequency (f_m) of 150 Hz (100% depth) (Fig. 1b). The high F0 ensured FFRs were of brainstem (rather than cortical) origin [12]. Stimuli were matched in duration (100 ms) and root mean squared level.

Frequency-following response recordings

FFRs were recorded (passive listening) using a vertical montage [Fpz (+) –linked A1/A2 (–); mid-forehead electrode = ground]. Impedances were <2 kΩ. Stimuli (2000 trials/token) were delivered binaurally using rarefaction polarity at 80 dB sound pressure level (SPL) (interstimulus interval = 50 ms) through ER-30 earphones (Etymotic Research, Elk Grove Village, Illinois). Extended acoustic tubing (20 ft) avoided stimulus artifact from contaminating FFRs. Neural activity was digitized (10 kHz) (SynAmps RT; Compumedics Neuroscan, Charlotte, NC), filtered (80–2500 Hz), epoched (0–150 ms), and averaged to derive FFRs for each condition. Sweeps >±50 μV were rejected from averaging.

We quantified the degree of nonlinearity in brainstem FFRs (FFR_{THD}) via total harmonic distortion [15] (equation 1):

$$\mathrm{FFR}_{\mathrm{THD}} = 10log\left(\frac{\sqrt{H_2^2 + H_3^2 + H_4^2 + \ldots + H_8^2}}{H_1}\right) \quad (1)$$

where H1–H8 are the first eight harmonic amplitudes of the F0(=H1) extracted from the FFR's spectrum. We considered ≤H8 in the $\mathrm{FFR}_{\mathrm{THD}}$ calculation given phase-locking limits of midbrain neurons (<1200 Hz) and the upper limit of recordable FFR in humans [16]. When expressed in decibels relative to the F0 carrier amplitude (dBc), $\mathrm{FFR}_{\mathrm{THD}}$ represents the amplitude of the distortional components relative to the input stimulus F0 frequency. Intuitively, it reflects the degree of nonlinear response present in the FFR relative to the response to the stimulus component alone. Higher (more positive) numbers reflect stronger harmonic distortion components in the neural response which we explore here as a proxy of stronger cochlear nonlinearity. Importantly, $\mathrm{FFR}_{\mathrm{THD}}$ is a relative measure, which circumvents problems with using absolute measures in detecting suprathreshold auditory deficits via scalp potentials [17], which are too susceptible to non-biological factors [13].

We measured FFR onset latency at the peak negative deflection between 8 and 12 ms, the expected latency of brainstem FFRs [12,16]. FFR neural timing delays have been suggested as a biomarker for SIN perception [18]. All latencies were corrected for the acoustic delay of the headphone (17.8 ms) [16].

Distortion product otoacoustic emissions

DPOAE testing was similar to our previous studies [19,20]. Emissions were evoked by the simultaneous presentation of two primary tones (f_2 & f_1 with $f_2 > f_1$; $f_2/f_1 = 1.2$; 512 ms duration) and were recorded using ER-10C probe (Etymotic Research) assembly seated in the ear canal controlled by a Mimosa Acoustics Hear ID system. DPOAEs were measured at the cubic difference frequency (i.e. $2f_1 - f_2$). f_2 was fixed at 2016 Hz to match the SAM tone *fc*, and hence cochlear place, as FFR recordings. To map DPOAE I/O functions [5,20], we varied the primary levels (L_1 and L_2) between 25 and 65 dB SPL using the 'scissor' paradigm [4] where $L1 = 0.4L2 + 39$. This yields DPOAE I/O functions that closely mimic basilar membrane compression [4]. DPOAE averaging was terminated using the Mimosa's stopping rules (for details see SI Methods, Supplemental digital content 1, *http://links.lww.com/WNR/A587*). DPOAEs were measured in both ears and collapsed for analysis. Two subjects' DPOAE data were lost due to technical error resulting in n = 12 DPOAE recordings.

We measured the compression threshold (CT) from each listener's DPOAE I/O function [4,21]. Curves were first fit with a two-segment 'broken stick' function [3] using a piece-wise regression approach previously used to characterize behavioral and physiological I/O functions [3]. The breakpoint (i.e. L_2 level where the two segments join) was taken as the CT, which roughly corresponded to the –6 dB down point of the I/O function (Fig. 3a). CTs reflect the level where DPOAEs begin growing compressively [21] and have been related to subtle variations in hearing even among normal-hearing listeners [5,20].

Behavioral speech-in-noise test

The QuickSIN (Etymotic Research) consists of 6 sentences (70 dB SPL) embedded in four-talker babble noise. SNR is decreased in 5 dB steps from 25 dB (easy) to 0 dB (difficult). The number of target words (five per sentence) correctly recalled was used to calculate SNR-loss, reflecting 50% performance. Higher scores indicate poorer SIN performance. Two lists were administered binaurally via Sennheiser HD 280 headphones and were averaged to obtain a single SNR-loss per listener.

Results

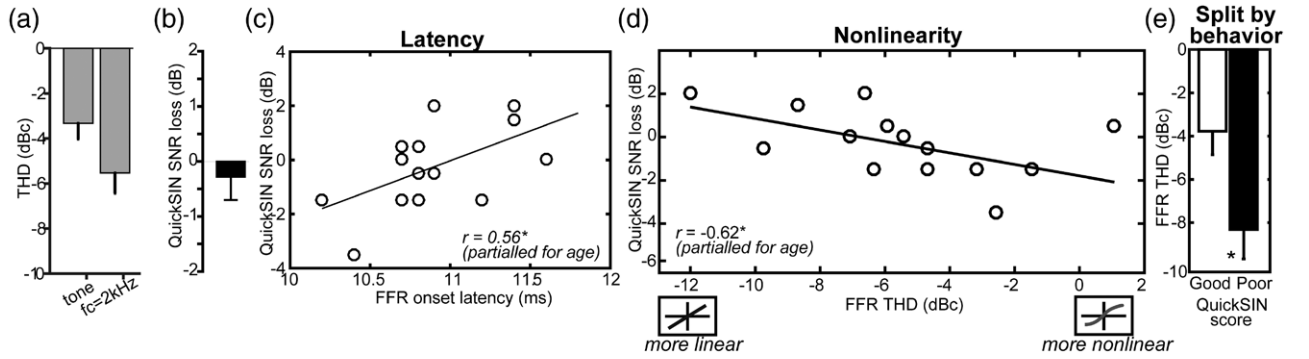
FFRs contained harmonically-related frequency components of the stimulus F0 (150 Hz) (Fig. 1b). This was true even for a pure tone input and presumably reflects nonlinearity stemming from cochlear haircell rectification [2,9]. Similarly, SAM tone stimuli contained no energy at the low F0, but nonlinear processing introduced this component into FFRs through phase-locked neural ensemble activity at the F0 [9,22]. SAM and pure tone stimuli elicited similar $\mathrm{FFR}_{\mathrm{THD}}$, indicating a similar degree of harmonic distortion in these two conditions ($t_{13} = 1.66$, $P = 0.12$, Cohen's $d = 0.03$) (Fig. 2a).

QuickSIN scores were <2 dB SNR loss, as expected in normal-hearing listeners (Fig. 2b). Brain-behavior correlations revealed associations between FFR response properties to SAM tones and QuickSIN scores (Fig. 2c–d). After accounting for age, we found more prolonged FFR latencies predicted poorer (i.e. larger) QuickSIN scores (Pearson's $r = 0.56$, $P = 0.046$]. Interestingly, listeners with more nonlinear FFRs (i.e. higher $\mathrm{FFR}_{\mathrm{THD}}$) showed better QuickSIN [$r = -0.62$, $P = 0.024$]. This implies stronger nonlinear cochlear processing (as reflected in higher $\mathrm{FFR}_{\mathrm{THD}}$) is related to superior SIN perception. No correlations were observed for the pure tone stimulus.

To further investigate the behavioral relevance of $\mathrm{FFR}_{\mathrm{THD}}$ to SIN perceptual skills, we divided our cohort in two groups based on their QuickSIN scores. We then assessed $\mathrm{FFR}_{\mathrm{THD}}$ in the top vs. bottom quartile of our sample as a function of SIN performance. This analysis revealed better SIN perceivers had more nonlinear FFRs (i.e. larger $\mathrm{FFR}_{\mathrm{THD}}$) compared to poorer SIN perceivers who had more linearized neural responses ($t_6 = 2.61$, $P = 0.0401$, $d = 1.85$; Fig. 2e).

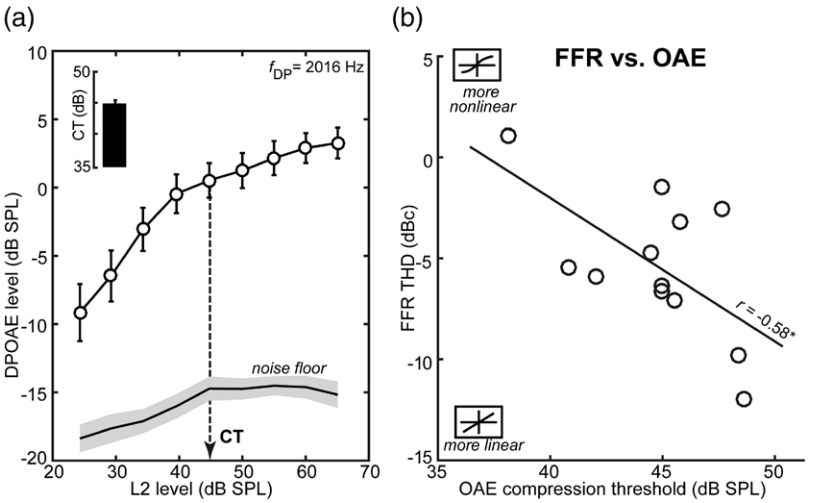
To validate our interpretation that $\mathrm{FFR}_{\mathrm{THD}}$ represents a reflection of the nonlinear cochlear output, we compared our neural metric against a gold-standard measure of cochlear nonlinearity: the DPOAE I/O function

Fig. 2



Relationship between neural measures of cochlear nonlinearity (FFR_{THD}) and behavioral SIN perception. (a) FFR_{THD}. Stronger harmonic distortion in FFRs (i.e. THDc = 0) indicates more robust cochlear nonlinearity. (b) QuickSIN scores. (c) FFR latency and (d) THD for the 2 kHz SAM responses were significantly correlated with QuickSIN scores; stronger FFR nonlinearity is associated with better SIN perception. (e) Split of the cohort (top vs. bottom quartile) based on their QuickSIN. Better SIN perceivers have more nonlinear FFRs (i.e. higher THD) whereas poorer SIN perceivers have more linearized neural responses. $^*P < 0.05$; errorbars = ±1 SEM. FFR, frequency-following response; SAM, sinusoidally amplitude modulated; SIN, speech-in-noise.

Fig. 3



Relationship between FFR_{THD} and DOAE-based index of cochlear nonlinearity. (a) Mean DPOAE I/O function at the 2016 Hz cochlear region. Compression threshold (CT) is indicated by the dashed line. (b) Scatterplot of FFR_{THD} vs. DOAE CTs across ears. Stronger neural nonlinearity (i.e. more positive FFR_{THD}) is associated with lower DOAE compression thresholds. $^*P < 0.05$; error bars/shading = ±1 SEM. DPOAE, distortion product otoacoustic emission; DOAE, distortion product otoacoustic emission; FFR, frequency-following response.

(Fig. 3). I/O functions revealed typical DPOAE growth with more linear responses at low levels and a compressive nonlinearity at mid-to-high levels [4,21]. On average, DPOAE-derived CTs were between 40 and 50 dB SPL, consistent with estimates from previous DPOAE studies [20]. Interestingly, comparisons between FFR and DPOAE measures revealed neural distortion (FFR_{THD}) was negatively associated with cochlear CTs [r = –0.58, P = 0.0363] (Fig. 3b): larger peripheral CTs were related to lower neural distortion in FFRs (i.e. more linear responses).

Discussion

By measuring brainstem FFRs in normal-hearing listeners, our data reveal that salient biomarkers of normal cochlear function (i.e. harmonic distortion) are robustly represented in scalp-recorded EEG potentials. Moreover, our findings suggest the degree of nonlinearity and neural timing in listeners' FFR is directly related to their cochlear nonlinearity (i.e. DPOAE CTs) and suprathreshold listening skills relevant to SIN perception.

Our findings corroborate previous FFR reports [9,22] by demonstrating biological distortion produced by the

auditory system is readily captured in scalp-recorded neural potentials. Distortion components in FFRs are thought to reflect a forward propagated version of those generated by basilar membrane mechanics [6,10,11,23], whereby cochlear haircell transduction shows a bias in the preferential phase of excitation that effectively half-wave rectifies the input signal [2,8,9]. Together with the compressive transfer function of the outer hair cell-driven cochlear amplifier [1], these nonlinearities yield a complex signal at the cochlear output when driven by pure tone input (Fig. 1). EEG-based FFRs dominantly originate from upper brainstem nuclei with other major contributions from auditory nerve [12]. Thus, harmonic distortion observed in neural FFRs most probably reflects cochlear generated nonlinearities that are propagated upstream to the FFR's site(s) of generation.

It is tempting to suggest our FFR_{THD} metric might offer an objective assay of cochlear health and the pre-clinical hearing loss [7,17,24]. While individual differences in SIN perception and FFR nonlinearity observed here even in normal-hearing individuals may reflect early neural degeneration/synaptopathy [7], we caution this interpretation. Attempts to establish a viable electrophysiological correlate of cochlear synaptopathy have experienced mixed results stemming from poor test-retest reliability and construct validity of the numerous diagnostic indices proposed in the literature [for review, see Ref. 17]. Hidden hearing loss remains highly controversial [24] and there is not, at present, strong evidence that it exists in humans, let alone relates to SIN processing [14,24,25]. High-powered human studies fail to find a relation between clinical auditory brainstem response measures and SIN perception [25]. Alternatively (and more likely), observed variations in SIN perception and cochlear/neural nonlinearity may simply reflect intrinsic variations in auditory system function (see SI Discussion).

Still, new objective measures of hearing function that circumvent previous shortcomings are needed to advance early detection and characterization of hearing deficits. Signal nonlinearities (compression, harmonic distortion) are robust indicators of cochlear health because they diminish with SNHL as basilar membrane responses are linearized with hearing impairment [3]. FFR distortion products are also much stronger than their weaker DPOAE version; the latter is attenuated in the ear canal via its roundtrip acoustic pathway whereas the former is forward propagated to the nervous system directly from its cochlear source [6,23]. Thus, lesser cochlear distortion broadcast into the FFR may provide a sensitive measure for detecting early SNHL, as implied by the negative correlation between DPOAEs and FFR_{THD}. Moreover, DPOAEs and FFR_{THD} together might help differentiate pre-neural and neural abnormalities in nonlinearity [6]. Future studies in SNHL listeners are needed to test these possibilities. Collectively, our data suggest a more global measure of neural nonlinearity/distortion may offer an alternate way to evaluate individual differences in suprathreshold hearing skills (e.g. SIN perception) and perhaps subtle changes in auditory health not captured by normal hearing evaluation.

Acknowledgements

We thank Jessany Knapp and Victoria Heitzmann for their help in data collection.

This work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award number NIH/NIDCD R01DC016267 (G.M.B.).

Conflicts of interest

There are no conflicts of interest.

References

1 Moore BCJ. *Introduction to the Psychology of Hearing*. San Diego: Academic Press; 2003.
2 Ruggero MA. Physiology and coding of sound in the auditory nerve. In: Popper AN, Fay RR, editors. *The Mammalian Auditory Pathways: Neurophysiology*. New York: Springer-Verlag; 1992. pp. 34–93.
3 Plack CJ, Drga V, Lopez-Poveda EA. Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss. *J Acoust Soc Am* 2004; **115**:1684–1695.
4 Neely ST, Gorga MP, Dorn PA. Cochlear compression estimates from measurements of distortion-product otoacoustic emissions. *J Acoust Soc Am* 2003; **114**:1499–1507.
5 Dorn PA, Konrad-Martin D, Neely ST, Keefe DH, Cyr E, Gorga MP. Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears. *J Acoust Soc Am* 2001; **110**:3119–3131.
6 Elsisy H, Krishnan A. Comparison of the acoustic and neural distortion product at 2f1-f2 in normal-hearing adults. *Int J Audiol* 2008; **47**:431–438.
7 Kujawa SG, Liberman MC. Synaptopathy in the noise-exposed and aging cochlea: primary neural degeneration in acquired sensorineural hearing loss. *Hear Res* 2015; **330**:191–199.
8 Oxenham AJ, Bernstein JG, Penagos H. Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci U S A* 2004; **101**:1421–1425.
9 Lins OG, Picton PE, Picton TW, Champagne SC, Durieux-Smith A. Auditory steady-state responses to tones amplitude-modulated at 80-110 Hz. *J Acoust Soc Am* 1995; **97**:3051–3063.
10 Bhagat SP, Champlin CA. Evaluation of distortion products produced by the human auditory system. *Hear Res* 2004; **193**:51–67.
11 Smalt CJ, Krishnan A, Bidelman GM, Ananthakrishnan S, Gandour JT. Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hear Res* 2012; **292**:26–34.
12 Bidelman GM. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage* 2018; **175**:56–69.
13 Shaheen LA, Valero MD, Liberman MC. Towards a diagnosis of cochlear neuropathy with envelope following responses. *J Assoc Res Otolaryngol* 2015; **16**:727–745.
14 Johannesen PT, Buzo BC, Lopez-Poveda EA. Evidence for age-related cochlear synaptopathy in humans unconnected to speech-in-noise intelligibility deficits. *Hear Res* 2019; **374**:35–48.
15 Shmilovitz D. On the definition of total harmonic distortion and its effect on measurement interpretation. *IEEE Transactions on Power Delivery* 2005; **20**:526–528.
16 Bidelman G, Powers L. Response properties of the human frequency-following response (FFR) to speech and non-speech sounds: level dependence, adaptation and phase-locking limits. *Int J Audiol* 2018; **57**:665–672.
17 Guest H, Munro KJ, Prendergast G, Plack CJ. Reliability and interrelations of seven proxy measures of cochlear synaptopathy. *Hear Res* 2019; **375**:34–43.
18 Anderson S, Skoe E, Chandrasekaran B, Kraus N. Neural timing is linked to speech perception in noise. *J Neurosci* 2010; **30**:4922–4926.

19 Bidelman GM, Bhagat SP. Right-ear advantage drives the link between olivocochlear efferent 'antimasking' and speech-in-noise listening benefits. *Neuroreport* 2015; **26**:483–487.
20 Bhagat SP. Modeling DPOAE input/output function compression: comparisons with hearing thresholds. *J Am Acad Audiol* 2014; **25**:746–759.
21 Abdala C, Kalluri R. Exploiting dual otoacoustic emission sources. Proceedings of the 12th International Workshop on the Mechanics of Hearing. AIP Conference Proceedings; Cape Sounio, Greece, June 23–29, 2014; 2015; p. 1703.
22 Greenberg S, Marsh JT, Brown WS, Smith JC. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hear Res* 1987; **25**:91–114.
23 Pandya PK, Krishnan A. Human frequency-following response correlates of the distortion product at 2F1-F2. *J Am Acad Audiol* 2004; **15**:184–197.
24 Guest H, Munro KJ, Prendergast G, Howe S, Plack CJ. Tinnitus with a normal audiogram: relation to noise exposure but no evidence for cochlear synaptopathy. *Hear Res* 2017; **344**:265–274.
25 Smith SB, Krizman J, Liu C, White-Schwoch T, Nicol T, Kraus N. Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. *Hear Res* 2019; **371**:66–74.