INVITED REVIEW

On some limitations of the frequency following response

Hedwig E. Gockel*

*MRC Cognition and Brain Sciences Unit, University of Cambridge,
15 Chaucer Road, Cambridge, CB2 7EF, United Kingdom*

Abstract: In recent years, there has been increased interest in the scalp-recorded frequency following response (FFR), which is an electrical signal that reflects sustained phase locking to sound of large populations of neurons mainly in the upper brainstem in response to stimulus-related periodicities. It provides a non-invasive measure of neural processing in humans, which can be compared to behavioural responses concerning the listener's perception. It has been argued that the FFR reflects processes important for the perception of pitch and that changes in the FFR with experience and/or training provide a measure of neural plasticity at the level of the brainstem. This paper reviews recent work aimed at elucidating the origin and the specifics of the information present in the FFR. It is argued that the neural responses measured by the FFR preserve temporal information important for pitch to a certain degree, but do not necessarily represent pitch-related processing over and above that present in the auditory periphery. In addition, multiple generators may affect the overall measure to various degrees, depending on the repetition rate of the stimulus.

Keywords: Pitch, Brainstem, Phase locking

PACS number: 43.66.Hg, 43.64.Qh [doi:10.1250/ast.41.83]

1. INTRODUCTION

The scalp-recorded frequency following response (FFR) is an electrical signal that reflects sustained phase locking of large populations of auditory neurons, mainly in the upper brainstem, in response to stimulus-related periodicities in sound and that is recorded during ongoing stimulation [1–3]. Sometimes, when related to the repetition rate of complex tones, it is also called envelope following response (EFR). It provides a non-invasive measure of neural processing in humans, which can be compared to behavioural responses concerning the listener's perception.

In its simplest form, the FFR can be measured with only three electrodes, and the choice of electrode configuration is likely to affect the relative contribution of different anatomical generators to the measure [4,5]. The anatomical generators contributing to the FFR are often determined from the latency of the FFR. Here we are concerned with the FFR having a latency between 6 and 10 ms [2,3,6], suggesting a main generation site at the level of the inferior colliculus (IC) or lateral lemniscus (LL).

In recent years, there has been increased interest in the FFR, not only as a potentially useful tool for understanding the neural mechanisms underlying various aspects of

auditory perception in humans, but also as a potential objective measure in clinical assessment. Several authors have suggested that the FFR reflects a representation of pitch at the level of the brainstem [4,7–10] and that it is plastic and/or may have clinical applications. More specifically, it has been shown that the FFR is enhanced in musicians [11,12] and speakers of a tone language [13], and that it can also be influenced by short-term learning [14]. Russo *et al.* [15] reported that the FFR may be abnormal in clinical populations, such as in children with autism spectrum disorders, for whom prosodic information processing is impaired. The strength of the FFR is negatively correlated with age, i.e. older people show a decrease in some FFR measures [16,17], and with frequency difference limens (FDLs), i.e. listeners with stronger FFRs tend to have lower FDLs [17]. It has been reported that middle-aged musicians may have enhanced speech in noise perception relative to middle-aged non-musicians [18], although this could not be replicated by [19], and it has been suggested that musical and auditory training may offset age-related delays in neural timing and improve speech in noise perception [20,21]. The FFR has also been investigated as a tool for assessing the possible connection between cochlear synaptopathy [22,23] and “hidden hearing loss” i.e. reduced suprathreshold coding abilities despite normal absolute hearing thresholds [24–26].

*e-mail: hedwig.gockel@mrc-cbu.cam.ac.uk

The aim of the present paper is to review some work done in (or in collaboration with) our laboratory over the recent past, measuring the FFR in humans, to better understand its origin and its limitations in relation to human perception.

2. THE FFR AND PITCH

Greenberg and colleagues were among the first to suggest that the FFR reflects the neural representation of the pitch of complex sounds at the level of the brainstem [7–9]. Some of the observations that have been used in support of this were: (1) the FFR to harmonic complex tones contains a frequency component corresponding to the fundamental frequency (F0), even when that component is absent in the spectrum of the stimulus and in the presence of a low-pass noise, mirroring the well-known behavioral finding of the percept of the “missing fundamental” (the residue pitch) even in the absence of that component in the stimulus spectrum and in the presence of a low-pass noise that would be able to mask a distortion product at that frequency [27]; (2) the sensitivity of the FFR to the phases of the components in a complex tone depends on the harmonic numbers of the components, in a similar way as human pitch perception for the complex tone [9].

While many authors refer to effects on pitch “encoding” at the brainstem level when describing effects of stimulus manipulation or listener group on the periodicity information present in the FFR, sometimes stronger statements have been made to the effect that the brainstem “extracts” pitch [15] or that the FFR reflects “voice pitch processing” [13]. But while the FFR clearly reflects temporal information that could be used to estimate pitch, this does not necessarily mean that the FFR reflects neural processes involved in the extraction of pitch. Instead, it might simply reflect (a lowpass filtered version of) the neural representation of sounds in the auditory periphery. More specifically, a spectral component in the FFR at the missing F0 could arise as a consequence of true neural coding of pitch, or as a consequence of distortion products generated in the auditory system due to the nonlinearities in the auditory periphery and in later stages of the processing up to and including the generation of the FFR, or as a consequence of the beating of components, i.e. interaction between adjacent components in the auditory periphery, where the beat rate corresponds to the frequency spacing and thus to the F0 of a harmonic complex tone. In order to shed some light on these questions, Gockel *et al.* [28] measured the FFR in two experiments for two very different types of pitch-evoking stimuli.

2.1. Experiment 1: Dichotic Complex Tones

In experiment 1, Gockel *et al.* [28] measured the FFR to dichotic complex tones, i.e. complex tones in which

individual harmonics were presented to opposite ears. Here they made use of the finding of Houtsma and Goldstein [29] that listeners can derive a missing F0 from components presented to opposite ears. Thus, the pitch is determined by the combined input from the two ears. In contrast, distortion products and beating between components are determined by the input to each ear separately, allowing for a dissociation between more peripheral factors and processes involved in the extraction of pitch.

The stimuli used by Gockel *et al.* [28] were harmonics 2, 3, and 4 of a 244-Hz F0. The FFR was measured in four conditions. In the monaural condition, all three harmonics were presented to the left ear. In the dichotic condition, harmonics 2 and 4 were presented to the left ear and harmonic 3 was presented to the right ear. In a third condition (“harm 2+4”), only harmonics 2 and 4 were presented to the left ear, and in a fourth condition (“harm 3”) only harmonic 3 was presented to the right ear. Importantly, in all conditions all harmonics were gated on simultaneously in both ears to enhance fusion across ears. Then slowly, over the next 40 ms, the unwanted component(s) was faded out. This was done to enhance fusion of harmonics across ears in the dichotic condition. Specifically, in the dichotic condition, the pitch corresponded to that of a 244-Hz F0 throughout the duration of the stimulus. The overall stimulus duration was 450 ms, but the analysis of the FFR was restricted to the time window from 100–350 ms after stimulus onset. The question addressed was whether the FFR in the dichotic condition would correspond to that for the monaural condition, in which the same pitch was perceived and for which a spectral component at 244 Hz was expected to be present in the FFR, or whether it would be more similar to the combined FFRs of conditions harm 2+4 and harm 3, in which different pitches were perceived and for which no spectral component at 244 Hz was expected in the FFRs.

The FFR was recorded using a “vertical” electrode montage (+, Fz; –, C7; ground, Fpz) from five subjects. Stimuli were presented in quiet, at a level of 70 dB SPL per component, through μ -metal-shielded Etymotic Research ER2 insert earphones, with alternating starting polarity. The FFR to alternating polarity stimuli was added (addition waveform) to emphasize contributions from neural activity related to the stimulus envelope. The FFR to inverted-polarity stimuli was subtracted from the FFR to original-polarity stimuli (subtraction waveform) to emphasize contributions from phase locking to the signal frequencies (the temporal fine structure, TFS) [30].

Figure 1 shows the averaged magnitude spectra (in decibels re 0.01 μ V) of the FFR for the addition waveform for the four conditions. In the monaural condition (Panel a), clear spectral peaks were visible at the F0 of 244 Hz (corresponding to the pitch) and its integer multiples. In

Fig. 1 Magnitude spectra of the FFR with the two polarities added (envelope-related FFR).

condition harm 2+4 (Panel b), spectral peaks were present at 488 Hz (the F0 of this stimulus) and its integer multiples. In condition harm 3 (Panel d), there was no spectral peak, consistent with the absence of a fluctuating temporal envelope for the pure tone. Importantly, in the dichotic condition (Panel c) the spectrum resembled that for condition harm 2+4; there was no spectral component at 244 Hz (corresponding to the pitch).

Figure 2 shows the averaged magnitude spectra of the FFR for the subtraction waveform. In the monaural condition, spectral peaks were visible at the component

Fig. 2 Magnitude spectra of the FFR with the two polarities subtracted (TFS-related FFR).

frequencies and at the F0. In condition harm 2+4, spectral peaks were present at the component frequencies of 488 and 976 Hz. Likewise, in condition harm 3, a spectral peak was present at 732 Hz. Importantly, in the dichotic condition the spectrum resembled the sum of the spectra for condition harm 2+4 and condition harm 3; there was no spectral component at 244 Hz (corresponding to the pitch). Thus, neither the envelope-related FFR nor the TFS-related FFR showed a peak corresponding to the pitch perceived in the dichotic condition, indicating that the peak at the F0 observed in the monaural condition was due to distortion products and/or beating rather than the result of pitch extraction as such.

2.2. Experiment 2: Frequency-shifted Complex Tones

In experiment 2, Gockel *et al.* [28] compared the FFR to harmonic complex tones with that to frequency-shifted complex tones i.e. harmonic complex tones in which all components were shifted by a fixed amount in Hertz either upwards or downwards. This led to a change in pitch that was proportional to but smaller than the frequency shift, while the envelope rate was unaffected [31–33], again allowing a dissociation between more peripheral factors and processes involved in the extraction of pitch to show up in the FFR.

All complex tones were derived from harmonics 2–4 or 3–5 of a 244-Hz F0. They were either harmonic, or all components were shifted up or down in frequency by 50% or 25% of the F0. Pitch shifts for frequency-shifted complex tones are well described by a slightly modified version of de Boer’s rule [31], also known as “the first effect of pitch shift” [32], and its size depends on the lowest harmonic present. The questions addressed were whether the FFR would have a spectral component at the frequency of the pitch match and whether the periodicity of the FFR would correspond to the frequency of the pitch match. The stimulus duration was 100 ms and FFR analysis was restricted to the time range from 12–100 ms after stimulus onset. Otherwise, the methods were the same as used in experiment 1.

Figure 3 shows the averaged magnitude spectra of the FFR for the 25% frequency shift for five conditions. Spectra for the addition and subtraction waveforms are shown by dashed and solid lines, respectively. The left and right columns give the results for the complex tones derived from harmonics 2–4 and from harmonics 3–5, respectively. The first row is for the harmonic conditions, while the second and third rows are for the downward and upward frequency shifts, respectively.

While the spectra of the envelope-related FFRs (dashed lines) showed a peak at the 244-Hz envelope rate (and its integer multiples), and the spectra of the TFS-related FFRs (solid lines) showed peaks at the component frequencies

Fig. 3 Magnitude spectra of the addition (dashed lines) and subtraction FFRs (solid lines) for harmonic and frequency-shifted complex tones.

and related distortion product frequencies, neither showed a spectral peak at the frequency corresponding to human pitch matches; see also [10]. The same pattern was observed for the 50% frequency shift. This finding contrasted with that of Greenberg *et al.* [9], who reported a spectral component close to the pitch match in the subtraction waveform for a frequency shift of 50% of the F0 for similar stimuli, but only showed the results of one subject.

The periodicity of the FFR was assessed by calculating the autocorrelation function (ACF). The ACF of the addition waveform had its highest peak at or close to a lag of $1/244$ Hz, i.e. at the period of the envelope, for all conditions. In contrast, the ACF of the subtraction waveform had its highest peak at a lag close to $1/(\text{matched pitch})$. Thus the TFS-related FFR had a periodicity similar to the matched pitch. This result is shown in Fig. 4 (circles connected by dashed lines), where the solid line shows the pitch matches described by de Boer's rule as a function of the frequency shift of the lowest harmonic present. Importantly, the summary autocorrelation function (SACF) of an auditory nerve (AN) model [34–36] also predicted the

Fig. 4 Pitch predictions for the frequency-shifted complex tones.

pitch shifts (downward pointing triangles), indicating that temporal information in the AN was preserved (to a certain degree) in the FFR and might be used at a more central stage to extract pitch.

3. THE FFR AND DISTORTION PRODUCTS (DPs)

In the previous section it was argued that the spectral peaks in the FFR at the envelope rate of a complex tone are not the result of a neural process that has extracted pitch, but rather correspond to distortion products (DPs) that can arise from any stage in the auditory system where two or more components interact and where the neural response is a nonlinear function of the input signal (e.g. rectification or compression). In contrast, audible DPs, like the cubic difference tone (CDT, with frequency $2F_1 - F_2$, where F_1 and F_2 are the frequencies of two harmonics present in the complex and $F_2 > F_1$) or the quadratic difference tone (QDT, at $F_2 - F_1$), are assumed to be propagated from the site of generation in the cochlea to their characteristic place, where they act “as if” they were present in the input signal [37].

To gain some insight into the origin of the DPs observed in the FFR, Gockel *et al.* [38] estimated the *effective* magnitudes of difference tones in the FFR and compared these with the estimated levels of audible DPs. To do this, they measured the FFR to a frequency-shifted complex tone, derived from a 300-Hz F0 harmonic complex consisting of components 2, 3 and 4, where all components were shifted down in frequency by 56 Hz, and presented at a level of 70.2 dB SPL per component. This complex potentially leads to a CDT at 244 Hz and a QDT at 300 Hz, where the latter also corresponds to the envelope

rate of the complex. They also measured the FFR to a pure tone at 244 Hz, for a large range of levels, and determined the level of the pure tone needed to give a spectral component in the FFR of the same magnitude as that observed at the DP frequency in the FFR for the complex tone (i.e. the *effective* level of the DP in the FFR).

The results showed that the effective level of the CDT in the FFR was only 4.7 dB below the input level of the primaries (components 2, 3, and 4), while that of the QDT or envelope related component in the FFR was actually 5 dB above the input level of the primaries. This contrasts with estimates of the levels of audible (i.e. propagated) DPs obtained in psychophysical studies, which showed that, for stimuli like the ones used here, the CDT is at or just below the audibility region [39,40] and the QDT is at least 20 dB below the level of the primaries [41–43]. Thus, the DPs in the FFR mainly reflect contributions other than those giving rise to the audible DPs. The former include contributions from nonlinearities at any stage in the auditory pathway up to the site of FFR generation/measurement. The QDT component in the FFR may additionally be a result of beating between adjacent partials, i.e. wave interference on the basilar membrane.

4. FREQUENCY SPECIFICITY OF THE FFR ASSESSED USING ADAPTATION

The goal of a study by Gockel *et al.* [44] was to determine the characteristic frequency of neurons from which the FFR originates. This was explored using the phenomenon of adaptation, namely the decline of the FFR in response to a steady stimulus. If the FFR is driven by neurons tuned to the target frequency, then the response to the target should be reduced most when the frequency of the adaptor is close to that of the target.

Gockel *et al.* [44] measured the FFR for a 100-ms pure tone target (T) that was preceded by a 200-ms pure tone adaptor (A) with either the same or different frequency as T. A and T either had a low frequency (L) of 213 Hz (presented at 80 dB SPL) or a high frequency (H) of 504 Hz (presented at 75 dB SPL), and they were combined in an orthogonal design, leading to four conditions (ALTL, ALTH, AHTL, and AHTH). The FFR was analyzed for five different 50-ms time windows: three over the time course of the adaptor, and two over the time course of the target (see caption of Fig. 5 for details of the time windows).

Figure 5 shows that: (i) the spectral magnitude of the FFR was significantly lower for the H than for the L tone throughout as expected, due to the well-known low-pass characteristic of the FFR; (ii) there was adaptation over the time course of A for both frequencies; (iii) most importantly, at T-start, the effect of A depended on the frequency of A only, i.e. the high-frequency adaptor reduced the FFR for T more (and by the same amount for both TH and TL)

Fig. 5 Mean spectral magnitude of TFS-related FFR at the frequency of the pure-tone stimulus (adaptor A, or target T) and the corresponding standard errors across 10 subjects. The three left-hand groups of four bars show the FFR strength during three time ranges of stimulation with the adaptor: 12–62 ms after onset (A-start), 62–112 ms after onset (A-mid), and 150–200 ms after onset (A-end). The two right-hand groups of four bars show the FFR strength during two time ranges of stimulation with the target: 12–62 ms after onset (T-start), and 50–100 ms after onset (T-end). Within each group of four bars, the two left-hand bars are for the 213-Hz target (TL), while the remaining two bars are for the 504-Hz target (TH). * $p < 0.05$; ** $p < 0.01$.

than the low-frequency adaptor. Thus, there was no evidence for audio-frequency specific adaptation in the FFR to the target. Instead, the high-frequency adaptor produced more adaptation than the low-frequency adaptor, even for the low-frequency target. The latter finding is consistent with Dau's [45] suggestion that the FFR for low-frequency tones mainly stems from activity of neurons on the upper skirt of the excitation pattern (which overlap substantially for different tone frequencies) and not from neurons tuned to frequencies around the signal frequency.

5. ATTENTIONAL MODULATION OF THE FFR AT LOWER BUT NOT HIGHER MODULATION RATES

Holmes *et al.* [46] investigated whether frequency specific attention, presumably depending on cortical activity, affects the FFR, which is usually assumed to be generated in the brainstem. Listeners were presented with three simultaneous 20-s long tone sequences. Each sequence consisted of amplitude modulated (AM) sinusoids that varied in duration but had a fixed carrier frequency (CF) and AM rate. Across sequences, the CF and the AM

rate differed. In experiment 1, which used relatively low rates, the three CFs (and associated AM rates) were 1,027 (93, Low), 1,343 (99, Mid) and 2,913 (109, High) Hz. In experiment 2, which used relatively high rates, the three CFs (and AM rates) were 1,753 (217, Low), 2,257 (223, Mid) and 4,537 (233, High) Hz. In different conditions, subjects attended either to a series of visual stimuli that were presented simultaneously with the tone sequences and performed a visual task, or they attended to one of the tone sequences. In each trial, either the sequence with the low CF or the sequence with the high CF (and AM rate) started slightly earlier than the other two sequences of tones. This cued subjects which sequence to attend to in the auditory condition; their task was to detect an occasional deviant tone of shorter duration in the attended sequence.

For experiment 1 only (lower AM rates), the results showed that the FFR strength at a given AM rate (the EFR) was significantly higher when the subjects attended to the tone sequence with that AM rate than when they attended to the sequence with a different AM rate. This did not occur for experiment 2, in which the AM rates of the attended sequences were higher. Importantly, there was a significant interaction between the factor experiment and attention, indicating that the pattern of results indeed differed significantly between the two ranges of AM rates used. Therefore, the EFRs were modulated by attention only at the lower AM rate.

Traditionally, scalp recorded EFRs have been assumed to arise mainly from the brainstem, not from the cortex, for AM rates above about 70 Hz. However, recent evidence questions this assumption. For example, a study of Coffey *et al.* [47] using magnetoencephalography, suggested that generators in the cortex contribute to the EFR at 98 Hz. The results of Holmes *et al.* [46] are consistent with the interpretation that the effect of attention on the EFR, observed at rates around 100 Hz, was due to a contribution of cortical generators to the EFR, rather than an effect of attention at the level of the brainstem.

6. SUMMARY AND CONCLUSIONS

The studies above indicated that: (i) For dichotic complex tones, the FFR resembled the sum of the responses to the stimuli presented to each ear alone. (ii) The peaks in the FFR spectra at the missing F0 were due to distortion products and/or monaural beats. (iii) The levels of the quadratic difference tone and the cubic difference tone were considerable higher than those for audible DPs. (iv) The neurons that mainly drive the FFR are not necessarily the neurons that are tuned to the signal frequency. It is concluded that the FFR reflects neural phase locking that may (or may not) be used to derive pitch, but does not reflect pitch “processing.” Thus, effects of training, experience, or language problems on the FFR

are not necessarily mediated by effects on pitch processing. Furthermore, and depending on the repetition rate of the stimulus, there might be small contributions from the cortex to the FFR, which may require re-interpretation of earlier studies.

ACKNOWLEDGMENTS

I thank Bob Carlyon, Emma Holmes, Ingrid Johnsrude, Fred Marmel, Anahita Metha, and Chris Plack for their collaboration in the work reported here, and Brian Moore for helpful comments on a previous version of this paper. The author was supported by intramural funding from the MRC (MC_UU_00005/3).

REFERENCES

- [1] J. T. Marsh, W. S. Brown and J. C. Smith, “Far-field recorded frequency-following responses: Correlates of low pitch auditory perception in humans,” *Electroencephalogr. Clin. Neurophysiol.*, **38**, 113–119 (1975).
- [2] J. C. Smith, J. T. Marsh and W. S. Brown, “Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources,” *Electroencephalogr. Clin. Neurophysiol.*, **39**, 465–472 (1975).
- [3] E. M. Glaser, C. M. Suter, R. Dasheiff and A. Goldberg, “The human frequency-following response: Its behavior during continuous tone and tone burst stimulation,” *Electroencephalogr. Clin. Neurophysiol.*, **40**, 25–32 (1976).
- [4] G. C. Galbraith, “Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli,” *Electroencephalogr. Clin. Neurophysiol.*, **92**, 321–330 (1994).
- [5] A. Krishnan, “Frequency-following response,” in *Auditory Evoked Potentials: Basic Principles and Clinical Application*, R. F. Burkard, M. Don and J. J. Eggermont, Eds. (Lippincott, Williams, and Wilkins, Philadelphia, 2006), pp. 313–333.
- [6] E. Skoe and N. Kraus, “Auditory brain stem response to complex sounds: A tutorial,” *Ear Hear.*, **31**, 302–324 (2010).
- [7] S. Greenberg, J. C. Smith, J. T. Marsh and W. S. Brown, “Human frequency following response to synthetic vowels,” *J. Acoust. Soc. Am.*, **63**, S76 (1978).
- [8] J. C. Smith, J. T. Marsh, S. Greenberg and W. S. Brown, “Human auditory frequency-following responses to a missing fundamental,” *Science*, **201**, 639–641 (1978).
- [9] S. Greenberg, J. T. Marsh, W. S. Brown and J. C. Smith, “Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones,” *Hear. Res.*, **25**, 91–114 (1987).
- [10] D. Wile and E. Balaban, “An auditory neural correlate suggests a mechanism underlying holistic pitch perception,” *PloS one*, **2**, e369 (2007).
- [11] P. C. Wong, E. Skoe, N. M. Russo, T. Dees and N. Kraus, “Musical experience shapes human brainstem encoding of linguistic pitch patterns,” *Nat. Neurosci.*, **10**, 420–422 (2007).
- [12] G. M. Bidelman, J. T. Gandour and A. Krishnan, “Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch,” *Brain Cogn.*, **77**, 1–10 (2011).
- [13] A. Krishnan, Y. Xu, J. Gandour and P. Cariani, “Encoding of pitch in the human brainstem is sensitive to language experience,” *Brain Res. Cogn. Brain Res.*, **25**, 161–168 (2005).
- [14] S. Carcagno and C. J. Plack, “Subcortical plasticity following

- perceptual learning in a pitch discrimination task,” *J. Assoc. Res. Otolaryngol.*, **12**, 89–100 (2011).
- [15] N. M. Russo, E. Skoe, B. Trommer, T. Nicol, S. Zecker, A. Bradlow and N. Kraus, “Deficient brainstem encoding of pitch in children with autism spectrum disorders,” *Clin. Neurophysiol.*, **119**, 1720–1731 (2008).
- [16] C. G. Clinard, K. L. Tremblay and A. R. Krishnan, “Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings,” *Hear. Res.*, **264**, 48–55 (2010).
- [17] F. Marmel, D. Linley, R. P. Carlyon, H. E. Gockel, K. Hopkins and C. J. Plack, “Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently,” *J. Assoc. Res. Otolaryngol.*, **14**, 757–766 (2013).
- [18] A. Parbery-Clark, E. Skoe, C. Lam and N. Kraus, “Musician enhancement for speech-in-noise,” *Ear Hear.*, **30**, 653–661 (2009).
- [19] D. Ruggles, A. Riddell, R. Freyman and A. J. Oxenham, “Intelligibility of voiced and whispered speech in noise in listeners with and without musical training,” *J. Acoust. Soc. Am.*, **133**, 3379 (2013).
- [20] A. Parbery-Clark, S. Anderson, E. Hittner and N. Kraus, “Musical experience offsets age-related delays in neural timing,” *Neurobiol. Aging*, **33**, 1483.e1–1483.e4 (2012).
- [21] S. Anderson, T. White-Schwoch, A. Parbery-Clark and N. Kraus, “Reversal of age-related neural timing delays with training,” *Proc. Natl. Acad. Sci. USA*, **110**, 4357–4362 (2013).
- [22] S. G. Kujawa and M. C. Liberman, “Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss,” *J. Neurosci.*, **29**, 14077–14085 (2009).
- [23] A. C. Furman, S. G. Kujawa and M. C. Liberman, “Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates,” *J. Neurophysiol.*, **110**, 577–586 (2013).
- [24] H. M. Bharadwaj, S. Masud, G. Mehraei, S. Verhulst and B. G. Shinn-Cunningham, “Individual differences reveal correlates of hidden hearing deficits,” *J. Neurosci.*, **35**, 2161–2172 (2015).
- [25] G. Prendergast, H. Guest, K. J. Munro, K. Kluk, A. Leger, D. A. Hall, M. G. Heinz and C. J. Plack, “Effects of noise exposure on young adults with normal audiograms I: Electrophysiology,” *Hear. Res.*, **344**, 68–81 (2017).
- [26] H. Guest, K. J. Munro, G. Prendergast, R. E. Millman and C. J. Plack, “Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure,” *Hear. Res.*, **364**, 142–151 (2018).
- [27] J. C. R. Licklider, “Auditory frequency analysis,” in *Information Theory*, C. Cherry, Ed. (Academic Press, New York, 1956), pp. 253–268.
- [28] H. E. Gockel, R. P. Carlyon, A. Mehta and C. J. Plack, “The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch,” *J. Assoc. Res. Otolaryngol.*, **12**, 767–782 (2011).
- [29] A. J. M. Houtsma and J. L. Goldstein, “The central origin of the pitch of complex tones: Evidence from musical interval recognition,” *J. Acoust. Soc. Am.*, **51**, 520–529 (1972).
- [30] T. J. Gobleck and R. R. Pfeiffer, “Time-domain measurements of cochlear nonlinearities using combination click stimuli,” *J. Acoust. Soc. Am.*, **46**, 924–938 (1969).
- [31] E. de Boer, *On the “residue” in hearing* (Ph.D. Thesis, University of Amsterdam, Amsterdam, 1956).
- [32] J. F. Schouten, R. J. Ritsma and B. L. Cardozo, “Pitch of the residue,” *J. Acoust. Soc. Am.*, **34**, 1418–1424 (1962).
- [33] R. D. Patterson and F. L. Wightman, “Residue pitch as a function of component spacing,” *J. Acoust. Soc. Am.*, **59**, 1450–1459 (1976).
- [34] R. Meddis and M. Hewitt, “Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery. I: Pitch identification,” *J. Acoust. Soc. Am.*, **89**, 2866–2882 (1991).
- [35] R. Meddis and L. O’Mard, “A unitary model of pitch perception,” *J. Acoust. Soc. Am.*, **102**, 1811–1820 (1997).
- [36] C. J. Summer, E. A. Lopez-Poveda, L. P. O’Mard and R. Meddis, “A revised model of the inner-hair cell and auditory-nerve complex,” *J. Acoust. Soc. Am.*, **111**, 2178–2188 (2002).
- [37] M. Furst, W. M. Rabinowitz and P. M. Zurek, “Ear canal acoustic distortion at $2f_1 - f_2$ from human ears: Relation to other emissions and perceived combination tones,” *J. Acoust. Soc. Am.*, **84**, 215–221 (1988).
- [38] H. E. Gockel, R. Farooq, L. Muhammed, C. J. Plack and R. P. Carlyon, “Differences between psychoacoustic and frequency following response measures of distortion tone level and masking,” *J. Acoust. Soc. Am.*, **132**, 2524–2535 (2012).
- [39] J. L. Goldstein, “Auditory nonlinearity,” *J. Acoust. Soc. Am.*, **41**, 676–689 (1967).
- [40] G. F. Smoorenburg, “Audibility region of combination tones,” *J. Acoust. Soc. Am.*, **52**, 603–614 (1972).
- [41] J. W. Hall, “Auditory distortion products $f_2 - f_1$ and $2f_1 - f_2$,” *J. Acoust. Soc. Am.*, **51**, 1863–1871 (1972).
- [42] D. Pressnitzer and R. D. Patterson, “Distortion products and the pitch of harmonic complex tones,” in *Physiological and Psychophysical Bases of Auditory Function*, D. J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs and R. Schoonhoven, Eds. (Shaker, Maastricht, 2001), pp. 97–104.
- [43] A. J. Oxenham, C. Micheyl and M. V. Keebler, “Can temporal fine structure represent the fundamental frequency of unresolved harmonics?” *J. Acoust. Soc. Am.*, **125**, 2189–2199 (2009).
- [44] H. E. Gockel, A. Krugliak, C. J. Plack and R. P. Carlyon, “Specificity of the human frequency following response for carrier and modulation frequency assessed using adaptation,” *J. Assoc. Res. Otolaryngol.*, **16**, 747–762 (2015).
- [45] T. Dau, “The importance of cochlear processing for the formation of auditory brainstem and frequency following responses,” *J. Acoust. Soc. Am.*, **113**, 936–950 (2003).
- [46] E. Holmes, D. W. Purcell, R. P. Carlyon, H. E. Gockel and I. S. Johnsrude, “Attentional modulation of envelope-following responses at lower (93–109 Hz) but not higher (217–233 Hz) modulation rates,” *J. Assoc. Res. Otolaryngol.*, **19**, 83–97 (2018).
- [47] E. B. Coffey, S. C. Herholz, A. M. Chepesiuk, S. Baillet and R. J. Zatorre, “Cortical contributions to the auditory frequency-following response revealed by MEG,” *Nat. Commun.*, **7**, 11070 (2016).