Cochlear, brainstem, and psychophysical responses show spectrotemporal tradeoff in human auditory processing

Gavin M. Bidelman[a,b] and Shaum P. Bhagat[a]

Auditory filter theory posits a tradeoff in time–frequency analysis: high temporal precision is achievable only at the expense of poorer frequency resolution and vice versa. Here, we examined the hierarchy of brain mechanisms of these spectrotemporal tradeoffs through a series of physiological and behavioral measures aimed to tap temporal and spectral acuity at different levels of the auditory neuroaxis (cochlea→brainstem→percept). Cochlear and behavioral frequency selectivity was measured by stimulus–frequency otoacoustic emissions (SFOAE) and psychophysical tuning curves; temporal acuity was measured physiologically and behaviorally by paired click recovery of auditory brainstem responses (ABRs) and gap detection thresholds (GDTs), respectively. Comparison of physiological and behavioral estimates of temporal acuity and frequency tuning showed high consistency between measurement domains with temporal thresholds of ~ 3–4 ms and filter tuning $Q_3 \approx 10$ across brain and behavioral measures. Cochlear SFOAE estimates of tuning inversely predicted listeners' temporal acuity estimated from both brainstem ABRs and behavioral GDTs. The high predictive power of cochlear responses on temporal thresholds and similarity between time–frequency tradeoffs measured at progressively higher levels of the processing hierarchy (brainstem, behavior) suggest that the temporal resolution of human hearing established in the cochlea might be inherited at progressively higher levels of the hearing pathway. *NeuroReport* 28:17–22 Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

NeuroReport 2017, **28**:17–22

Keywords: auditory brainstem response, cochlear filtering, pitch discrimination, stimulus–frequency otoacoustic emissions, temporal resolution, time–frequency analysis

[a]School of Communication Sciences and Disorders and [b]Institute for Intelligent Systems, University of Memphis, Memphis, Tennessee, USA

Correspondence to Gavin M. Bidelman, PhD, School of Communication Sciences and Disorders, University of Memphis, 4055 North Park Loop, Memphis, TN 38152, USA
Tel: + 1 901 678 5826; fax: + 1 901 525 1282; e-mail: g.bidelman@memphis.edu

Received 13 September 2016 **accepted** 2 November 2016

Introduction

Peripheral auditory processing involves a spectral decomposition of the sound input imposed by the filtering properties of the cochlea. Theoretically, there exists an inherent tradeoff between frequency and temporal resolution such that a filter's bandwidth and its time constant (i.e. impulse response duration) are inversely related [1]; improved frequency resolution is achieved at the expense of poorer temporal resolution and vice versa [2]. These predictions are supported by basilar membrane and auditory nerve fiber responses [3,4] that show that broader filters (e.g. with hearing loss) produce an improvement (decrease) in response latency (i.e. first spike timing) [4].

In human listeners, psychophysical studies show that the width of the cochlear filters (measured by masking experiments) increases [5] and the temporal resolution [measured by gap detection thresholds (GDTs)] decreases at higher center frequencies (CFs), consistent with time–frequency tradeoffs of filter theory [6]. Yet, psychophysical data involve perceptual masking paradigms that engage the entire auditory system. Thus, it remains unclear whether time–frequency tradeoffs in auditory processing are initiated in cochlear filtering *per se.* Moreover, tradeoffs have been observed at higher levels of the auditory system (e.g. brainstem) [7]. Without a direct physiological assay of cochlear frequency tuning, it remains unclear whether spectrotemporal tradeoffs observed in previous reports [7,8] truly reflect limits in cochlear filtering or processes higher in neural hierarchy.

Here, we extend previous work [7,8] by using two behavioral and physiological measures of frequency and temporal resolution to investigate the interplay between these factors and spectrotemporal tradeoffs across different levels of the human auditory system. Frequency resolution was measured behaviorally by psychophysical tuning curve (PTCs) and physiologically with stimulus–frequency otoacoustic emission (SFOAE) tuning curves [9,10]. Temporal resolution was assessed behaviorally with GDTs and neurally by paired click recovery of auditory brainstem responses (ABRs) [7,8]. Comparisons between frequency selectivity (SFOAE, PTCs) and temporal (GDTs, ABR) measures allowed us to assess possible spectrotemporal tradeoffs in auditory processing. Physiological (SFOAE, ABR) and psychophysical (PTCs, GDTs) comparisons evaluated time–frequency relations across different stages of the auditory neuroaxis (i.e. cochlea→brainstem→perception).

Methods

Participants

Eleven, right-handed, normal-hearing adults (five women; age: 28.9 ± 3.7 years) participated in the experiment. Audiometry confirmed normal hearing thresholds (< 5 dB HL) bilaterally between 250 and 8000 Hz. None

0959-4965 Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved. DOI: 10.1097/WNR.0000000000000714

reported a history of neuropsychiatric illness. Participants provided written informed consent in compliance with a protocol approved by the Institutional Review Board at The University of Memphis.

Behavioral tasks

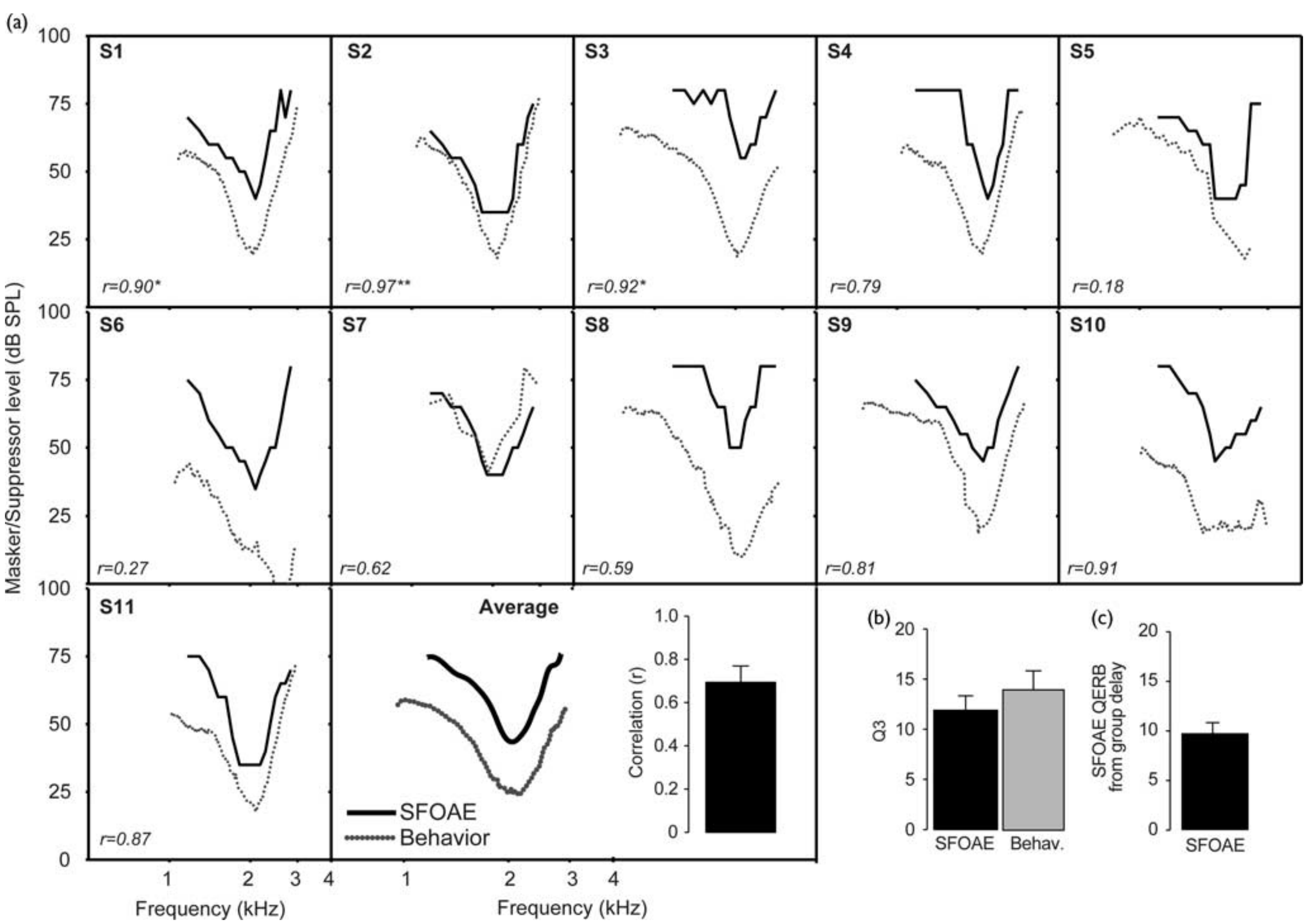
Psychophysical tuning curves

Behavioral frequency selectivity was assessed by fast-mapped PTCs [11,12]. Briefly, listeners monitored a low-intensity (18 dB SPL) 2 kHz probe tone concurrent with a narrowband noise masker (320 Hz bandwidth) that varied in CF. The probe was a 500 ms, pulsed (ISI: 200 ms) pure tone. Masker CF swept upward from 700 to 3000 Hz over 4 min. Masker level was varied continuously according to a Békésy track (2 dB/s). The run began with initial masker set at 50 dB SPL. Participants held a button so long as the probe tone remained audible and released it when it became inaudible. Masked threshold plotted against masker CF provided an estimate of a listener's PTC at 2 kHz. Data from two runs were averaged to construct the final PTC per listener. Filter 'sharpness' was then quantified from smoothed (two-point average [12]) PTCs by measuring the quality (Q) factor, a normalized measure of filter 'sharpness'. Q_3 was computed as $Q_3 = f_c/BW_{3\,dB}$, where $BW_{3\,dB}$ was the bandwidth + 3 dB above the filter's center frequency (f_c). Higher Q denotes superior tuning.

Gap detection thresholds

Behavioral temporal acuity was measured by GDTs [13] in a three-alternative forced-choice task using the PsyAcoustX MATLAB GUI [14]. Participants heard three sequential intervals (assigned randomly): two contained a contiguous 200-ms noise (3 ms gating) and one

Fig. 1



Cochlear and behavioral frequency selectivity. (a) Psychophysical tuning curves (gray dotted lines) measured by simultaneous masking [12] and cochlear tuning curves measured by suppression of SFOAEs (black) measured at 2 kHz. Single panels, individual participants; bottom right, grand average. The inset shows the correlation between physiological and psychophysical filter estimates. (b) Filter quality factors (i.e. sharpness) measured by Q_3. Cochlear and behavioral tuning estimates show similar degrees of tuning, suggesting that perceptual frequency selectivity is predicted on the basis of initial peripheral (cochlear) processing. (c) Alternate measures of cochlear tuning (Q_{ERB}) on the basis of analysis of the group delay of SFOAE responses [16]. Sharper auditory filters cause longer round-trip travel times of the cochlear response. Q_{ERB} estimates agree with Q_3 measures of tuning curves. $^*P < 0.05$, $^{**}P < 0.01$.

contained a brief gap (temporally centered). The noise was filtered (1200–3100 Hz) and spectrally centered at 2-kHz to restrict cochlear location and mimic SFOAE/ABR recordings (detailed below). Participants identified the interval containing the gap. GDTs were measured using the 2-down, 1-up adaptive tracking rule (i.e. 71% performance). Following two correct responses, gap duration (Δt) was decreased (i.e. made harder) for the subsequent trial and was increased (i.e. made easier) following a single incorrect response. The geometric mean of the final 8/14 reversals was used to compute each listener's GDT. Stimulus level was set at 80 dB SPL. Two separate runs were averaged to provide a more stable estimate of GDTs.

Stimulus–frequency otoacoustic emission tuning curves

SFOAE TCs were measured by a well-established suppression method [10,15] using a low-noise ER-10C microphone (Etymotic Research Elk Grove Village, Illinois, USA). Briefly, ear-canal sound-pressure level was recorded during presentation of a single probe tone ($f_{\mathrm{probe}} = 2$ kHz) in isolation and when paired with a suppressor tone (f_{supp}). The vector difference in ear canal pressure at the emission frequency (f_{probe}) with and without the presence of the suppressor tone resulted in an SFOAE residual, corresponding to the part of the emission suppressed by f_{supp} [9]. Sixteen suppressor tones were varied within the range of 0.6–1.4 f_{probe}. Probe tone level was set between 25 and 35 dB SPL. f_{supp} Level was then varied using a tracking procedure (5 dB step size) and was terminated when the SFOAE residual was within ± 2 dB of a residual criterion of 0 dB SPL [9]. The tracking procedure was terminated if the residual criterion was not fulfilled when the level of the suppressor reached 85 dB SPL.

SFOAE tuning was assessed by Q_3. In addition, tuning can be measured by the emission's group delay (called Q_{ERB}) under the assumption that longer round-trip travel times of the cochlear response are indicative of sharper cochlear filters [16,17]. Group delay (τ_{SFOAE}) was estimated from the negative slope of the unwrapped phase of the response. Q_{ERB} was then computed as $Q_{\mathrm{ERB}} = k (N_{\mathrm{SFOAE}}/2)$, where $N_{\mathrm{SFOAE}} = f \times \tau_{\mathrm{SFOAE}}$, the SFOAE delay in periods at the emission frequency f, τ_{SFOAE} is the group delay, and $k = 2.3 f^{-0.07}$ [16,17].

Physiological temporal resolution: auditory brainstem responses

We measured auditory temporal resolution physiologically by recovery of ABR amplitude to paired click stimuli [7,8]. Similar variants of this approach have been used widely to study temporal processing at several levels of the auditory system including auditory nerve [18,19], brainstem [7,8,20], and thalamocortical levels [21,22]. Briefly, filtered clicks were generated by applying 0.25 ms ramps ($\cos^2$ window) to a 0.67 ms, 2-kHz sinusoid. Filtered clicks allowed us to obtain more frequency-specific ABRs and directly compare neural, cochlear, and behavioral responses at roughly the same cochlear location (2 kHz). Paired-click stimuli (×3000 trials) were created with interclick intervals (ICIs) from 25 to 0 ms see Fig. 2). Stimuli were presented binaurally at 80 dB SPL (comparable level to GDTs) through ER-2 insert earphones.

EEGs were recorded differentially between Ag/AgCl electrodes placed on the scalp at the high forehead (~Fpz) referenced to linked mastoids (A1/A2) with the mid-forehead as ground [7]. Interelectrode impedance was up to 3 kΩ. EEGs were digitized at 10 kHz (SynAmps RT amps; Compumedics Neuroscan Charlotte, North Carolina, USA), epoched (0–40 ms window), averaged in the time domain for each ICI, and filtered (250–1200 Hz) for ABR quantification [7]. Artifactual sweeps ($> \pm 50\,\mu$V) were rejected before averaging.

ABRs to the first and second click overlapped in time. For each condition, the second click was isolated by point-by-point subtraction of the ABR to the single click stimulus (ICI = 0 ms) from the ABR to the paired click stimulus [7,8]. ABR recovery was defined as the proportion between the peak response amplitude to the second click divided by the peak response amplitude to the single click stimulus [8]. Temporal thresholds were then estimated from ABR recovery functions as the ICI yielding 50% ABR amplitude recovery [7,8].

Results

Cochlear and behavioral frequency selectivity

Overlays of behavioral PTCs and SFOAE tuning curves measured at the 2 kHz region of the cochlea are shown in Fig. 1. Both classes of tuning curves show the typical 'V-shape' with an extended low-frequency tail, highly selective tip frequency, and steep high-frequency skirt characteristic of auditory filters measured with a variety of techniques [3,7,11,12]. Across listeners, SFOAE and PTC curves show close correspondence with one another (Fig. 1a, bar inset) (Spearman's $r_s = 0.67 \pm 0.2$), indicating close correspondence between behavioral frequency selectivity and peripheral cochlear tuning.

Quantitative analysis of filters showed similar degrees of Q_3 or 'sharpness' between cochlear and behavioral estimates of tuning ($t_{10} = -0.76$, $P = 0.46$) (Fig. 1b and c). This suggests that behavioral frequency selectivity largely parallels cochlear frequency selectivity for low-intensity stimuli. In addition, estimates of cochlear tuning by filter group delay (i.e. Q_{ERB}) [16,23] showed close correspondence to filter Q_3 measures (Fig. 1c). The convergence between metrics indicates that both provide similar estimates of the 'sharpness' of the auditory filter.

Neurophysiological and behavioral temporal resolution

ABRs to paired click stimuli are shown in Fig. 2a. Time-waveforms indicate a monotonic increase in response amplitude with increasing ICI, ranging from total suppression (< 30%, 1-ms click spacing) to near full recovery (>90%, 25 ms spacing) (Fig. 2b). A mixed-model analysis of variance (ICI as fixed, patients as random factor) showed a significant effect of ICI on ABR amplitude recovery ($F_{9,90}=9.53$, $P<0.0001$, $\eta_p^2=0.49$). Bonferroni-adjusted multiple comparisons indicated that recovery was much lower for short (0.7, 1, 1.5 ms) compared with longer (7, 10, 25 ms) ICIs ($t_{90}=8.80$, $P<0.0001$). Neurophysiological thresholds (i.e. ICI yielding 50% recovery [7,8]) showed temporal resolutions of ~4-ms (t-test against a null of 0 ms threshold; $t_{10}=4.69$, $P<0.001$). These estimates on the basis of brainstem ABRs are consistent with the temporal resolution measured psychophysically [24,25].

Brain–behavioral relationships reflecting spectrotemporal resolution tradeoffs

To explore potential spectrotemporal tradeoffs between levels of auditory processing, we carried out a correlational analysis between tuning (behavioral PTC Q_3, SFOAE Q_3) and temporal resolution measures (behavioral GDTs, ABR recovery thresholds).

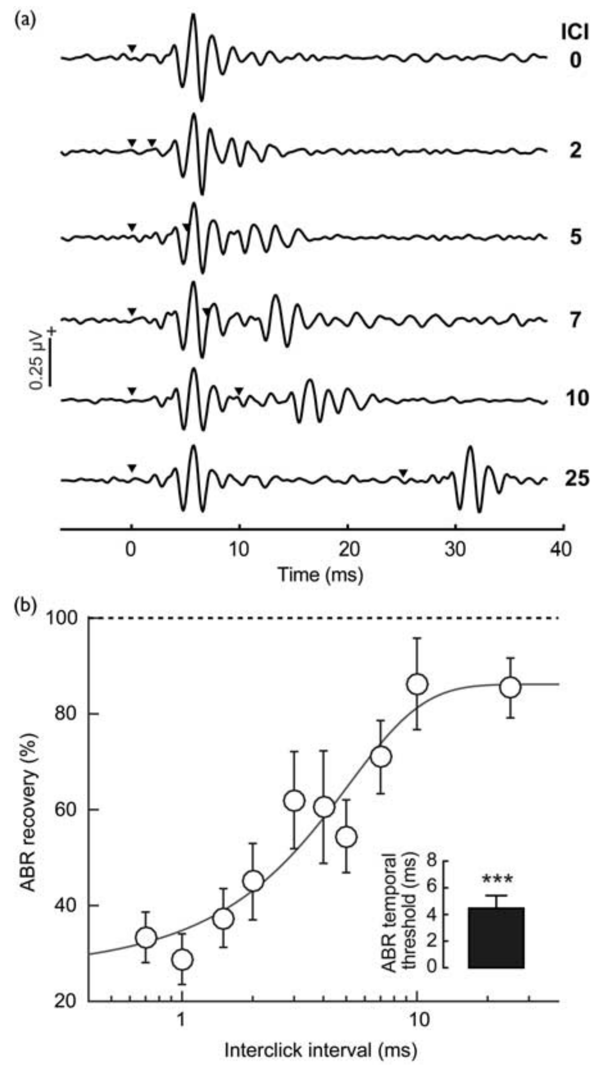
We found that cochlear tuning, measured by SFOAE filter Q_3, showed a robust correlation with psychophysical GDTs ($r_s=0.76$, $P=0.015$) (Fig. 3a). As predicted by the tradeoffs of filter theory, sharper cochlear tuning (i.e. higher filter Q_3) was associated with poorer gap detection acuity. Although the association was weaker, we found a similar positive prediction between SFOAE filter tuning and temporal thresholds for ABR recovery ($r_s=0.64$, $P=0.04$) (Fig. 3b), where sharper cochlear tuning was associated with poorer neurophysiological temporal acuity. Collectively, these findings indicate an inverse relationship in auditory spectral and temporal resolution, whereby sharper, more selective cochlear tuning (i.e. higher Q_3) predicts poorer auditory temporal processing. The fact that these spectrotemporal tradeoffs are observed between both physiological (OAE vs. ABR) and behavioral (OAE vs. behavior) measures indicates that temporal processing at multiple levels of the auditory pathway might be governed by initial cochlear frequency decomposition.

Discussion

The results of this study relate to three main observations. We found the following: (i) comparable filter sharpness between cochlear (SFOAE) and behavioral PTCs ($Q_3\approx10$), suggesting that retrocochlear frequency resolution might be largely determined by cochlear tuning (e.g. 10); (ii) behavioral and neural (ABR) estimates of temporal resolution showed high agreement (~3–4 ms), indicating close correspondence between physiological and behavioral temporal processing; and (iii) OAE estimates of cochlear tuning inversely predicted listeners' temporal acuity estimated from both neural (ABR) and behavioral (GDT) assays, suggesting that temporal processing at multiple higher levels of the auditory pathway might be governed by initial cochlear frequency decomposition.

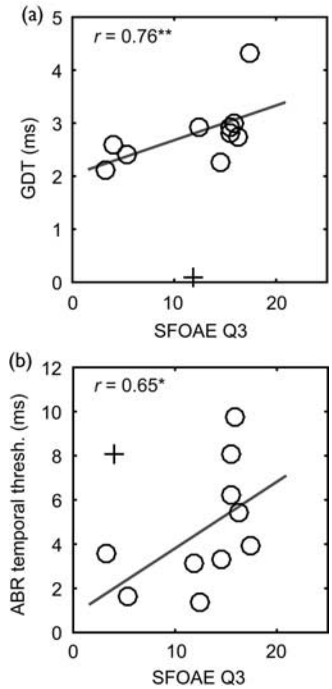
Conceivably, differences in temporal acuity and frequency tuning across listeners could reflect differences in peripheral auditory filter bandwidth. This notion is

Fig. 2



ABR responses to paired click stimuli and neurophysiological temporal resolution thresholds. (a) Grand–averaged time-waveforms for select ICIs. Arrows, click onsets. (b) ABR amplitude recovery across ICIs. Recovery is defined as the amplitude of the response to the second click expressed as a percentage of the amplitude evoked by the single click (ICI = 0 ms). Note the larger response suppression at shorter (1 ms) ICIs and near full recovery at longer ICIs (20 ms). ABR temporal thresholds were defined as the ICI yielding 50% ABR recovery [7]. (inset) Estimated auditory temporal resolution from ABRs is ~4 ms. Error bars, ± SEM ***$P<0.001$. ICIs, interclick intervals.

Fig. 3



Neural correlates of spectrotemporal tradeoff in the human auditory system. (a) Correspondence between cochlear tuning and behavioral gap detection thresholds (GDTs). Peripheral cochlear tuning (i.e. filter 'sharpness') is positively associated with GDTs such that superior tuning predicts poorer temporal processing. (b) Physiological cochlear tuning is similarly associated with neural ABR temporal recovery thresholds; sharper peripheral frequency tuning is associated with poorer (i.e. larger) neurophysiological temporal thresholds. Starred points (+) denote data points deemed influential observations by Cook's [26] excluded from the regression analyses before least-squares fitting. $^*P<0.05$, $^{**}P<0.01$.

supported by psychophysical findings in hearing-impaired listeners [27] and computational modeling [8], both of which indicate that wider filters generate faster temporal responses (i.e. spectrotemporal tradeoff). Although our cohort all had normal hearing, we have also recently shown that salient forms of auditory listening experience (e.g. musical training) can sharpen human cochlear tuning, estimated by SFOAE tuning curves [10]. This suggests that previous listening experience can modify cochlear tuning and, by proxy, might impact temporal resolution by the inherent time–frequency tradeoffs of cochlear filtering. Indeed, listeners in this study with sharper cochlear filters (i.e. Q_3 of SFOAE tuning curves) showed poorer temporal processing at both neural and behavioral levels as indexed by higher pair-click ABR recovery thresholds and GDTs, respectively. It is possible that normal variations in cochlear tuning (perhaps because of previous listening experience) may account for the distribution and correlations with ABR and behavioral temporal thresholds observed in this and previous studies [7].

Collectively, our data here show a strong reciprocal relation between cochlear spectral resolution and both (i) neural ABR (brainstem) and (ii) behavioral (GDT) temporal resolvability. Taken together with our other studies on the hierarchy of spectrotemporal tradeoffs [7,8], the high predictive power of cochlear responses but similarity between brainstem and behavioral temporal thresholds leads us to infer that the perceptual resolution of human temporal acuity (at least at 2 kHz) is established in the cochlea and is maintained at progressively higher levels of the hearing pathway.

Acknowledgements

The authors thank Ameenuddin Syed Khaja for assistance in data collection. This work was supported by grants from the American Hearing Research Foundation (AHRF) and the American Academy of Audiology (AAA) Foundation awarded to G.M.B.

Conflicts of interest

There are no conflicts of interest.

References

1 Duifhuis H. Consequences of peripheral filter selectivity for nonsimultaneous masking. *J Acoust Soc Am* 1973; **54**:1471–1488.
2 Viemeister NF, Plack CJ. Time analysis. In: Yost WA, Popper AN, Fay RR, editors. *Human Psychophysics*. New York, NY: Springer; 1993. pp. 116–154.
3 Ruggero MA, Narayan SS, Temchin AN, Recio A. Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. *Proc Natl Acad Sci USA* 2000; **97**:11744–11750.
4 Henry KS, Kale S, Heinz MG. Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers. *Front Sys Neurosci* 2014; **8**:20.
5 Glasberg BR, Moore BC. Derivation of auditory filter shapes from notched-noise data. *Hear Res* 1990; **47**:103–138.
6 Fitzgibbons P, Wightman F. Gap detection in normal and hearing-impaired listeners. *J Acoust Soc Am* 1982; **72**:761–765.
7 Bidelman GM, Syed Khaja A. Spectrotemporal resolution tradeoff in auditory processing as revealed by human auditory brainstem responses and psychophysical indices. *Neurosci Lett* 2014; **572**:53–57.
8 Henry KS, Gall MD, Bidelman GM, Lucas JR. Songbirds tradeoff auditory frequency resolution and temporal resolution. *J Comp Physiol A* 2011; **197**:351–359.
9 Charaziak KK, Souza P, Siegel JH. Stimulus-frequency otoacoustic emission suppression tuning in humans: comparison to behavioral tuning. *J Assoc Res Oto* 2013; **14**:843–862.
10 Bidelman GM, Nelms C, Bhagat SP. Musical experience sharpens human cochlear tuning. *Hear Res* 2016; **335**:40–46.
11 Bidelman GM, Schug JM, Jennings SG, Bhagat SP. Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians. *J Acoust Soc Am* 2014; **136**:EL33–EL39.
12 Sek A, Alcantara J, Moore BCJ, Kluk K, Wicher A. Development of a fast method for determining psychophysical tuning curves. *Int J Audiol* 2005; **44**:408–420.
13 Florentine M, Buus S. Temporal gap detection in sensorineural and simulated hearing impairments. *J Speech Hear Res* 1984; **27**:449–455.
14 Bidelman GM, Jennings SG, Strickland EA. PsyAcoustX: A flexible MATLAB package for psychoacoustics research. *Front Psychol* 2015; **6**:1–11.
15 Shera CA, Guinan JJ. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J Acoust Soc Am* 1999; **105**:782–798.
16 Shera CA, Guinan JJ Jr., Oxenham AJ. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc Natl Acad Sci USA* 2002; **99**:3318–3323.
17 Bhagat SP, Kilgore C. Efferent-mediated reduction in cochlear gain does not alter tuning estimates from stimulus-frequency otoacoustic emission group delays. *Neurosci Lett* 2014; **559**:132–135.

18 Ohashi T, Ochi K, Nishino H, Kenmochi M, Yoshida K. Recovery of human compound action potential using a paired-click stimulation paradigm. *Hear Res* 2005; **203**:192–200.
19 Parham K, Zhao HB, Kim DO. Responses of auditory nerve fibers of the unanesthetized decerebrate cat to click pairs as simulated echoes. *J Neurophysiol* 1996; **71**:17–29.
20 Poth EA, Boettcher FA, Mills JH, Dubno JR. Auditory brainstem responses in younger and older adults for broadband noises separated by a silent gap. *Hear Res* 2001; **161**:81–86.
21 Atcherson SR, Gould HJ, Mendel MI, Ethington CA. Auditory N1 component to gaps in continuous narrowband noises. *Ear Hear* 2009; **30**:687–695.
22 Joliot M, Ribary U, Llinás R. Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. *Proc Natl Acad Sci USA* 1994; **91**:11748–11751.
23 Koppl C. Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl. *(Tyto alba) J Neurosci* 1997; **19**:3312–3321.
24 Moore BC, Peters RW, Glasberg BR. Detection of temporal gaps in sinusoids: effects of frequency and level. *J Acoust Soc Am* 1993; **93**:1563–1570.
25 Eddins DA, Green DM. Temporal integration and temporal resolution. In: Moore BC, editor. *Hearing*. San Diego, CA: Academic Press; 1995. pp. 207–242.
26 Cook RD. Influential observations in linear regression. *J Amer Statist Assoc* 1979; **74**:169–174.
27 Strelcyk O, Dau T. Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am* 2009; **125**:3328–3345.