The Journal of the Acoustical Society of America

JULY 11 2024

Limitations in human auditory spectral analysis at high frequencies

Daniel R. Guest; Neha Rajappa; Andrew J. Oxenham

J. Acoust. Soc. Am. 156, 326–340 (2024)
https://doi.org/10.1121/10.0026475

Articles You May Be Interested In

Profile analysis and ripple discrimination at high frequencies

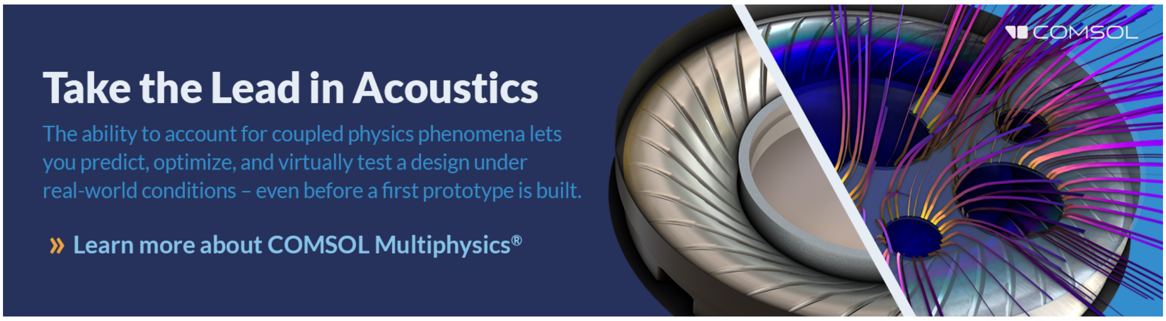
J Acoust Soc Am (April 2022)

Measuring harmonic benefit in musicians and non-musicians in several tasks

J Acoust Soc Am (April 2022)

Discrimination of depth of sinusoidal amplitude modulation with and without roved carrier levels

J. Acoust. Soc. Am. (January 2006)



ASA
ACOUSTICAL SOCIETY
OF AMERICA

JASA ARTICLE

Limitations in human auditory spectral analysis at high frequencies

Daniel R. Guest,[1,a)] Neha Rajappa,[2] and Andrew J. Oxenham[2]
[1]Department of Biomedical Engineering, University of Rochester, Rochester, New York 14642, USA
[2]Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455, USA

ABSTRACT:
Humans are adept at identifying spectral patterns, such as vowels, in different rooms, at different sound levels, or produced by different talkers. How this feat is achieved remains poorly understood. Two psychoacoustic analogs of spectral pattern recognition are spectral profile analysis and spectrotemporal ripple direction discrimination. This study tested whether pattern-recognition abilities observed previously at low frequencies are also observed at extended high frequencies. At low frequencies (center frequency ~500 Hz), listeners were able to achieve accurate profile-analysis thresholds, consistent with prior literature. However, at extended high frequencies (center frequency ~10 kHz), listeners' profile-analysis thresholds were either unmeasurable or could not be distinguished from performance based on overall loudness cues. A similar pattern of results was observed with spectral ripple discrimination, where performance was again considerably better at low than at high frequencies. Collectively, these results suggest a severe deficit in listeners' ability to analyze patterns of intensity across frequency in the extended high-frequency region that cannot be accounted for by cochlear frequency selectivity. One interpretation is that the auditory system is not optimized to analyze such fine-grained across-frequency profiles at extended high frequencies, as they are not typically informative for everyday sounds. *© 2024 Acoustical Society of America.*

https://doi.org/10.1121/10.0026475

(Received 18 October 2023; revised 4 June 2024; accepted 7 June 2024; published online 11 July 2024)

[Editor: Emily Buss] Pages: 326–340

I. INTRODUCTION

The term profile analysis refers to the ability of listeners to discriminate between complex tones based on differences in the relative amplitudes of their spectral components (Spiegel *et al.*, 1981; Green, 1983; Green and Kidd, 1983; Green *et al.*, 1983; Green and Mason, 1985; Richards *et al.*, 1989; Kidd *et al.*, 1991; Bernstein and Green, 1987, 1988; Gockel and Colonius, 1997; Gockel, 1998; Zera *et al.*, 1993; Lentz *et al.*, 1999; Maxwell *et al.*, 2020). In a typical profile-analysis experiment, listeners discriminate between a set of equal-amplitude frequency components (the reference stimulus) and the same set of frequency components with the level of one component increased, relative to the other components (the target stimulus) [Fig. 1(a)]. One hallmark of profile analysis is that it is remarkably robust to random variations in overall sound level between presentations. For example, profile-analysis thresholds [expressed in units of $20\log_{10}(\Delta A/A)$, where ΔA is the amplitude of the signal tone, added in phase to the background tone of amplitude A at the same frequency] are typically elevated by at most a few dB when sound levels are randomized across intervals compared to when they are fixed, even with level variations of 30 dB or more (Spiegel *et al.*, 1981) [Fig. 1(b)]. To overcome the effects of level roving or randomization, listeners must identify the interval in which the target level is higher than that of the other background tones, rather than simply the interval containing the highest absolute target level. Estimation of this relative level is thought to necessitate comparison of information from different frequency channels within a single interval (i.e., comparisons of intensity information across frequency, rather than across time). Such across-channel mechanisms are also implicated by the observation that performance improves when additional background components are added to the spectral edges of the stimulus, even when the frequencies of those additional components lie well outside the critical band around the target component (Green *et al.*, 1983). The results from profile-analysis experiments are of interest because they may provide a quantitative measure of listeners' ability to analyze spectral shapes, independent of overall level, in ways that are likely critical to everyday perceptual tasks, such as recognizing speech sounds.

The most frequently tested profile-analysis stimulus consists of inharmonic complex tones with logarithmically spaced frequency components (Green *et al.*, 1983; Green and Mason, 1985; Bernstein and Green, 1987, 1988; Kidd *et al.*, 1991), although some limited work has explored harmonic complex tones as well (Zera *et al.*, 1993). Although the target is typically a single central component in the complex, other variations, such as alternating patterns of attenuated and amplified components (Lentz *et al.*, 1999), have also been tested. From the numerous studies of profile analysis conducted over the past 40 years, some general

---

a)Email: Daniel_Guest@urmc.rochester.edu

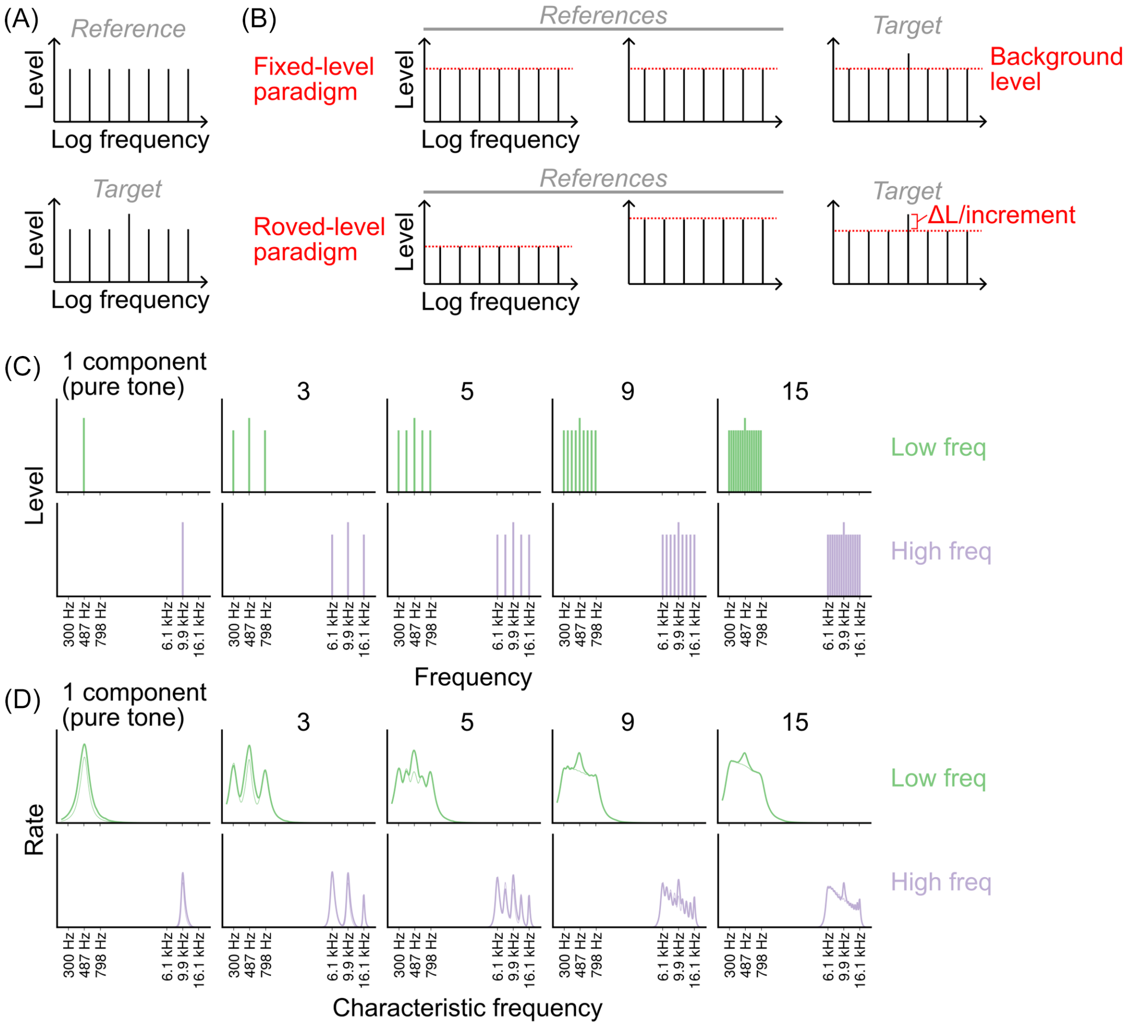


FIG. 1. (Color online) Schematic magnitude spectra for the profile-analysis stimuli. (A) An example reference stimulus spectrum (top) and target stimulus spectrum (bottom). Here, ΔL indicates the difference in level between each background component and the target component. (B) Schematic depicting the difference between profile analysis with a fixed (top row) or roved (bottom row) level. Each row shows two example reference stimulus spectra and one example target stimulus spectrum with a red dashed line denoting the background level in each stimulus. In the bottom right subplot, the cue in the target stimulus (the increment in the level of the middle component) is marked in red and labeled ΔL. (C) Schematic spectra for the profile-analysis stimuli tested in Experiment 1a. Color indicates frequency condition, with low- and high-frequency conditions plotted in green and purple, respectively. (D) Average rates as a function of characteristic frequency (CF) in response to the profile-analysis stimuli depicted in (C) from low-spontaneous-rate auditory-nerve fibers simulated using the model of Zilany *et al.* (2014). The stimulus was either a reference stimulus (thin line) or a target stimulus with an increment of 0 dB signal level, relative to the standard (SRS) (thick line; see Sec. II for details).

conditions in which listeners perform best include when components are: (1) far enough apart to avoid significant peripheral interactions between them, so as to permit an accurate estimate of the target level based on information from cochlear channels tuned near the target component; (2) sufficiently numerous and spanning a sufficient frequency range to allow for a robust and accurate estimate of the background level based on information from cochlear channels tuned away from the target component; and (3) not so far apart as to make level comparisons across relevant cochlear channels difficult (a difficulty that has also been demonstrated in tasks such as pure-tone level discrimination and loudness comparisons; see Marks, 1994, and Oxenham and Buus, 2000).

This simple interpretation of profile analysis is complicated, however, by work suggesting that both absolute and relative target frequency can impact performance. Listeners generally achieve better thresholds when the target component is the middle component of the tone complex than when it is a higher- or lower-ranked component (Green and Mason, 1985; Bernstein and Green, 1988). Performance also appears to worsen as the absolute frequency of the entire tone complex is shifted upward (Green and Mason, 1985), although this effect has only been measured in a small number of listeners for one particular combination of stimulus parameters (excepting a similar result seen for harmonic complex tones; Zera *et al.*, 1993).

There are some theoretical reasons to predict that profile analysis may deteriorate at high frequencies. First, a recent study has suggested that the ability to compare slowly fluctuating amplitudes across frequencies is poorer at high (> 6 kHz) than at low frequencies (Whiteford *et al.*, 2020), suggesting that similar deficits might be observed for the discrimination of static patterns of amplitude differences, such as those used in profile-analysis tasks. Second, one approach to explaining profile analysis has emphasized the role of hearing out the pitch of the target (Gockel and Colonius, 1997; Gockel, 1998), which has been shown to be more difficult or even impossible at frequencies above about 4 kHz for target tones embedded in both harmonic (Gockel and Carlyon, 2022) and inharmonic (Moore *et al.*, 2006) complexes. If profile analysis relies on hearing out the target and its pitch, then profile analysis should also be more difficult at high frequencies. Third, if listeners rely in any way on temporal fine structure (TFS) information to make their judgments, the roll-off of phase locking to TFS at frequencies beyond 2–3 kHz in the auditory nerve (Rose *et al.*, 1967; Weiss and Rose, 1988; Verschooten *et al.*, 2019) would also predict degradation of profile analysis at high frequencies. Based on these considerations, a reasonable hypothesis would be that profile analysis should be severely degraded at high frequencies.

Because profile analysis is thought to be determined in part by the limits of peripheral tuning, differences in tuning bandwidth across the tonotopic range must also be taken into account. In the context of the standard inharmonic profile-analysis stimuli with logarithmically spaced frequency components described previously, the relevant quantity to consider is filter sharpness, as measured by the Q factor (characteristic frequency, CF, divided by bandwidth). The Q factor of auditory filters increases at higher CFs (Shera *et al.*, 2002; Oxenham and Shera, 2003), and thus one might instead make the opposite prediction that profile analysis should improve at high frequencies relative to low frequencies, rather than worsen. Although such frequency-selectivity estimates have typically been limited to 8 kHz in humans, a study using simultaneous notched noise at extended high frequencies (Zhou, 1995) has suggested that frequency-selectivity estimates can be extrapolated reasonably well from estimates made at lower frequencies (e.g., Glasberg and Moore, 1990). The consequences of sharper filters at high frequencies is demonstrated at the level of the auditory nerve in Fig. 1(d) based on simulations from the auditory-nerve model of Zilany *et al.* (2014). Based on a qualitative evaluation of these simulations of peripheral auditory processing, it would be reasonable to predict that profile analysis should be at least as good, if not better, at high than at low frequencies.

To test these opposing predictions, we measured profile analysis at both low and high frequencies. Because level discrimination is known to differ somewhat at low and high frequencies (Jesteadt *et al.*, 1977; Florentine *et al.*, 1987), we also measured baseline pure-tone level discrimination thresholds, as well as profile-analysis thresholds in the absence of a level rove. Because the spacing between components in the stimulus is known to affect performance (Green and Mason, 1985; Bernstein and Green, 1987, 1988; Lentz *et al.*, 1999), but it was unclear *a priori* whether the component spacing that yields the best thresholds would be the same at low and high frequencies, performance was measured using several different component densities. To test whether results from profile analysis generalize to other tasks involving spectral pattern analysis, we also measured spectrotemporal ripple discrimination, the ability of listeners to discriminate between upward and downward spectrotemporal ripples (Chi *et al.*, 1999).

Consistent with the first hypothesis, we found that performance in profile analysis was much poorer at high than at low frequencies, and was poorer than predicted by differences in basic level discrimination abilities at low and high frequencies. Spectrotemporal ripple discrimination was also poorer at high frequencies than at low frequencies, especially when the information available to listeners was restricted to frequencies above about 10 kHz. The results suggest that across-channel comparisons are severely impaired at high frequencies.

II. METHODS

A. Participants and equipment

The participants were all students at the University of Minnesota with clinically normal hearing. They were recruited either through a Department of Psychology research participant pool or through an in-house participant database. All participants provided informed written consent to the experimental procedure prior to participation, and all experimental procedures were approved by the Institutional Review Board of the University of Minnesota. Participants were either paid or received course credit for their participation. A total of 98 listeners participated in some part of the study, including the screening stages described in the following. One participant was screened out based on their audiogram and another 40 participants were screened out based on their elevated extended high-frequency hearing thresholds. Of the remaining 58 participants, ten completed Experiment 1a and one completed 86% of Experiment 1a (all 11 were included in the figures and analyses), 12 completed Experiment 2a, and 22 completed Experiments 1b, 2b, and 3. Others participated in unreported pilot experiments or did not return for subsequent sessions.

Stimuli were presented to listeners over HD650 headphones (Sennheiser, Old Lyme, CT) via a Lynx E22 sound card (Lynx Studio Technologies, Costa Mesa, CA) in sound-attenuating booths. Listeners completed the experiment via a graphical user interface implemented in MATLAB (The MathWorks, Natick, MA) using the AFC framework (Ewert, 2013).

B. Screening Procedure A

Before listeners participated in Experiments 1a or 2a, they were required to pass a two-stage screening procedure.

The first stage was a standard audiogram to ensure the listeners had audiometrically normal hearing (i.e., absolute thresholds $\leq$ 20 dB hearing level, HL, at octave frequencies from 0.25 to 8 kHz) in both ears. The second stage consisted of measuring detection thresholds for diotic high-frequency pure tones presented in diotic broadband threshold-equalizing noise (TEN; Moore *et al.*, 2000), presented to each ear at 35 dB sound pressure level (SPL) within the estimated equivalent rectangular bandwidth (ERB) at 1 kHz (Glasberg and Moore, 1990). On each trial, a 350-ms pure tone with 20-ms raised-cosine onset and offset ramps was temporally centered in one of the two 500-ms noise samples, also gated with 20-ms raised-cosine onset and offset ramps, separated by a 500-ms interstimulus interval. The task of the listener was to indicate which of the two intervals contained the tone, in a forced-choice paradigm (which was also used in all other tasks). Visual feedback indicating whether the response was correct or incorrect was provided after each trial. The level of the pure tone was adaptively varied using a 3-down 1-up staircase procedure (Levitt, 1971) from a starting value of 50 dB SPL. The initial step size of the staircase procedure of 3 dB was reduced to 2 dB and then 1 dB following the second and fourth reversals of the staircase procedure, respectively. The procedure was terminated after six reversals at the smallest step size and the threshold was defined as the mean tone level at the last six reversals. Tone frequencies of 14 and 16 kHz were each tested three times. To pass the screening, listeners needed thresholds better (lower) than 45 dB SPL at both frequencies, averaged across the three runs at each frequency. This value was 5 dB lower than the lowest possible presentation level of individual spectral components in the main experiment, thereby ensuring their audibility. TEN was included in the screening stimuli even though it was not included in the experimental stimuli because the screening procedure was also used to recruit participants for pilot studies and other studies on high-frequency hearing with stimuli that included TEN (e.g., Guest and Oxenham, 2022).

C. Screening Procedure B

Before listeners participated in Experiments 1b, 2b, or 3 described in the following, they were required to pass a two-stage screening procedure. The first stage was a standard audiogram, as in Screening Procedure A. The second stage consisted of measuring detection thresholds for pure tones presented monaurally in quiet. Each trial consisted of two 350-ms intervals, visually indicated using a virtual lightbox, separated by a 350-ms silent gap. One of the intervals, selected randomly on each trial, contained a 16-kHz 350-ms pure tone with 20-ms raised-cosine ramps. The task of the listener was to indicate which interval contained the tone. Visual feedback indicating whether the response was correct or incorrect was provided after each trial. The level of the tone was varied adaptively, using the same tracking procedure and method for calculating thresholds as in Screening Procedure A. Listeners completed pairs of runs, one in the left ear and one in the right ear, and then removed their headphones and reseated them before moving to the next pair. To pass the screening, the average threshold in a single ear across two consecutive runs needed to be better (lower) than 40 dB SPL. Listeners were allowed to make multiple attempts to pass at the discretion of the experimenter, but all those who passed and were included in the data presented in the following passed on their first two attempts. The ear with the better average threshold was used for monotic presentation of the stimuli in Experiments 1b, 2b, and 3.

D. Experiment 1a: Level discrimination and profile analysis in quiet

Listeners who participated in Experiment 1a always completed the tasks in the following order: First, the first half of the level-discrimination task and the first half of the profile-analysis task with no level rove were completed, followed by the first half of the profile-analysis task with a level rove. This sequence was then repeated to complete the experiment. A fixed order of presentation was selected over counterbalanced or randomized orders because the tasks progressed from simpler to more complex. The additional exposure to the stimuli via the simpler task ahead of the more complex task was intended to give listeners more familiarity with the stimuli and improve their performance on the second task. Within each task, different conditions (i.e., combinations of frequency range and component density) were always presented in randomized order. In all tasks, visual feedback indicating whether the response was correct or incorrect was provided after each trial. All stimuli were presented diotically. The experiment took each participant between four and five 2-h sessions to complete.

1. Level discrimination

In the level-discrimination task, listeners were presented with two pure tones in each trial and asked to indicate which tone was louder. The tone frequency was either 487 Hz (low-frequency condition) or 9909 Hz (high-frequency condition). The tones were 350 ms in duration, including 25-ms raised-cosine onset and offset ramps, and were separated by 350 ms of silence. On each trial, the reference or standard tone was presented at a level of 60 dB SPL and the level of the target tone was determined by the adaptive tracking procedure. Discrimination thresholds were measured using a 3-up-1-down adaptive staircase procedure tracking the 79.4% correct point on the psychometric function (Levitt, 1971). The tracking variable was the level of the signal, added in phase to the standard tone, in units of $20\log_{10}(\Delta A/A)$ (referred to as dB signal re: standard, or dB SRS), to create the target tone. The staircase began at a value of 0 dB SRS (i.e., the target tone was 6 dB higher in level than the reference tone). The initial step size of the procedure was 3 dB, and this step size was adjusted to 1.5 and 0.75 dB following the second and fourth reversals of the staircase, respectively. The procedure was terminated after 6 reversals at the smallest step size and the threshold was defined as the mean of the tracked variable values at the last six reversals. The maximum permitted value of the procedure was 13 dB SRS (so the target was about 14.8 dB higher than the reference). If the procedure called for values above 13 dB on six trials within a run, the run was terminated early and a threshold value of 13 dB was recorded for that run. Each listener completed 12 runs in each of the two frequency conditions (low and high). Initial analysis, motivated by strong training effects shown in past studies of profile analysis (Kidd *et al.*, 1986; Kidd *et al.*, 1991; Drennan and Watson, 2001a), revealed that thresholds systematically improved over the first few runs of each condition before an asymptote was reached after three to six runs. To ensure that the results focused on asymptotic performance, the first six runs were discarded and each listener's threshold was calculated as the mean value across the final six runs. The threshold-estimation procedure was completed in all runs that entered into the final threshold estimate for each listener (i.e., the procedure was never terminated early in this task).

2. Profile analysis

Listeners were presented with two consecutive inharmonic complex tones in each trial and asked to indicate which had a middle component (the signal) that was higher in level relative to the other components in the tone (the background). The amplitude of the middle component of the target was A + ΔA, where A was the amplitude of each background component and ΔA was the amplitude of the signal component added in phase to the background component of the same frequency. The components within each complex were uniformly spaced on a logarithmic frequency scale from 300 to 792 Hz (low-frequency condition) or from 6100 to 16 100 Hz (high-frequency condition). These frequency ranges were selected to match the (logarithmic) stimulus bandwidth at low and high frequencies (1.4 octaves, in both cases) and avoid frequencies above 16 kHz while minimizing the extent to which any TFS in the high-frequency stimuli would elicit a phase-locked response in the auditory nerve (which is expected to falloff steeply above 2–3 kHz; Weiss and Rose, 1988). The number of components in the complex was either 3, 5, 9, or 15. A schematic of the spectral profiles is shown in Fig. 1(c). The tones were 350 ms in duration, including 25-ms raised-cosine onset and offset ramps, and were separated by 350 ms of silence.

In runs without level randomization, the reference stimulus was presented at a level of 60 dB SPL per component, while the target stimulus consisted of the background components at a level of 60 dB SPL per component, with a signal component added at a level determined by the adaptive tracking procedure. The listener was instructed to select the louder interval as the target, with feedback presented after every trial. In runs with level randomization or roving, the level per component for the background tones in each interval was selected randomly from a uniform distribution spanning 50–70 dB SPL. The level of the signal, relative to the background tones in that interval, was determined by the adaptive tracking procedure. When the level was roved, the absolute level was no longer a reliable indicator of which tone was the target (i.e., the reference tone often had a higher overall level, and was thus louder, than the target tone). Thus, listeners were instructed to select the interval where one part of the sound was louder than the other parts of that sound, or to select the interval with a difference in quality, pitch, or timbre of the sound, again with feedback provided after every trial.

The adaptive tracking procedure was performed in the same way as described previously for level discrimination to estimate thresholds for profile analysis. Each listener completed 12 runs at each of the two frequency conditions (low and high), each of the four component conditions (3, 5, 9, and 15), and each of the two level roving conditions (no rove, 20 dB level rove), yielding a total of 192 threshold estimates per listener. As in the level discrimination task, the threshold for each listener was calculated as the mean of the final six threshold estimates in each condition, with an individual run estimate set to 13 dB SRS in cases where no threshold could be measured. Among the runs included in the final threshold estimates, no threshold could be measured (and thus a value of 13 dB SRS was used instead) in 0 of 264 low-frequency no-rove runs (0%), 71 of 264 high-frequency no-rove runs (26.9%), 4 of 249 low-frequency roved runs (1.6%), and 207 of 248 high-frequency roved runs (83.5%).

E. Experiment 1b: Level discrimination and profile analysis in TEN

Experiment 1b tested a subset of conditions from Experiment 1a in a different group of listeners using modified stimuli and procedures. Experiment 1b measured level discrimination and profile analysis in the 9-component condition only. Listeners only completed six, rather than 12, runs per condition. Stimuli were presented monaurally to the better ear based on absolute-threshold measurements from Screening Procedure B, rather than diotically. Additionally, stimuli were presented in TEN at a level of 45 dB SPL in the ERB centered at 1 kHz, rather than in quiet. The TEN was gated on 300 ms before the beginning of the first interval and was gated off 300 ms after the end of the second interval. Finally, in contrast to the fixed task order of Experiment 1a, the task order was randomized across participants. Thus, about half the participants first completed level discrimination and then profile analysis ($n = 10$) while the other half first completed profile analysis and then level discrimination ($n = 12$). Otherwise, the task, stimulus parameters, and protocol were the same as those of Experiment 1a. Unlike Experiment 1a, clear asymptotic performance may not have been fully achieved by the end of the experiment due to its shorter overall length. Nevertheless, to reduce the effects of training on the data, the first three runs of data from each condition, where the largest training effects were observed, were discarded and thresholds were calculated based on the last three runs only.

The experiment took each participant about two 2-h sessions to complete. Among the runs included in the final threshold estimates, no threshold could be measured (and thus a value of 13 dB SRS was used instead) in 1 of 66 low-frequency no-rove runs (1.5%), 33 of 66 high-frequency no-rove runs (50%), 11 of 66 low-frequency roved runs (16.7%), and 54 of 66 high-frequency roved runs (81.8%).

F. Experiment 2a: Spectrotemporal ripple detection and discrimination in quiet

Listeners who participated in this experiment always completed the ripple-detection task first, followed by the ripple-discrimination task. All stimuli were presented diotically. The experiment took each listener between one and two 2-h sessions to complete.

1. Ripple detection

Listeners were presented with three stimuli in each trial. One stimulus was a spectrotemporal ripple, synthesized by adding 300 random-phase pure tones, each with superimposed sinusoidal amplitude modulation. The carrier frequencies were spaced evenly on a logarithmic scale from 500 Hz to 18 kHz and were presented at a level of 45 dB SPL per component. The modulator of each carrier, S, was synthesized according to the equation (Chi *et al.*, 1999)

$$S(x,\, t) = 1 + m\sin(2\pi[\omega t + \Omega x] + \Phi),$$

where x indicates the carrier frequency f in units of $\log_2(f/f_0)$, where $f_0 = 500$ Hz, ω indicates the ripple rate in Hz, Ω indicates the ripple density in cycles/octave, Φ indicates the ripple phase offset in radians, and m indicates the modulation index. For the present stimuli, the ripple rate (ω) was set to 2 Hz and the ripple density (Ω) was set to four cycles/octave; these values were selected because they produce good ripple detection thresholds (Chi *et al.*, 1999). The ripple phase offset (Φ) was set randomly on each interval. The modulation index (m) was adaptively varied in the psychophysical procedure. The other two stimuli were noises generated by synthesizing ripples with $m = 0$ (i.e., no modulation). Before playback, the stimuli were bandpass filtered with 8th-order zero-phase Butterworth bandpass filters into 1.5 octave bands spanning either 600–1700 Hz (low frequency) or 6–17 kHz (high frequency), resulting in an overall level of about 64 dB SPL, which was then roved over a ±3 dB range between intervals. The stimuli were 1 s in duration (allowing for two full cycles of the modulation in each channel to take place in each stimulus), were ramped on and off with 75-ms raised-cosine ramps, and were separated by 100-ms silent interstimulus intervals. The order of the three stimuli was randomized on each trial, and listeners were tasked with identifying the modulated stimulus (i.e., the ripple), with feedback provided after each trial. Detection thresholds were estimated via a 3-down 1-up adaptive staircase procedure tracking the 79.4% correct point on the psychometric function (Levitt, 1971). The tracked variable was $20\log_{10}m$ (the modulation index in dB) and the staircase began at a value of –10 dB. The initial step size of the staircase procedure was 8 dB, and this step size was adjusted to 4 dB and 2 dB following the second and fourth reversals of the staircase, respectively. The procedure was terminated after 6 reversals at the smallest step size and the threshold was defined as the mean value of the tracked variable at the last six reversal points. The maximum permitted value of the tracked variable was 0 dB ($m = 1$). Thus, stimuli were not permitted to be overmodulated. If the staircase procedure called for values exceeding 0 dB on six trials within a given run, the run was terminated early and a threshold value of 0 dB was recorded for that run. Each listener completed six runs at each of the two frequency conditions (low and high), yielding a total of 12 threshold estimates per listener. Unlike with profile analysis, the initial analysis did not reveal any learning effects, so the final thresholds for each listener and condition were calculated as the mean threshold value from all six runs. The threshold-estimation procedure was completed in all runs that entered into the final threshold estimate for each listener (i.e., the procedure was never terminated early in this task).

2. Ripple discrimination

Listeners were presented with three stimuli in each trial. Two stimuli were ripples sweeping in one direction, while the other stimulus was a ripple sweeping in the opposite direction (achieved by inverting the sign of the ripple rate parameter ω in the equation noted previously). The modulation depth of all three stimuli was the same in each trial and was adaptively varied between trials. Listeners were tasked with identifying the odd stimulus out (i.e., which stimulus contained a ripple sweeping in the opposite direction from the other two), with feedback provided after each trial. Otherwise, the task was the same as the ripple detection task. Each listener completed six runs at each of the two frequency conditions (low and high), yielding a total of 12 threshold estimates per listener. As with ripple detection, the initial analysis did not reveal any learning effects, so the threshold for each listener and condition was calculated as the mean threshold value from all six runs. Among the runs included in the final threshold estimates, no threshold could be measured (and thus a value of 0 dB was used instead) in 11 of 72 low-frequency runs (15.3%) and 17 of 72 high-frequency runs (23.6%).

G. Experiment 2b: Spectrotemporal ripple detection and discrimination in TEN

Experiment 2b repeated the measurements from Experiment 2a in a different group of listeners using modified stimuli. Stimuli were presented monaurally to the better ear based on absolute-threshold measurements from Screening Procedure B, rather than diotically. Additionally, stimuli were presented in TEN at a level of 45 dB SPL in the ERB centered at 1 kHz, rather than in quiet. As in Experiment 1b, the TEN was gated on 300 ms before the beginning of the first interval and gated off 300 ms after the end of the second interval. Finally, the lower cutoff of the bandpass filter used in synthesizing the ripple stimuli was increased from 0.6 to 1 kHz in the low-frequency condition and from 6 to 10 kHz in the high-frequency condition, resulting in bandwidths 1–1.7 kHz and 10–17 kHz, respectively. Otherwise, parameters and procedures were the same as those of Experiment 2a. Initial analysis of the data revealed that two participants had thresholds at the ceiling (0 dB) in one or both conditions of the detection task, so their data in all conditions were excluded from figures and statistical analysis. Among the runs included in the final threshold estimates, no runs were terminated early for the detection task. For the direction-discrimination task, no threshold could be measured (and thus a value of 0 dB was used instead) in 24 of 120 low-frequency runs (20%) and 81 of 120 high-frequency runs (67.5%).

H. Experiment 3

Experiment 3 measured detection thresholds for pure tones presented in TEN to verify that TEN had the expected effect of equating masked thresholds throughout the tested frequency range (Moore *et al.*, 2000). The stimuli and protocol were identical to that of Screening Procedure A, except that listeners only completed two runs per condition (instead of three) and tone frequencies of 300, 487, and 798 Hz and 6.1, 9.9, and 16.1 kHz were tested (instead of 14 and 16 kHz). These frequencies were selected because they were approximately the target frequency, the lower edge frequency, and the upper edge frequency of the low- and high-frequency conditions of Experiments 1a and 1b.

I. Statistical analysis

The behavioral data were analyzed with linear mixed-effects models. For Experiment 1a, the data were analyzed in units of dB SRS. Fixed effects included the number of components and the frequency region (both treated as categorical variables) as well as interactions of these terms and a maximal random effects structure with respect to the listener (Barr *et al.*, 2012). Two statistical models were fit to the different subsets of the data for analysis. First, the level-discrimination data and the unroved profile-analysis data were combined and analyzed by treating level discrimination as another category of the "number of components" condition (i.e., level discrimination was treated as a special case of unroved profile analysis where the number of components was one). Next, the unroved and roved profile-analysis data were combined and analyzed jointly, with an additional factor indicating the presence or absence of a rove (and interactions between rove and other model terms). For Experiment 1b, level-discrimination and profile-analysis data were analyzed jointly. The fixed effects included frequency region, task, and their interaction, while the random effects included random intercepts and slopes for listener. For Experiment 2a, the data were analyzed in a single statistical model using units of $20\log_{10}m$. The fixed effects were the experimental task (detection vs discrimination) and frequency region, while random effects included random intercepts and slopes for listener. For Experiment 2b, the data were analyzed in the same way as in Experiment 2a. For Experiment 3, the data were analyzed in units of dB SPL. The only fixed effect was target frequency, which was coded as a categorical variable. Random effects included random intercepts and slopes for listener.

The statistical models were implemented using the R programming language and the lme4 package (Bates *et al.*, 2015) via penalized maximum likelihood estimation. Before proceeding with the analysis and interpretation of the models, diagnostic checks by visual inspection of QQ plots, scale-location plots, plots of standardized residuals versus fitted values, and plots of standardized residuals versus experimental conditions were made. Models were analyzed in two ways: first, the significance of fixed effects was examined by likelihood ratio F-tests in a type II analysis of variance (ANOVA), calculated using the Kenward-Rogers approximation for the denominator degrees of freedom via the car (Fox and Weisberg, 2019) and pbkrtest packages (Halekoh and Højsgaard, 2014); second, the significance of linear contrasts of model coefficients was examined by χ^2-tests via the phia package (Rosario-Martinez *et al.*, 2015). All p-values within a single set of tests associated with a given model were jointly corrected using the Holm-Bonferroni method. Corrected p-values are reported and compared against a criterion of $\theta = 0.05$ to assess statistical significance.

J. Auditory-nerve modeling

The auditory-nerve model of Zilany *et al.* (2014) was used to simulate average discharge rates in response to our stimuli at the level of the auditory nerve [Fig. 1(d)]. Model responses were computed using a custom wrapper in Julia and slightly modified versions of the original source code in C (modifications were restricted to removal of MATLAB-specific elements of the source code). Stimuli matching those in Experiment 1a were synthesized at a level of 60 dB SPL per component and then processed with the model. Simulated fibers were high-threshold low-spontaneous-rate fibers (Sachs and Abbas, 1974; Liberman, 1978) with "humanized" parameters, including sharper filters than are assumed for the standard cat model (Ibrahim and Bruce, 2010). An approximate implementation of power-law adaptation was used to increase simulation speed and fractional Gaussian noise was disabled in the inner-hair-cell auditory-nerve synapse stage to yield a deterministic prediction of instantaneous spike rate, ignoring sources of randomness in nerve responses (Zilany *et al.*, 2009). A response was simulated and then averaged over time to arrive at an average discharge rate for each of 301 CF values spanning logarithmically from one octave below the low-frequency target frequency to one octave above the high-frequency target frequency.

III. RESULTS

A. Experiment 1a

Thresholds for the level discrimination and profile analysis tasks are shown in Fig. 2. For the level-discrimination and unroved profile-analysis data, an ANOVA revealed significant main effects of frequency (low or high) [$F(1, 10) = 374$, $p < 0.001$] and component density (1, 3, 5, 9, or 15 components) [$F(4, 7) = 59.3$, $p < 0.001$], as well as a significant interaction between them [$F(4, 7) = 6.11$, $p = 0.019$]. For the profile analysis tasks (roved and unroved, excluding level discrimination), an ANOVA revealed significant main effects of frequency (low or high) [$F(1, 9.92) = 332$, $p < 0.001$], component density (1, 3, 5, 9, or 15 components) [$F(3, 7.89) = 29.8$, $p < 0.001$], and rove (roved or unroved) [$F(1, 9.96) = 89.5$]. Significant two-way interactions were also identified between frequency and component density [$F(3, 7.95) = 8.49$, $p = 0.007$], frequency and rove [$F(1, 9.88) = 60.8$, $p < 0.001$], and component density and rove [$F(3, 7.92) = 27.7$, $p < 0.001$]. The three-way interaction between all model terms was not significant [$F(3, 7.92) = 2.91$, $p = 0.102$].

1. Performance in level discrimination and profile analysis without roving

Conceptually, neither level discrimination nor profile analysis without level roving requires across-channel comparisons to be performed, and could be performed simply by analyzing the level of either the target component alone or, in the case of unroved profile analysis, the overall level of the complex tone. Performance in both tasks was found to be dependent on stimulus frequency: for level discrimination, thresholds were over 8 dB higher (worse) at high frequencies than at low frequencies (estimated difference $= 8.4$ dB, $\chi^2_1 = 98.6$, $p < 0.001$). This difference is roughly consistent in magnitude with prior measurements of level discrimination across the frequency range (Jesteadt *et al.*, 1977; Florentine *et al.*, 1987), and is likely due, at least in part, to the lower sensation level (SL) of the high-frequency tones and limited spread of excitation, due to the proximity of stimulation to the base of the cochlea (Florentine and Buus, 1981). Average profile-analysis thresholds without a rove were also worse at high than at low frequencies (estimated difference $= 11.1$ dB, $\chi^2_1 = 206$, $p < 0.001$); however, this difference is unlikely to be explained in terms of spread-of-excitation cues, due to the limits on detectable spread imposed by the flanking components at both low and high frequencies.

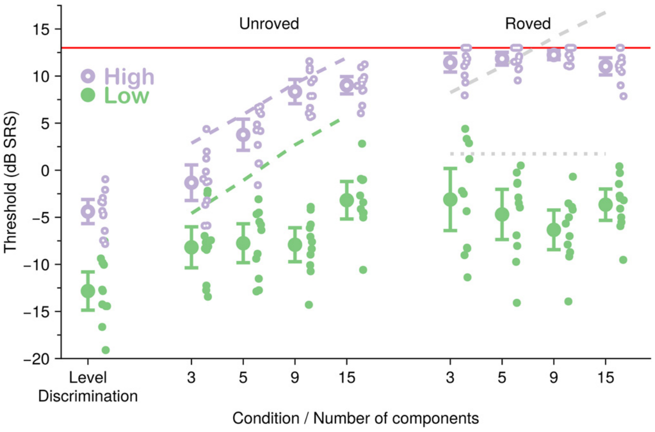


FIG. 2. (Color online) Behavioral data from Experiment 1a (level discrimination and profile analysis in quiet). Large symbols indicate mean thresholds across participants, while error bars indicate the 95% confidence intervals (± 1.96 standard error of the mean). Smaller symbols to the right of each mean show individual thresholds. The isolated symbols on the left indicate thresholds for level discrimination, while the remainder indicates thresholds for unroved (middle) and roved (right) profile analysis. Color and marker type indicate frequency condition, with low-frequency thresholds in filled green symbols and high-frequency thresholds in open purple symbols. Dashed and dotted lines indicate predicted thresholds according to different task strategies (e.g., "pick the loudest interval," see text for details). The solid red line at the top of the plot demarcates the upper limit of the adaptive procedure (13 dB SRS).

The presence of flanking components affected performance in the unroved profile analysis data, with performance being poorer for profile analysis than for level discrimination, although the influence of flankers was different at low and high frequencies. Specifically, the difference between pure-tone level-discrimination thresholds and profile-analysis thresholds was larger at high than at low frequencies for 9-component stimuli (estimated difference of differences $= 7.83$ dB, $\chi^2_1 = 20.2$, $p < 0.001$) and 15-component stimuli (estimated difference of differences $= 3.77$ dB, $\chi^2_1 = 6.33$, $p = 0.036$), but not for 3-component stimuli (estimated difference of differences $= 1.58$ dB, $\chi^2_1 = 1.00$, $p = 0.316$) or 5-component stimuli (estimated difference of differences $= 3.08$ dB, $\chi^2_1 = 2.78$, $p = 0.192$). One possible explanation for this result could be that listeners were utilizing different strategies at low and high frequencies. At low frequencies, listeners may have been better able to perceptually segregate the target component or attend selectively to channels dominated by the target component and thus achieve good performance. In contrast, at high frequencies, listeners may have simply selected the interval with a higher overall level as the target interval. The dashed sloping lines in Fig. 2 show predictions of thresholds when listeners are (1) assumed to select the interval with the highest overall sound level and (2) achieve threshold performance when the overall change in sound level between reference and target stimuli matches their average pure-tone level-discrimination thresholds. The close correspondence of these predictions to the obtained thresholds in the high-frequency condition suggests that many listeners may have adopted such a suboptimal strategy at high frequencies. This explanation would be consistent with the idea that listeners have trouble "hearing out" the target tone or attending to specific spectral channels at high frequencies (Moore and Ohgushi, 1993; Hartmann *et al.*, 1990; Moore *et al.*, 2006; Gockel and Carlyon 2018, 2022), and is also consistent with our initial hypothesis that across-channel comparisons become difficult at high frequencies (e.g., Whiteford *et al.*, 2020).

2. Effects of level roving

We hypothesized that performance in roved profile analysis would be worse at high frequencies than at low frequencies, above and beyond any baseline differences in basic intensity processing between low and high frequencies. As in the other tasks, we found that performance was indeed worse overall at high than at low frequencies (estimated difference = 16.0 dB, $\chi^2_1 = 239$, $p < 0.001$). Importantly, the difference between low and high frequencies was larger in the presence than in the absence of the rove (estimated difference of differences = 4.37 dB, $\chi^2_1 = 70.1$, $p < 0.001$). High-frequency profile-analysis thresholds were at or very close to the ceiling of the adaptive procedure (13 dB SRS) in all conditions (horizontal solid line in Fig. 2), suggesting that listeners could not perform roved profile analysis at high frequencies even with the signal level at the maximum 13 dB allowable in the adaptive procedure. As a result, our estimate of the difference between low and high frequencies is, if anything, an underestimate of the true difference.

It is worth emphasizing just how remarkably poor performance was in the high-frequency roved profile-analysis task: for a high-frequency pure tone, listeners could discriminate a change in level on the order of –4 dB SRS, corresponding to a roughly 4-dB change in the sound level of the pure tone in question. In contrast, in the profile-analysis task, listeners could not reliably discriminate a level difference on the order of 13 dB SRS, which corresponds to a nearly 15-dB change in the sound level of the target component. The flat dotted gray line presented alongside the roved profile-analysis data in Fig. 2 indicates predicted thresholds at the 79.4% correct point under the assumption that listeners always selected the interval containing the *higher target-component sound level* on each trial. This is equivalent to approximately 1/3 of the overall rove range, meaning a change in level (ΔL) at the target frequency of 7 dB, or 2 dB SRS, for all numbers of components and both frequencies. The sloping dashed gray line presented alongside the roved profile-analysis data indicates predicted thresholds at the 79.4% correct point under the assumption that listeners selected the interval with the *higher overall sound level* on each trial. Note that these predictions do not factor in the baseline level discrimination abilities of listeners, but instead indicate what performance would be expected for a theoretical system with perfect level discrimination but a suboptimal decision-making strategy. Listeners' poor profile-analysis thresholds at high frequencies, relative to these predicted thresholds, suggest that listeners may have relied on overall loudness cues to perform the task at high frequencies.

Qualitatively, component density had a non-monotonic effect on roved performance at low frequencies, consistent with prior data (Green and Mason, 1985; Bernstein and Green, 1987; Lentz *et al.*, 1999). In contrast, for the non-roved low-frequency conditions, performance remained roughly constant for component numbers between 3 and 9, rising only for 15 components, where spectral resolution is likely to play a role [see Fig. 1(d)].

B. Experiment 1b

The results from Experiment 1a suggest a major deficit in profile analysis at high frequencies. However, before attributing the deficit to the frequency range of the stimuli, it is important to rule out two potential confounds. First, although our screening procedure ensured that the stimuli were audible up to 16 kHz, the SL of the low-frequency stimuli was likely greater and more uniform across the stimulus than the SL of the high-frequency stimuli. Prior work has shown that profile-analysis thresholds vary by less than 5 dB for median SLs over the range from 30 to 70 dB SL (Mason *et al.*, 1984), but it is possible that thresholds worsen at very low SLs. Thus, the effects of frequency may be due to differences in SL, rather than frequency *per se*. Second, the order of presentation was fixed, with the first six runs of the non-roved conditions presented first; listeners may therefore have learned a strategy based on level discrimination, which they then continued to use (non-optimally) in the roved conditions.

To investigate these issues, we measured low- and high-frequency profile analysis again in a separate group of listeners using modified procedures and stimuli. First, a modified screening procedure (Screening Procedure B; see Sec. II for details) verified that all listeners had absolute thresholds at 16 kHz better than 40 dB SPL. These thresholds were measured separately for the left and right ears; stimuli were subsequently presented monaurally to the ear with the better (lower) 16-kHz threshold. Next, profile analysis was measured with and without a rove for the 9-component stimulus in the presence of TEN at a level of 45 dB SPL in the 1-kHz ERB to make the SL of the stimuli more uniform across the spectrum (both within and across the low- and high-frequency regions). Finally, the task order (i.e., order of unroved and roved tasks) was randomized across participants to examine whether exposing listeners to the unroved task first unintentionally encouraged suboptimal decision strategies for the roved task.

The results of Experiment 1b are shown in Fig. 3. As for Experiment 1a, an LME model was fit to the data from Experiment 1b and analyzed with an ANOVA to identify significant effects. The ANOVA showed a significant main effect of frequency (low or high) [$F(1, 20) = 120$, $p < 0.001$] and a significant main effect of rove (roved or unroved) [$F(1, 20) = 14.3$, $p = 0.001$]. A marginally significant interaction between order (roved first or unroved first) and rove was also observed [$F(1, 20) = 3.52$, $p = 0.075$]. Other terms did not achieve statistical significance [all $F(1, 20) < 0.504$, $p > 0.485$].

The significant main effects of frequency and rove were expected and matched the trends seen in Experiment 1a; performance was worse at high than low frequencies and thresholds were elevated on average by the rove. Importantly, a large main effect of frequency was observed

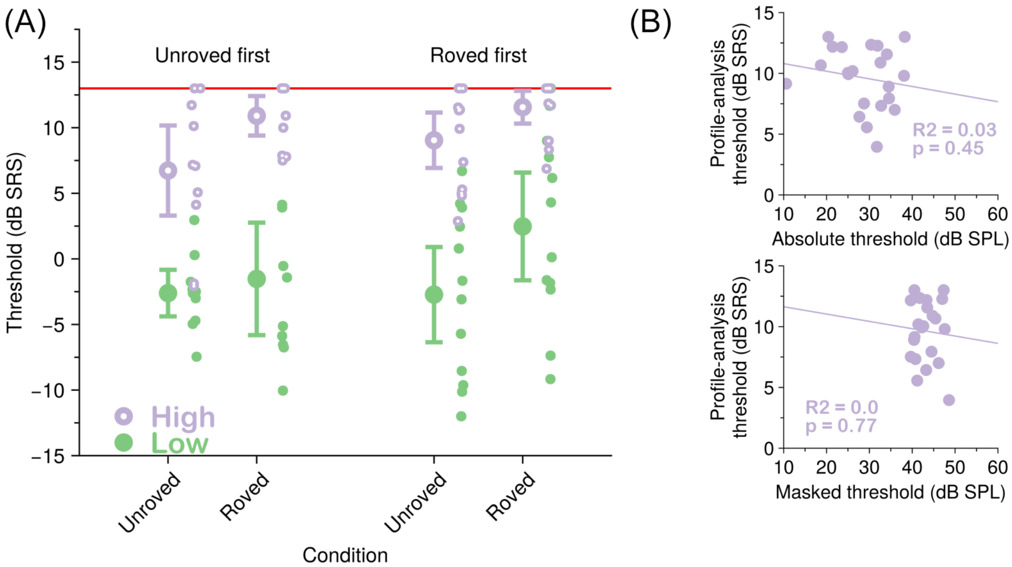


FIG. 3. (Color online) Behavioral data for Experiment 1 b (profile analysis in TEN). (A) Larger symbols indicate mean thresholds across participants, while error bars indicate the 95% confidence intervals (±1.96 standard error of the mean). Smaller symbols to the right of each mean indicate individual thresholds. Markers on the left half indicate data from listeners who completed the unroved conditions first, followed by the roved conditions. Markers on the right half indicate data from listeners who completed the roved conditions first, followed by the unroved conditions. Color and marker type indicate frequency condition, with low-frequency thresholds in filled green symbols and high-frequency thresholds in open purple symbols. The solid red line at the top of the plot demarcates the upper limit of the adaptive procedure (13 dB SRS). (B) Correlations between average high-frequency profile-analysis thresholds and absolute thresholds at 16 kHz as measured in Screening Procedure B (top) or masked thresholds in TEN at 16.1 kHz as measured in Experiment 3 (bottom). Values shown in each panel indicate the proportion of variance explained by linear regression between the two variables (R2) and the *p*-value corresponding to an uncorrected significance test of the resulting correlation coefficient via Fisher's r-to-Z transformation (*p*).

despite the equated (masked) SL in the low- and high-frequency conditions. That is, Experiment 1b demonstrates that profile analysis remains substantially better at low than at high frequencies, even when SLs are roughly equated (see Experiment 3 results). Moreover, profile-analysis thresholds in the high-frequency condition were uncorrelated with either absolute or masked thresholds near 16 kHz at the individual level [Fig. 3(b)], suggesting that audibility or effective presentation level were not major factors contributing to the observed effects of frequency. The background noise, in addition to equating SL, also may have made the task more difficult. For example, the background noise added an additional source of variance to across-channel level estimates within the stimulus passband. Moreover, the background noise provided an additional pedestal against which the target level could be judged, but in the presence of the level rove, this comparison would not have provided useful information and may have distracted listeners. Consistent with this observation, thresholds were qualitatively somewhat poorer on average in Experiment 1b than in Experiment 1a.

The lack of significant effects of task order (roved or unroved first) or its interactions in the ANOVA revealed that the order in which listeners were exposed to the profile-analysis task had little effect on their profile-analysis thresholds. Qualitatively, listeners tended to do relatively better in the task that they completed second, as might be expected based on simple learning effects. This is illustrated by the observation that listeners who completed roved profile analysis first had a larger difference between low-frequency roved and unroved thresholds than did listeners who completed the tasks in the opposite order, although this comparison was not statistically significant (estimated difference of differences = 4.11 dB, $\chi^2_1 = 3.13$, $p = 0.154$). Thus, the results of Experiment 1b indicate that the key trends in Experiment 1a were not due to listeners adopting an ineffective strategy in high-frequency profile analysis due to the fixed task order of that experiment.

C. Experiment 2a

Profile analysis is thought to require across-channel comparisons. To further test our hypothesis that across-channel comparisons are degraded at high frequencies, we employed another task that was also designed to necessitate across-channel comparisons, but that has typically been used to assess the limits of spectral or temporal resolution: spectrotemporal ripple discrimination (Chi *et al.*, 1999), in which listeners discriminate rising from falling sweeps. Our task was similar in nature to those of Denham (2005) and Archer-Boyd *et al.* (2018), with the overall modulator phase (i.e., the starting point of the sweep) randomized between intervals. This randomization makes the task difficult or impossible to perform based on information from a single frequency channel; instead, listeners must integrate information from more than one frequency channel to successfully perform the task. To rule out the possibility that listeners simply were less able to detect the modulation at high frequencies, we also measured thresholds for detecting the presence of the ripples versus unmodulated noise.

Thresholds for the ripple detection and direction discrimination tasks are shown in Fig. 4(a). An ANOVA revealed a significant main effect of task (detection vs discrimination) [$F(1,\ 11) = 34.3$, $p < 0.001$] and a significant interaction between task and frequency (high or low) [$F(1,\ 11) = 9.73$, $p = 0.01$], but the main effect of frequency was not significant [$F(1,\ 11) = 0.02$, $p = 0.89$]. Overall performance was better for the detection task than for the discrimination task (estimated difference = 7.34 dB, $\chi^2_1 = 34.3$, $p < 0.001$). That is, listeners were able to detect the presence of the ripples at a lower modulation depth than they were able to discriminate between upward and downward ripples. Overall performance, averaged across tasks, was not significantly different at low and high frequencies (estimated difference = 0.9 dB, $\chi^2_1 = 0.19$, $p = 0.89$). When considered separately, no significant differences were detected between performance at low and high frequencies for the detection task (estimated difference = 1.37 dB, $\chi^2_1 = 3.02$, $p = 0.25$) or for the discrimination task (estimated difference = –1.18 dB, $\chi^2_1 = 2.47$, $p = 0.25$). However, an interaction contrast comparing the difference in performance between low and high frequencies in the detection task and the discrimination task was significant (estimated difference of differences = 2.55 dB, $\chi^2_1 = 9.73$, $p = 0.007$). In other words, the worsening of performance in the discrimination task relative to the detection task was larger at high than at low frequencies by approximately 2.5 dB. This difference can be seen more clearly in Fig. 4(b), which shows the difference in thresholds between the two ripple tasks at the group-average and individual levels. This difference was larger on average at high frequencies, and all but one participant in the experiment exhibited this pattern at the individual level.

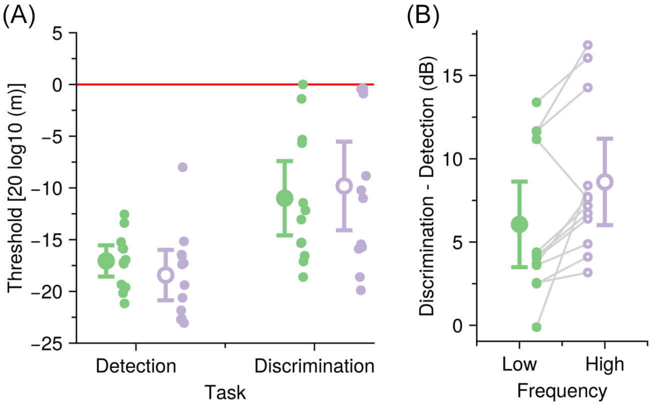


FIG. 4. (Color online) Behavioral data from Experiment 2a. (A) Larger symbols indicate mean thresholds across participants, while error bars indicate the 95% confidence intervals (± 1.96 standard error of the mean). Smaller symbols to the right of each mean indicate individual thresholds. Color and marker type indicate frequency condition, with low-frequency thresholds shown in filled green symbols and high-frequency thresholds shown in open purple symbols. (B) Differences between thresholds in the discrimination task and the detection task. Markers indicate the same as in (A). Threshold differences from the same listener in the low- and high-frequency conditions are linked by light gray lines.

D. Experiment 2b

In the context of our original hypothesis, the differences in the magnitudes of the frequency effects observed in profile analysis (Experiments 1a and 1b) and observed for spectrotemporal ripples (Experiment 2a) could be interpreted to suggest that across-channel integration is less necessary for spectrotemporal ripple discrimination than for profile analysis. However, a key difference between the two tasks must first be addressed. Namely, in profile analysis, across-channel comparisons must include the target channel, or some channel near it, that contains information about the level increment. In the high-frequency conditions, that information is presented at the stimulus center frequency near 9.9 kHz. In ripple discrimination, however, the task could theoretically be performed based on comparisons of nearly any two channels. In Experiment 2a, for example, listeners could have relied on frequencies near the lower edge of the stimulus passband, such as 6 and 6.5 kHz, instead of frequencies closer to the 9.9-kHz region necessary for the profile-analysis task. This crucial difference between the two tasks may explain why one (profile analysis) exhibited a larger effect of frequency region than the other (ripple discrimination).

To test this potential explanation, we measured low- and high-frequency ripple detection and discrimination again in a separate group of listeners using modified procedures and stimuli. First, the same modified screening procedure used for Experiment 1b (Screening Procedure B; see Sec. II for details) verified that all listeners had absolute thresholds at 16 kHz better than 40 dB SPL. These thresholds were measured separately for the left and right ears; stimuli were subsequently presented monaurally to the ear with the lower (better) threshold. Next, ripple detection and discrimination thresholds were measured as before, but with the stimuli presented in TEN at a level of 45 dB SPL in the 1-kHz ERB and with the lower cutoff of the stimulus passband adjusted to either 1 kHz (low-frequency condition) or 10 kHz (high-frequency condition).

The results of Experiment 2b are presented in Fig. 5. As for Experiment 2a, an LME model was fit to the data and analyzed with an ANOVA to assess statistical significance. As in Experiment 2a, the ANOVA revealed significant main effects of frequency (low or high) [$F(1,\ 21.5) = 23.4$, $p < 0.001$] and task (detection or discrimination) [$F(1,\ 21.5) = 249$, $p < 0.001$], as well as a significant two-way interaction between frequency and task [$F(1,\ 21.8) = 16.8$, $p < 0.001$].

Thresholds were somewhat poorer overall in Experiment 2b than in 2a, likely due to the presence of the background TEN and to the increase in the lower cutoff frequency of the stimulus and the consequent reduction of the stimulus passband. Discrimination thresholds were higher than detection thresholds, with an effect size comparable to Experiment 2a (estimated difference = 7.70 dB, $\chi^2_1 = 243$, $p < 0.001$). In the detection task, differences between low- and high-frequency thresholds did not achieve statistical

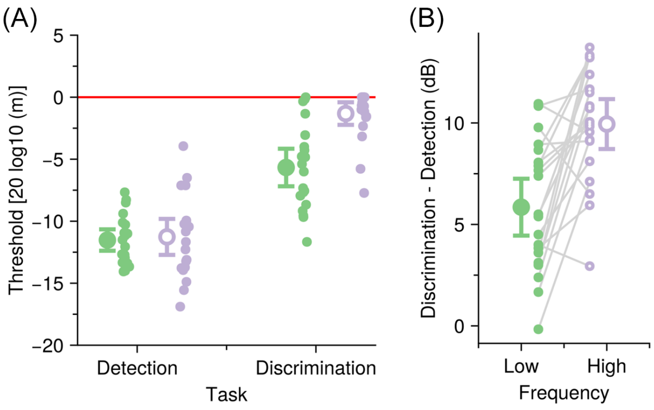


FIG. 5. (Color online) Behavioral data from Experiment 2b. (A) Larger symbols indicate mean thresholds across participants, while error bars indicate the 95% confidence intervals (± 1.96 standard error of the mean). Smaller symbols to the right of each mean indicate individual thresholds. Color and marker type indicate frequency condition, with low-frequency thresholds shown in filled green symbols and high-frequency thresholds shown in open purple symbols. (B) Differences between thresholds in the discrimination task and the detection task. Markers indicate the same as in (A). Threshold differences from the same listener in the low- and high-frequency conditions are linked by light gray lines.

significance (estimated difference $= 0.40$ dB, $\chi^2_1 = 0.396$, $p = 0.529$). In the discrimination task, however, high-frequency thresholds were significantly higher than low-frequency thresholds (estimated difference $= 4.31$ dB, $\chi^2_1 = 39.3$, $p < 0.001$). The increase in discrimination thresholds relative to detection thresholds was greater at high frequencies than at low frequencies (estimated difference of differences $= 3.91$ dB, $\chi^2_1 = 16.8$, $p < 0.001$).

The significant difference between low- and high-frequency discrimination thresholds, and the greater increase in discrimination thresholds relative to detection thresholds at high frequencies, are consistent with our initial hypothesis. Qualitatively, these effects were larger than those observed in Experiment 2a, suggesting that the differences in the stimulus, such as the more restrictive stimulus passband, accentuated the high-frequency deficit in the discrimination task. It is also worth noting that, as in Experiment 2a, the individual data are quite consistent, with almost all listeners showing a larger discrimination vs detection difference at high frequencies. Many high-frequency discrimination thresholds were at or near the ceiling value of 0-dB modulation depth, suggesting that, if anything, the present results may underestimate the true differences between low and high frequencies, as was also found for profile analysis.

E. Experiment 3

Experiments 1b and 2b used TEN to equate SL for the low- and high-frequency stimuli, but the original validation of TEN was only conducted for frequencies up to 10 kHz (Moore *et al.*, 2000). To verify that the TEN indeed achieved its intended purpose in our experiments up to 16 kHz, we measured masked detection thresholds for pure tones in TEN with a level of 45 dB SPL in the ERB centered at 1 kHz for the edge and center frequencies of the profile-analysis stimuli (300, 487, and 798 Hz for low-frequency stimuli, and 6.1, 9.9, and 16.1 kHz for high-frequency stimuli). The results of this experiment are shown in Fig. 6.

As for the other experiments, an LME model was fit to the data and analyzed with an ANOVA to assess statistical significance. The ANOVA revealed a significant main effect of frequency [$F(1, 16.8) = 6.07$, $p = 0.002$]. Average thresholds were within 0.5 dB of 41 dB SPL at all of the lower test frequencies, indicating that listeners could detect pure tones at levels about 4 dB lower than the level of the TEN in the ERB centered at 1 kHz (45 dB SPL) across the low-frequency stimulus range. Qualitatively, results were similar at 6.1 kHz, but some listeners had higher thresholds at 9.9 and 16.1 kHz. To assess the significance of these changes, we performed *post hoc* tests on the significance of each high-frequency condition's mean, relative to the average low-frequency threshold of approximately 41 dB SPL. This analysis revealed no significant differences at 6.1 kHz (estimated difference $= 0.44$ dB, $\chi^2_1 = 2.09$, $p = 0.148$), but thresholds were significantly elevated at 9.9 kHz (estimated difference $= 1.75$ dB, $\chi^2_1 = 6.38$, $p = 0.023$) and at 16.1 kHz (estimated difference $= 1.89$ dB, $\chi^2_1 = 10.2$, $p = 0.004$).

These results indicate that masked SL was equated in Experiments 1b and 2b across the entire low-frequency range. In the high-frequency range, SL was equated near the low-frequency edge of the stimulus, but slightly poorer at the target frequency of 9.9 kHz and up to the high-frequency edge of the stimulus. However, although these differences in SL at high frequencies were significant, they were less than 2 dB on average, making them unlikely to have strongly affected performance in a task that included a 20-dB level rove range.

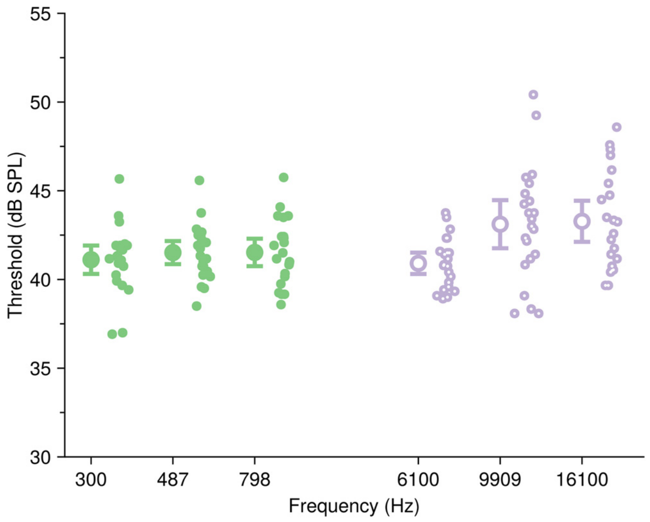


FIG. 6. (Color online) Behavioral data from Experiment 3. Masked thresholds in TEN as a function of frequency, with color and marker type indicating whether the frequency corresponds to the low-frequency condition (filled green symbols) or high-frequency condition (open purple symbols) of Experiments 1a and 1b. Larger points indicate mean thresholds across participants, while error bars indicate the 95% confidence intervals (± 1.96 standard error of the mean). Smaller points indicate individual-participant thresholds.

IV. DISCUSSION

The results demonstrate that listeners are remarkably poor at making certain types of judgments about complex sounds in the high- to extended-high frequency range (above 8 kHz). In profile analysis, when listeners were tasked with detecting an increase in the amplitude of one component in an inharmonic complex while ignoring random changes in overall sound level, these deficits were pronounced (Experiment 1a), even when stimuli were presented in TEN to equate SL across and within the two frequency regions (Experiments 1b and 3). This high-frequency deficit was larger than would be expected from the listeners' poorer pure-tone intensity discrimination thresholds at high relative to low frequencies. In ripple discrimination, when listeners were tasked with differentiating upward from downward spectrotemporal ripples while ignoring random changes in the starting modulator phase, a small but significant high-frequency deficit initially observed (Experiment 2a) became considerably more substantial when the spectral range was limited to frequencies above 9.9 kHz (Experiment 2b). As both profile analysis and ripple discrimination were designed to necessitate across-frequency comparisons to achieve good performance, the present results are consistent with our initial hypothesis that the ability to perform across-channel comparisons is impaired at high frequencies, similar to prior observations about comparing the phase of slow amplitude modulations across channels (Whiteford *et al.*, 2020).

Some alternative explanations of our results warrant consideration. One possible explanation concerns phase locking to TFS at the level of the auditory nerve, which degrades at high frequencies (Joris and Verschooten, 2013; Palmer and Russell, 1986; Verschooten *et al.*, 2019). It has been suggested that phase-locked (as opposed to spectral or rate-place) encoding of TFS is required for "hearing out" a component of interest within a complex (Moore and Ohgushi, 1993; Moore *et al.*, 2006; Gockel and Carlyon 2018, 2022), an ability that in turn has been connected to profile analysis by some investigators (Gockel and Colonius, 1997; Gockel, 1998). If these connections hold, they could explain a selective high-frequency deficit in profile analysis. The issue of whether it is necessary for listeners to "hear out" the target component in profile analysis, however, remains unclear. Prior work has shown that perceptual segregation of the target from the background via onset asynchrony cues can impair profile analysis (Green and Dai, 1992; Hill and Bailey, 1997), while making the target component easier to hear out by mistuning it (in a harmonic profile-analysis task) does not (Hill and Bailey, 2000). Clearly, more work is needed to determine the precise role of perceptual segregation in profile analysis.

It has also been suggested that phase-locked encoding of TFS underlies complex pitch perception (Cariani and Delgutte, 1996; Meddis and O'Mard, 1997). Listeners in profile analysis experiments often report pitch and timbre changes associated with the level increment (e.g., Gockel, 1998; Hill and Bailey, 2000), and various experimenters have invoked changes in pitch strength and differences in pitch strength between harmonic and inharmonic complexes to explain certain effects in profile analysis (Gockel, 1998; Drennan and Watson, 2001b). Thus, the loss of phase-locked TFS coding at high frequencies could explain why profile-analysis thresholds worsen at high frequencies. Richards *et al.* (1989) and Kidd *et al.* (1991), based on comparisons between psychophysical data and predictions of the envelope-weighted average instantaneous frequency model (Feth, 1974), argued against a role for pitch judgments in profile analysis. These prior computational analyses, however, are limited by the use of a particular pitch model that calculates shifts in pitch height of the overall complex tone, rather than the relative pitch strength or salience of the target component.

A recent study has suggested a different neural mechanism for profile analysis in the form of modulation-sensitive neurons in the auditory midbrain (Maxwell *et al.*, 2020). Profile-analysis stimuli evoke fluctuations in auditory-nerve responses due to peripheral interactions between nearby frequency components, and when the target component is incremented the fluctuations are attenuated in channels near the target frequency. As profile-analysis stimuli are shifted up to higher frequencies, the rates of modulations elicited by peripheral interaction between components likewise shift upward. These modulations could be decoded by modulation-sensitive cells in the auditory midbrain, but only over a limited range of modulation frequencies (typically less than ∼300–500 Hz in neurons in the auditory midbrain and thalamus; Kim *et al.*, 2020). This frequency effect could explain the present results. However, the model of Maxwell *et al.* (2020) has so far only been tested on the stimulus of Lentz *et al.* (1999), which used an alternating pattern of amplified and attenuated components and thus contained rich increment-dependent modulation cues over the entire frequency range of the stimulus. It remains to be seen whether their framework would generalize to more typical profile-analysis stimuli, such as those used in the present study. It is also worth noting that a framework based on detecting envelope fluctuations would likely only be sensitive in conditions where component spacing is limited to frequency differences below a few hundred Hertz, where fluctuations are detectable (Kohlrausch *et al.*, 2000), making it unlikely to provide a general account for the broad range of profile-analysis data that includes very widely spaced components.

The main finding of severe deficits in processing across-frequency cues at high frequencies suggests that there may be more to learn from many classic auditory paradigms by extending them to this extended high-frequency range and testing whether current theories can generalize accordingly. The present results have interesting implications for some current topics in auditory perception, such as the debates about the use of TFS in various tasks (Verschooten *et al.*, 2019) or perceptual implications of frequency-place mismatches in cochlear-implant users (Fu and Shannon, 1999a,b; Baskent and Shannon, 2003; Xu *et al.*, 2020), which are concerned with perception of lower-frequency stimuli that are presented to tonotopic regions normally sensitive to high frequencies.

The present results may be explained by posulating that accurate multichannel intensity processing is not needed to fulfill the ecological roles associated with high-frequency hearing (Hunter *et al.*, 2020). For instance, the fine-grained spectral information useful for speech recognition, such as the first few formant frequencies, is limited to frequencies below 6 kHz (Hillenbrand *et al.*, 1995), whereas high-frequency cues that extend beyond 6 kHz and may help distinguish different fricatives are more broadband in nature (Maniwa *et al.*, 2009; Monson *et al.*, 2012). Other spectral cues, such as those contained in head-related transfer functions at high frequencies, can be smoothed considerably without affecting perception (Kulkarni and Colburn, 1998), again suggesting a lack of sensitivity to high-frequency spectral profile cues. Future work could shed light on these issues by seeking out other behavioral tasks that show similar high-frequency deficits, by searching for potential differences in the neural implementations of across-channel comparisons in low- and high-frequency auditory areas, and by further investigating the ecological relevance of high-frequency hearing. These questions could also be addressed by studying whether individual differences in high-frequency spectral analysis abilities are consistent across different tasks, potentially reflecting a common underlying mechanism.

V. CONCLUSIONS

- Profile-analysis thresholds were worse at high than at low frequencies, beyond what would be expected from baseline differences in pure-tone intensity discrimination at low and high frequencies.
- High-frequency deficits in profile analysis were observed both when stimuli were presented at a fixed sound level and when sound levels were roved over a 20-dB range, in quiet and when stimuli were presented in TEN to equate SL at low and high frequencies.
- Spectrotemporal ripple discrimination was slightly worse at high than at low frequencies when high-frequency stimuli contained energy down to 6 kHz, but was substantially worse than at low frequencies when the high-frequency stimuli were restricted to frequencies of 10 kHz and above.
- These results are consistent with the hypothesis that comparing intensity information across frequency is more difficult at high than at low frequencies in ways that cannot be accounted for by cochlear spectral resolution.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R01 DC005216 (A.J.O.), by NIH fellowship F31 DC019247 (D.R.G.), and by NSF training grant NRT-UtB1734815 (A.J.O.). We thank Hedwig Gockel and two anonymous reviewers for many useful comments on earlier versions of this manuscript.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

Data and code used here are available at https://github.com/guestdaniel/GuestRajappaOxenham2024.

Archer-Boyd, A. W., Southwell, R. V., Deeks, J. M., Turner, R. E., and Carlyon, R. P. (**2018**). “Development and validation of a specto-temporal processing test for cochlear-implant listeners,” J. Acoust. Soc. Am. **144**, 2983–2997.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (**2012**). “Random effects structure for confirmatory hypothesis testing: Keep it maximal,” J. Mem. Lang. **68**, 255–278.

Baskent, D., and Shannon, R. V. (**2003**). “Speech recognition under conditions of frequency-place compression and expansion,” J. Acoust. Soc. Am. **113**, 2064–2076.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (**2015**). “Fitting linear mixed-effects models using lme4,” J. Stat. Softw. **67**(1), 1–48.

Bernstein, L. R., and Green, D. M. (**1987**). “The profile-analysis bandwidth,” J. Acoust. Soc. Am. **81**, 1888–1895.

Bernstein, L. R., and Green, D. M. (**1988**). “Detection of changes in spectral shape: Uniform vs. non-uniform background spectra,” Hear. Res. **34**, 157–165.

Cariani, P. A., and Delgutte, B. (**1996**). “Neural correlates of the pitch of complex tones. I. Pitch and pitch salience,” J. Neurophysiol. **76**, 1698–1716.

Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (**1999**). “Spectro-temporal modulation transfer functions and speech intelligibility,” J. Acoust. Soc. Am. **106**, 2719–2732.

Denham, S. L. (**2005**). “Perception of the direction of frequency sweeps in moving ripple noise stimuli,” in *Plasticity and Signal Representation in the Auditory System*, edited by J. Syka and M. M. Merzenich (Springer, New York) pp. 317–321.

Drennan, W. R., and Watson, C. S. (**2001a**). “Sources of variation in profile analysis. I. Individual differences and extended training,” J. Acoust. Soc. Am. **110**, 2491–2497.

Drennan, W. R., and Watson, C. S. (**2001b**). “Sources of variation in profile analysis. II. Component spacing, dynamic changes, and roving level,” J. Acoust. Soc. Am. **110**, 2498–2504.

Ewert, S. D. (**2013**). “AFC—A modular framework for running psychoacoustic experiments and computational perception models,” in *Proceedings of the International Conference on Acoustics AIA-DAGA*, March 18–21, Merano, Italy, pp. 1326–1329.

Feth, L. L. (**1974**). “Frequency discrimination of complex periodic tones,” Percept. Psychophys. **15**, 375–378.

Florentine, M., and Buus, S. (**1981**). “An excitation-pattern model for intensity discrimination,” J. Acoust. Soc. Am. **70**, 1646–1654.

Florentine, M., Buus, S., and Mason, C. R. (**1987**). “Level discrimination as a function of level for tones from 0.25 to 16 kHz,” J. Acoust. Soc. Am. **81**, 1528–1541.

Fox, J., and Weisberg, S. (**2019**). *An R Companion To Applied Regression* (Sage, Thousand Oaks, CA).

Fu, Q.-J., and Shannon, R. V. (**1999a**). “Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant,” Ear Hear. **20**, 332–344.

Fu, Q.-J., and Shannon, R. V. (**1999b**). “Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing,” J. Acoust. Soc. Am. **105**, 1889–1900.

Halekoh, U., and Højsgaard, S. (**2014**). “A Kenward-Rogers approximation and parametric bootstrap methods for tests in linear mixed models—The R Package pbkrtest,” J. Stat. Softw. **59**, 1–30.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). “Acoustic characteristics of American English vowels,” J. Acoust. Soc. Am. **97**, 3099–3111.

Glasberg, B. R., and Moore, B. C. J. (**1990**). “Derivation of auditory filter shapes from notched-noise data,” Hear. Res. **47**, 103–138.

Gockel, H. E. (**1998**). "On possible cues in profile analysis: Identification of the incremented component," J. Acoust. Soc. Am. **103**, 542–552.

Gockel, H. E., and Carlyon, R. P. (**2018**). "Detection of mistuning in harmonic complex tones at high frequencies," Acta Acust. **104**, 766–769.

Gockel, H. E., and Carlyon, R. P. (**2022**). "On mistuning detection and beat perception for harmonic complex tones at low and very high frequencies," J. Acoust. Soc. Am. **152**, 226–239.

Gockel, H. E., and Colonius, H. (**1997**). "Auditory profile analysis: Is there perceptual constancy for spectral shape for stimuli roved in frequency?," J. Acoust. Soc. Am. **102**, 2311–2315.

Green, D. M. (**1983**). "Profile analysis: A different view of auditory intensity discrimination," Am. Psychol. **38**, 133–142.

Green, D. M., and Dai, H. (**1992**). "Temporal relations in profile comparisons," in *Auditory Physiology and Perception: Proceedings of the 9th International Symposium on Hearing*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford, UK), pp. 471–478.

Green, D. M., and Kidd, G. (**1983**). "Further studies of auditory profile analysis," J. Acoust. Soc. Am. **73**, 1260–1265.

Green, D. M., Kidd, G., and Picardi, M. C. (**1983**). "Successive versus simultaneous comparison in auditory intensity discrimination," J. Acoust. Soc. Am. **73**, 639–643.

Green, D. M., and Mason, C. R. (**1985**). "Auditory profile analysis: Frequency, phase, and Weber's law," J. Acoust. Soc. Am. **77**, 1155–1161.

Guest, D. R., and Oxenham, A. J. (**2022**). "Human discrimination and modeling of high-frequency complex tones shed light on the neural codes for pitch," PLoS Comput. Biol. **18**, e1009889.

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). "Hearing a mistuned harmonic in an otherwise periodic complex tone," J. Acoust. Soc. Am. **88**, 1712–1724.

Hill, N. I., and Bailey, P. J. (**1997**). "Profile analysis with an asynchronous target: Evidence for auditory grouping," J. Acoust. Soc. Am. **102**, 477–481.

Hill, N. I., and Bailey, P. J. (**2000**). "Profile analysis of harmonic complexes: Effects of mistuning the target," J. Acoust. Soc. Am. **107**, 2291–2294.

Hunter, L. L., Monson, B. B., Moore, D. R., Dhar, S., Wright, B. A., Munro, K. J., Zadeh, L. M., Blankenship, C., Stiepan, S., and Siegel, J. H. (**2020**). "Extended high frequency hearing and speech perception implications in adults and children," Hear. Res. **397**, 107922.

Ibrahim, R. A., and Bruce, I. C. (**2010**). "Effects of peripheral tuning on the auditory nerve's representation of speech envelope and temporal fine structure cues," in *The Neurophysiological Bases of Auditory Perception*, edited by E. Lopez-Poveda, A. Palmer, and R. Meddis (Springer, New York).

Jesteadt, W., Wier, C. C., and Green, D. M. (**1977**). "Intensity discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 169–177.

Joris, P. X., and Verschooten, E. (**2013**). "On the limit of neural phase locking to fine structure in humans," in *Basic Aspects of Hearing*, edited by B. C. J. Moore (Springer, New York), pp. 101–108.

Kidd, G., Mason, C. R., and Green, D. M. (**1986**). "Auditory profile analysis of irregular sound spectra," J. Acoust. Soc. Am. **79**, 1045–1053.

Kidd, G., Mason, C. R., Uchanski, R. M., Brantley, M. A., and Shah, P. (**1991**). "Evaluation of simple models of auditory profile analysis using random reference spectra," J. Acoust. Soc. Am. **90**, 1340–1354.

Kim, D. O., Carney, L., and Kuwada, S. (**2020**). "Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body," J. Neurophysiol. **124**, 1198–1215.

Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723–734.

Kulkarni, A., and Colburn, H. S. (**1998**). "Role of spectral detail in sound-source localization," Nature **396**, 747–749.

Lentz, J. J., Richards, V. M., and Matiasek, M. R. (**1999**). "Different auditory filter bandwidth estimates based on profile analysis, notched noise, and hybrid tasks," J. Acoust. Soc. Am. **106**, 2779–2792.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442–455.

Maniwa, K., Jongman, A., and Wade, T. (**2009**). "Acoustic characteristics of clearly spoken English fricatives," J. Acoust. Soc. Am. **125**, 3962–3973.

Marks, L. E. (**1994**). " 'Recalibrating' the auditory system: The perception of loudness," J. Exp. Psychol. **20**, 382–396.

Mason, C. R., Kidd, G., Hanna, T. E., and Green, D. M. (**1984**). "Profile analysis and level variation," Hear. Res. **13**, 269–275.

Maxwell, B. N., Richards, V. M., and Carney, L. H. (**2020**). "Neural fluctuation cues for simultaneous notched-noise masking and profile-analysis tasks: Insights from model midbrain responses," J. Acoust. Soc. Am. **147**, 3523–3537.

Meddis, R., and O'Mard, L. (**1997**). "A unitary model of pitch perception," J. Acoust. Soc. Am. **102**, 1811–1820.

Monson, B. B., Lotto, A. J., and Story, B. H. (**2012**). "Analysis of high-frequency energy in long-term average spectra of singing, speech, and voiceless fricatives," J. Acoust. Soc. Am. **132**, 1754–1764.

Moore, B. C. J., Glasberg, B. R., Low, K. E., Cope, T., and Cope, W. (**2006**). "Effects of level and frequency on the audibility of partials in inharmonic complex tones," J. Acoust. Soc. Am. **120**, 934–944.

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántra, J. I. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.

Moore, B. C. J., and Ohgushi, K. (**1993**). "Audibility of partials in inharmonic complex tones," J. Acoust. Soc. Am. **93**, 452–461.

Oxenham, A. J., and Buus, S. (**2000**). "Level discrimination of sinusoids as a function of duration and level for fixed-level, roving-level, and across-frequency conditions," J. Acoust. Soc. Am. **107**, 1605–1614.

Oxenham, A. J., and Shera, C. A. (**2003**). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," J. Assoc. Res. Otolaryngol. **4**, 541–554.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hear. Res. **24**, 1–15.

Richards, V. M., Onsan, Z. A., and Green, D. M. (**1989**). "Auditory profile analysis: Potential pitch cues," Hear. Res. **39**, 27–36.

Rosario-Martinez, H. D., Fox, J., R Core Team, and Chalmers, P. (**2015**). "phia: Post-hoc interaction analysis," https://cran.r-project.org/package=phia (Last viewed 5/3/2024).

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769–793.

Sachs, M. B., and Abbas, P. J. (**1974**). "Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli," J. Acoust. Soc. Am. **56**, 1835–1847.

Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). "Revised estimates of cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. U.S.A. **99**, 3318–3323.

Spiegel, M. F., Picardi, M. C., and Green, D. M. (**1981**). "Signal and masker uncertainty in intensity discrimination," J. Acoust. Soc. Am. **70**, 1015–1019.

Verschooten, E., Shamma, S., Oxenham, A. J., Moore, B. C. J., Joris, P. X., Heinz, M. G., and Plack, C. J. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hear. Res. **377**, 109–121.

Weiss, T. F., and Rose, C. (**1988**). "A comparison of synchronization filters in different auditory receptor organs," Hear. Res. **33**, 175–180.

Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (**2020**). "The role of cochlear place coding in the perception of frequency modulation," eLife **9**, e58468.

Xu, K., Willis, S., Gopen, Q., and Fu, Q.-J. (**2020**). "Effects of spectral resolution and frequency mismatch on speech understanding and spatial release from masking in simulated bilateral cochlear implants," Ear Hear. **41**, 1362–1371.

Zera, J., Onsan, Z. A., Nguyen, Q. T., and Green, D. M. (**1993**). "Auditory profile analysis of harmonic signals," J. Acoust. Soc. Am. **93**, 3431–3441.

Zhou, B. (**1995**). "Auditory filter shapes at high frequencies," J. Acoust. Soc. Am. **98**, 1935–1942.

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," J. Acoust. Soc. Am. **135**, 283–286.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," J. Acoust. Soc. Am. **126**, 2390–2412.