bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title: Perception of frequency modulation is mediated by cochlear place coding

Authors: Kelly L. Whiteford*[a], Heather A. Kreft[a], and Andrew J. Oxenham[a]

[a]Department of Psychology, University of Minnesota, Minneapolis, MN 55455

*Corresponding author: Department of Psychology, 75 East River Parkway, Minneapolis, MN 55455; Phone: 612-626-3258; Email: whit1945@umn.edu

Keywords: Place coding, time coding, frequency modulation, sensorineural hearing loss

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Natural sounds convey information via frequency and amplitude modulations (FM and AM). Humans are acutely sensitive to the slow rates of FM that are crucial for speech and music. Two coding mechanisms are believed to underlie FM sensitivity, one based on precise stimulus-driven spike timing (time code) for slow FM rates, and another coarser code based on cochlear place of stimulation (place code) for fast FM rates. We tested this long-standing explanation by studying individual differences in listeners with varying degrees of hearing loss that resulted in widely varying fidelity of place-based or tonotopic coding. Our findings reveal that FM detection at both slow and fast rates is closely related to the fidelity of place coding in the cochlea, suggesting a unitary neural code for all FM rates. These insights into the initial coding of important sound features provide a new impetus for improving place coding in auditory prostheses.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Modulations in frequency (FM) and amplitude (AM) carry critical information in biologically relevant sounds, such as speech, music, and animal vocalizations (Attias and Schreiner, 1997; Nelken et al., 1999). In humans, AM is crucial for understanding speech in quiet (Shannon et al., 1995; Smith et al., 2002), while FM is particularly important for perceiving melodies, recognizing talkers, determining speech prosody and emotion, and segregating speech from other competing background sounds (Zeng et al., 2005; Strelcyk and Dau, 2009; Sheft et al., 2012). The perception of FM is often degraded in older listeners and people with hearing loss (Lacher-Fougère and Demany, 1998; Moore and Skrodzka, 2002; He et al., 2007; Strelcyk and Dau, 2009; Grose and Mamo, 2012; Paraouty et al., 2016; Wallaert et al., 2016; Paraouty and Lorenzi, 2017; Whiteford et al., 2017). This deficit likely contributes to the communication difficulties experienced by such listeners in noisy real-world environments, which may in turn help explain why age-related hearing loss has been associated with decreased social engagement, greater rates of cognitive decline, and increased risk of dementia (Lin et al., 2011, 2013; Lin and Albert, 2014; Deal et al., 2017; Thomson et al., 2017). Current assistive listening devices, such as hearing aids and cochlear implants, have been generally unsuccessful at reintroducing viable FM cues to the auditory system (Chen and Zeng, 2004; Ives et al., 2013). This lack of success is partly related to a gap in our scientific understanding regarding how FM is extracted by the brain from the information available in the auditory periphery.

The coding of AM begins in the auditory nerve with periodic increases and decreases in the instantaneous firing rate of auditory nerve fibers that mirror the fluctuations in the temporal envelope of the stimulus (Schreiner and Langner, 1988; Joris et al., 2004). As early as the inferior colliculus and extending to the auditory cortex, rapid AM rates are transformed to a code involving firing rate that is no longer time-locked to the stimulus envelope and instead relies on

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

overall firing rate, with different neurons displaying bandpass, lowpass, or highpass responses to different AM rates (Schreiner and Langner, 1988; Wang et al., 2008). The coding of FM is less straightforward. For a pure tone with FM, the temporal envelope of the stimulus is flat; however, the changes in frequency lead to dynamic shifts in the tone's tonotopic representation along the basilar membrane, resulting in a transformation of FM into AM at the level of the auditory nerve (Zwicker, 1956; Moore and Sek, 1995; Saberi and Hafter, 1995; Sek and Moore, 1995).

Although this FM-to-AM conversion provides a unified and neurally efficient code for both AM and FM based on periodic fluctuations in the instantaneous auditory-nerve firing rate in both cases (Saberi and Hafter, 1995), it falls short of explaining human behavioral trends in FM sensitivity, specifically at low carrier frequencies (f_c < ~4-5 kHz) and slow modulation rates (f_m <~ 10 Hz), where sensitivity tends to be considerably better than at higher carrier frequencies or fast modulation rates (Demany and Semal, 1989; Moore and Sek, 1995; Sek and Moore, 1995; Moore and Sek, 1996; He et al., 2007; Whiteford and Oxenham, 2015; Whiteford et al., 2017). This discrepancy is important, because low frequencies and slow modulation rates are the most important for human communication, including speech and music, as well as animal vocalizations. The enhanced sensitivity to slow FM at low carrier frequencies has been explained in terms of an additional neural code based on stimulus-driven spike timing in the auditory nerve that is phase-locked to the temporal fine structure of the stimulus (Moore and Sek, 1995). Although such a time-based code can potentially provide greater accuracy (Siebert, 1970; Heinz et al., 2001), and is used for spatial localization (Moiseff and Konishi, 1981; Grothe et al., 2010), it is not known whether or how this timing information is extracted by higher stages of the auditory system to code periodicity and FM.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

If the detection of FM at fast modulation rates depends on an FM-to-AM conversion, whereas the detection of FM at slow rates does not, then fast-rate FM detection thresholds should depend on the sharpness of cochlear tuning (Fig. 1), whereas slow-rate FM detection thresholds should not. Previous studies using normal-hearing listeners have not demonstrated such a relationship for either slow or fast FM rates (Whiteford and Oxenham, 2015; Whiteford et al., 2017). However, this failure to find a correlation may be due to lack of variability in cochlear filtering within a normal-hearing population. People with cochlear hearing loss often have poorer frequency selectivity (Glasberg and Moore, 1986; Moore et al., 1999), due to a broadening of cochlear tuning (Robertson and Manley, 1974; Liberman et al., 1986; Moore, 2007). In contrast, damage to the cochlea is not thought to lead to a degradation of auditory-nerve phase locking to temporal fine structure for sounds presented in quiet (Henry and Heinz, 2012), so we would not expect to find a strong relationship between slow-rate FM detection thresholds and hearing-loss-induced changes in cochlear tuning.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

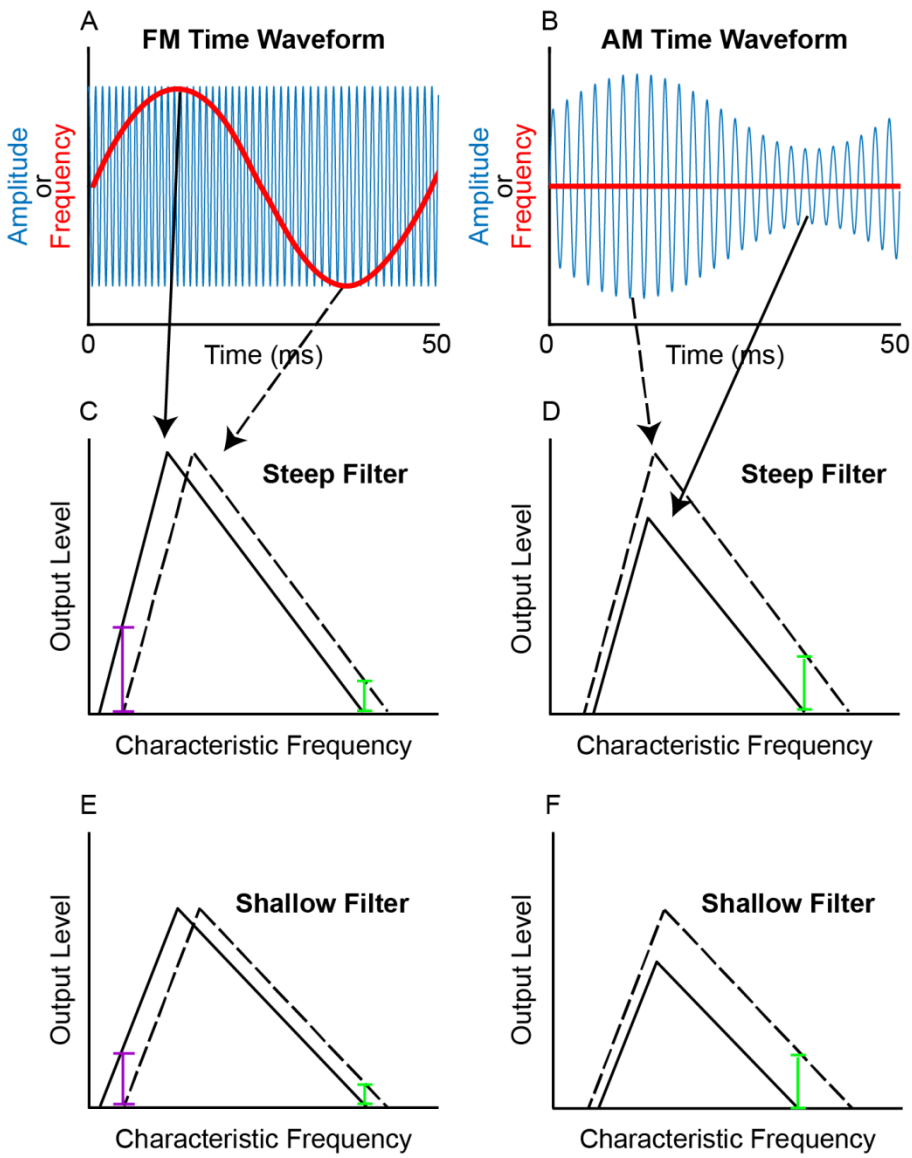

Fig. 1. Schematic of (A) FM and (B) AM time waveforms (f_c = 1 kHz; f_m = 20 Hz) and the resulting changes in basilar membrane excitation for people with (C & D) steep versus (E & F) shallow cochlear filters. In A and B, the blue time waveforms represent amplitude over time, while the superimposed red waveforms are the same stimuli plotted in frequency over time. Panels C and E demonstrate that a place code for FM would result in a greater change in output level on the low-frequency side of the excitation pattern (purple lines) relative to the high-frequency side (green lines).

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Here we measured FM and AM detection at slow (f_m = 1 Hz) and fast (f_m = 20 Hz) modulation rates in a large sample of listeners with hearing thresholds at the carrier frequency (f_c = 1 kHz) ranging from normal (~0 dB sound pressure level, SPL) to severely impaired (~70 dB SPL), consistent with sensorineural hearing loss (SNHL). The fidelity of cochlear frequency tuning was assessed using a psychophysical method to estimate the steepness of the forward masking function around 1 kHz. The results revealed a relationship between the estimated sharpness of cochlear tuning and sensitivity to FM at both fast and slow modulation rates. This relationship remained significant even after controlling for degree of hearing loss, sensitivity to AM, and age. Our results suggest that the fidelity of coding of slow FM depends on the fidelity of cochlear filtering, as predicted by a unified theory of AM and FM coding, and that an additional neural timing code may not be necessary to explain human perception of FM.

Results

Effects of hearing loss on masking functions.

The fidelity of place coding at the test frequency (1 kHz) was measured using pure-tone forward-masking patterns. Participants heard two tones, one at a time, and were instructed to select the tone that had a short 20-ms tone pip directly following it. The masker tones were fixed in frequency (1 kHz) and level, while the tone pip level was adaptively varied to measure the lowest sound level that the participant could detect. Without the presence of a masker, the level of the tone pip reflects the absolute threshold (Supplementary Fig. 1, unfilled circles). In the presence of a pure-tone forward masker, the level of the tone pip depends on the tone pip's frequency proximity to the masker and the shape of the individual's cochlear filters, where

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

detection for tone pips close in frequency to the masker are much poorer (i.e., the level must be higher) than for tone pips farther away in frequency. For each participant, the steepness of the low- and high-frequency slopes of the masking function were estimated by calculating linear regressions between the thresholds for the four lowest (800, 860, 920, and 980 Hz) and highest tone pip frequencies (1020, 1080, 1140, and 1200 Hz), with tone pip frequency transformed to logarithmic units for the regression. Within-subject test-retest reliability of the estimated slope functions was high (bootstrapped simulated test-retest correlations of r = .98 and r = .953 for the low and high slopes, respectively; see Methods). The range of measured masking function slopes in the present study spanned 152 dB/octave for the low slope (-24 – 128 dB/octave; $\bar{x}$ = 49.4) and 120 dB/octave for the high slope (-92.7 – 28.3 dB/octave; Fig. 2, y-axis; $\bar{x}$ = -23.3), which was much greater than that observed in a purely normal-hearing population at 500 Hz (Whiteford and Oxenham, 2015; Whiteford et al., 2017).

Consistent with expectations (Glasberg and Moore, 1986), the amount of hearing loss at the tone pip frequency correlated with the slopes of the masking functions (Fig. 2; low slope: r = -.685, p < .0001, CI = -.804, -.513; high slope: r = .717, p < .0001, CI = .559, .826), confirming that hearing loss is associated with poorer frequency tuning. However, frequency tuning is believed to be governed solely by basilar membrane mechanics and outer hair cell function (Moore, 2007), whereas overall hearing loss also includes contributions from other factors, such as the function of the inner hair cells and the auditory nerve. These additional factors may explain why filter slopes account for only approximately half the variance observed in absolute thresholds.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

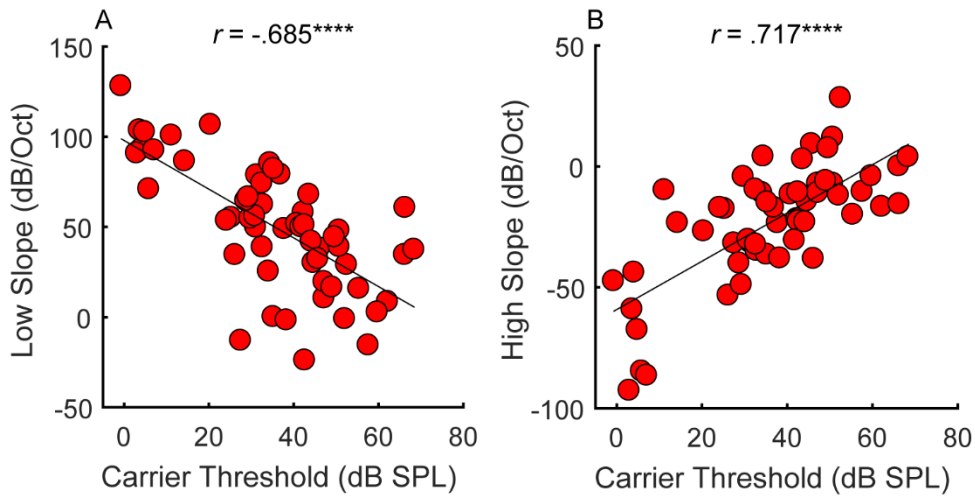

Fig. 2. Correlations between average absolute thresholds at 1 kHz (x-axis) and the steepness of the (A) low and (B) high side of the cochlear filter slopes (n=55). Participants with greater hearing loss at 1 kHz tended to have shallower filter slopes. Correlations marked with an * are significant after Holm's correction (****$p < .0001$).

Average FM and AM detection thresholds.

When compared to earlier results from normal-hearing listeners varying in age (Whiteford et al., 2017), the range of FM detection thresholds, indicated by the upper and lower whiskers in Fig. 3A, was much wider in the present study, whereas the range of AM detection thresholds (Fig. 3B) was comparable. This result suggests that cochlear hearing loss may affect FM more than AM thresholds. For AM, thresholds were generally lower (better) for the high rate than for the low rate (slow AM: $\bar{x}$ = -17.1, s = 4.06; fast AM: $\bar{x}$ = -24.9, s = 3.64; t_{54} = 17.7, p < .0001, d_z = 2.39, CI = 6.98, 8.76), whereas the opposite trend was observed for FM (slow FM: $\bar{x}$ = 2.16, s = 3.24; fast FM: $\bar{x}$ = 2.69, s = 2.72; t_{54} = -2.15, p = .018, d_z = -.29, CI = -1.02, -.036), consistent with earlier studies (Viemeister, 1979; Sheft and Yost, 1990; Moore and Sek, 1995,

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1996; Lacher-Fougère and Demany, 1998; Whiteford and Oxenham, 2015; Whiteford et al., 2017; Whiteford and Oxenham, 2017).

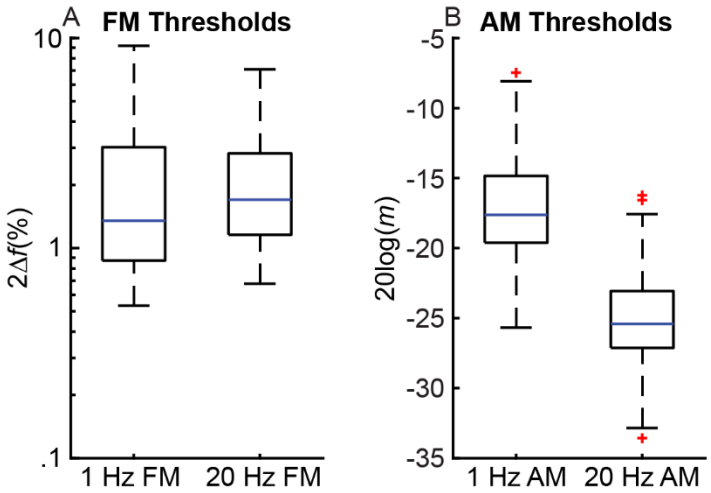

Fig. 3. Boxplots of sensitivity to modulations in (A) frequency and (B) amplitude across 55 listeners. The y-axes correspond to the smallest amount of modulation detectable at threshold, where lower represents better performance. Blue bars are the median of each group. Whiskers are the lowest and highest data points within 1.5 times the lower and higher inter-quartile ranges. Red crosses represent thresholds of individuals whose performance fell outside this range.

Correlations between FM and AM detection.

Test-retest reliability for the estimation of AM and FM detection thresholds was very high (average correlations using a bootstrapping procedure: slow FM, $r = .973$, $p < .0001$, CI = .954, .984; fast FM, $r = .97$, $p < .0001$, CI = .949, .983; slow AM, $r = .925$, $p < .0001$, .874, .956; fast AM, $r = .956$, $p < .0001$, CI = .925, .974; see Methods). If slow FM utilizes a time code, then across-listener variability in slow FM detection should partly reflect variability in time coding. This means that across-listener correlations in tasks known to use a shared code (fast FM, slow

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AM, and fast AM) should be greater than in tasks thought to use different codes (slow FM with any other task). Inconsistent with this prediction, slow and fast FM detection thresholds were strongly correlated (r = .826, p < .0001, CI = .718, .895), as were detection thresholds for slow and fast AM (r = .638, p < .0001, CI = .449, .773) and fast FM and fast AM (r = .317, p = .018, CI = .057, .537) (Fig. 4). The correlation between slow FM and slow AM was not significant (r = .199, p = .072, CI = -.07, .441), but this correlation was not significantly different from the correlation between fast FM and fast AM (Z = -.906, p = .365, two-tailed). Even though participants in the present study varied widely in peripheral place coding fidelity (Fig. 2), correlational trends between FM and AM thresholds generally mirrored those observed in groups of listeners with normal hearing (Whiteford and Oxenham, 2015; Whiteford et al., 2017).

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

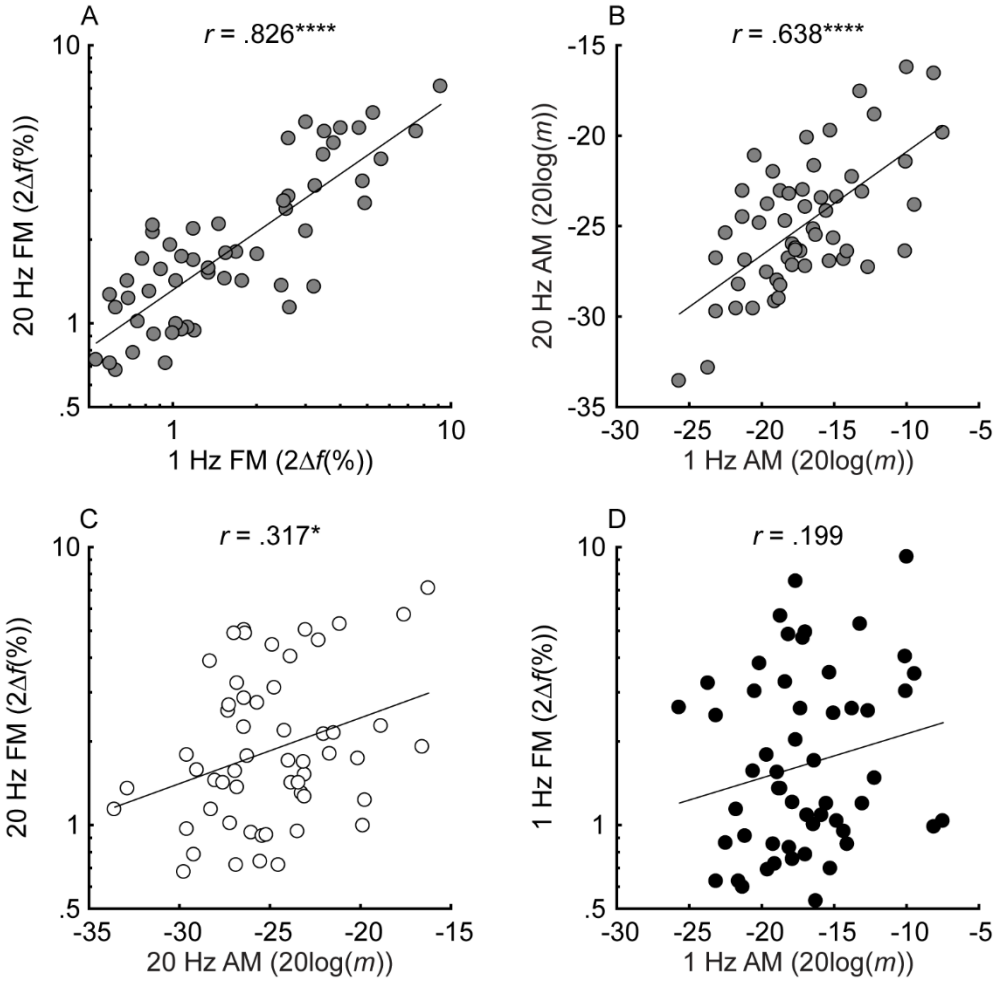

Fig. 4. Individual thresholds for slow (f_m = 1 Hz; black) and fast (f_m = 20 Hz; white) FM and AM detection (n=55). Grey circles represent different rates on the x- and y-axes. FM and AM thresholds are plotted in percent peak-to-peak frequency change (2Δ*f*(%)) and 20log(*m*), where Δ*f* is the frequency excursion from the carrier and *m* is the modulation depth (ranging from 0-1)). For all tasks, lower on the x- or y-axis represents better thresholds. Correlations marked with an * are significant after Holm's correction (****$p < .0001$, ***$p < .001$, **$p < .01$, and *$p < .05$).

The role of frequency selectivity in FM detection.

The unitary neural coding theory of FM and AM predicts that steeper masking functions (implying sharper cochlear tuning) should be related to better FM detection thresholds (Zwicker,

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1956). The current consensus is that theory applies to fast but not slow FM detection (Moore and Sek, 1995, 1996; Lacher-Fougère and Demany, 1998; Strelcyk and Dau, 2009). Our results contradict this consensus by showing that both slow and fast FM detection were similarly strongly related to the masking function slopes (Fig. 5). Age and sensitivity to AM could confound effects of cochlear filtering because they are known to influence FM detection in listeners with normal hearing (Whiteford and Oxenham, 2015; Paraouty et al., 2016; Whiteford et al., 2017). Audibility is not thought to affect FM for levels that are 25 dB or more above absolute threshold (Zurek and Formby, 1981), but it was included as a precaution, since a few listeners with the most hearing loss had stimuli presented at or near 20 dB sensation level (SL), and because hearing loss has been postulated to affect time coding, independent of place coding (Ewert et al., 2018). Partial correlations between FM detection and masking function slopes were conducted, controlling for age, absolute thresholds at 1 kHz (task 1), and AM detection at the corresponding rate, thereby isolating the role of place coding in FM detection. The correlations between the residuals (Fig. 6) demonstrate a significant relation between the low slope and FM detection threshold at both rates (slow FM: r_p = -.364, p = .016, CI = -.574, -.109; fast FM: r_p = -.377, p = .015, CI = -.584, -.124) and no relation between the high slope and FM (slow FM: r_p = -.064, p = .555, CI = -.323, .205; fast FM: r_p = -.084, p = .555, CI = -.341, .186). Because the low slope of the masking function (reflecting the upper slopes of the cochlear filters) is generally steeper than the high slope, it provides more stimulus information relative to the high side (Fig. 1, leftmost column), and is therefore predicted to dominate FM performance (Zwicker, 1956). Sensitivity to AM detection was not related to either the low slopes (slow AM: r = .058, p > .499, CI = -.211, .318; fast AM: r = .277, p = .076, CI = .013, .505) or high slopes (slow AM: r = .007, p > .499, CI = -.259, .272; fast AM: r = -.281, p = .076, CI = -.508, -.017) of the masking

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

functions, demonstrating that the relations between masking function slopes and modulation detection is specific to FM, as predicted by place coding. The results therefore provide strong support for the hypothesis that place coding is utilized for FM detection at both slow and fast rates. These conclusions were confirmed using multiple linear regression analyses (see Supplementary Text 1).

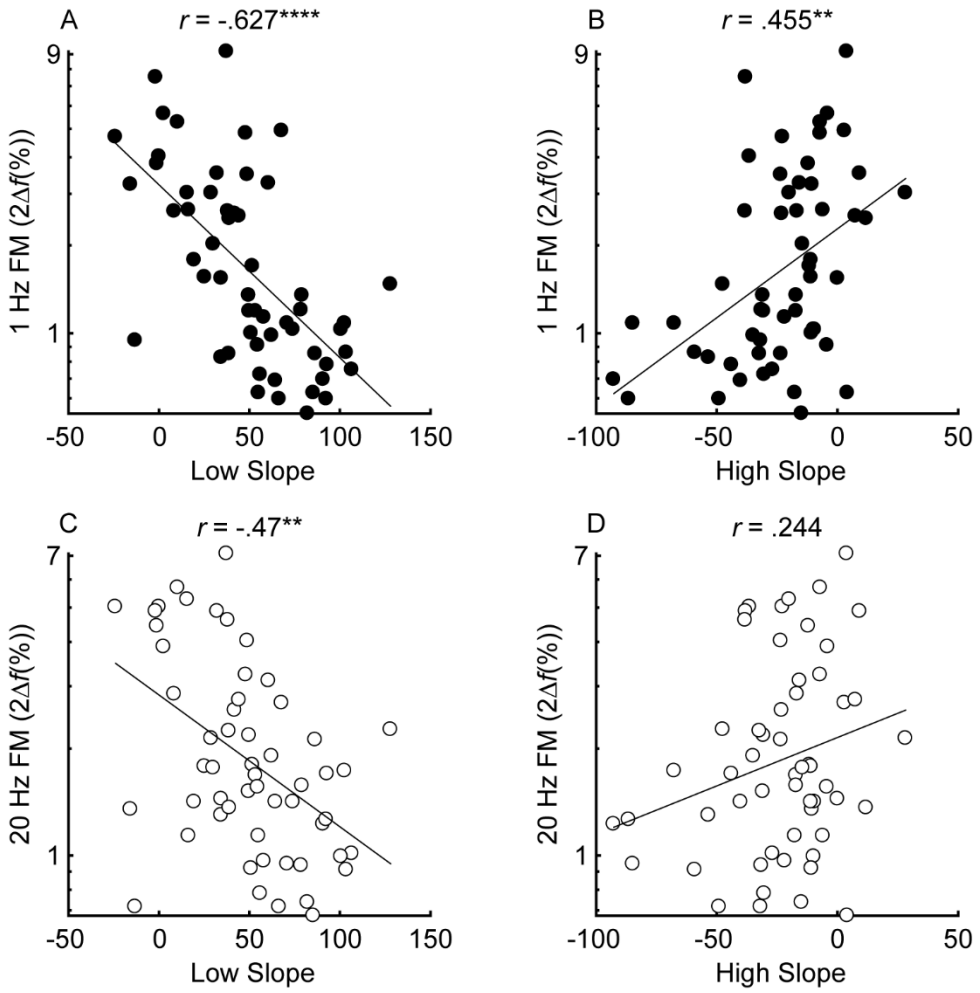

Fig. 5. Correlations between the low slope (A and C) and high slope (B and D) and slow (f_m = 1 Hz; black) and fast (f_m = 20 Hz; white) FM detection (n=55). Correlations marked with an * are significant after Holm's correction (****p < .0001, ***p < .001, **p < .01, and *p < .05).

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

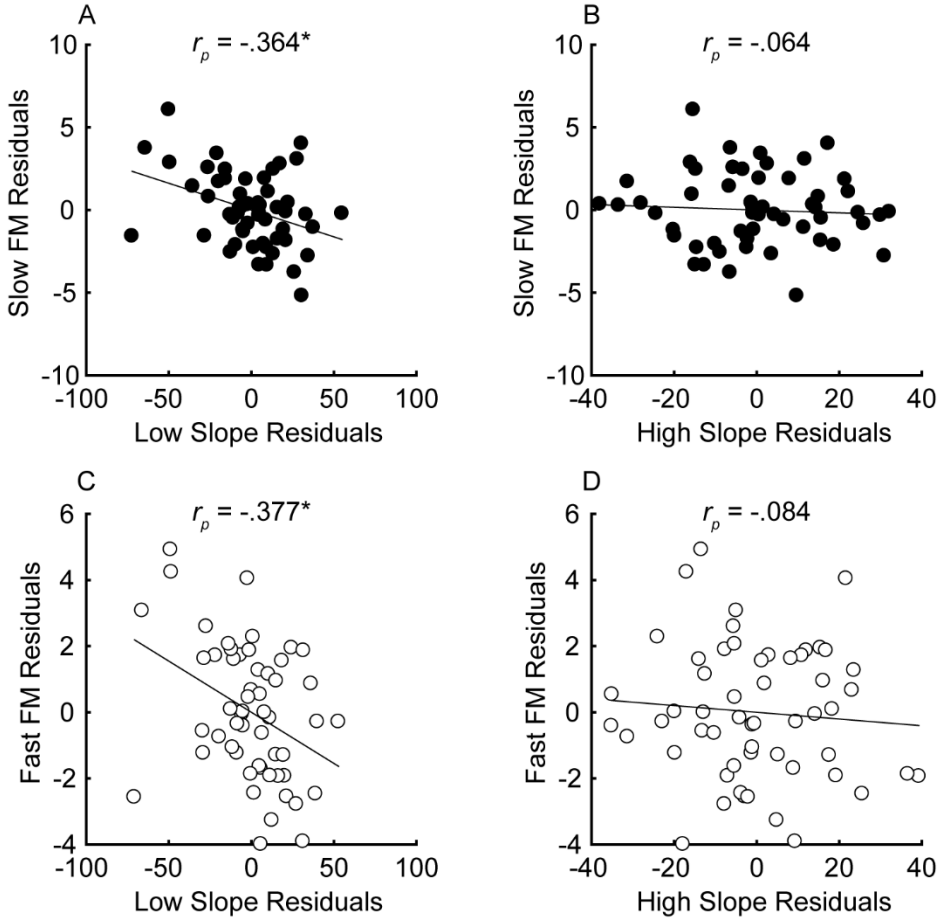

Fig. 6. Partial correlations between the steepness of the masking function slopes (x-axes) and FM detection (y-axes) after variance due to audibility, sensitivity to AM, and age has been partialled out for n=55 participants. Units of the x- and y-axes are arbitrary because they correspond to the residual variance for slow (f_m = 1 Hz; black) and fast FM detection (f_m = 20 Hz; white). Correlations marked with an * are significant after Holm's correction (****p < .0001, ***p < .001, **p < .01, and *p < .05).

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

A unitary code for FM

Our finding that cochlear place coding is equally important for both slow- and fast-rate FM detection was unexpected. Humans' acute sensitivity to slow changes in frequency at carriers important for speech and music has been thought to result from precise neural synchronization to the temporal fine structure of the waveform (Demany and Semal, 1989; Moore and Sek, 1995, 1996; Sek and Moore, 1995; Lacher-Fougère and Demany, 1998; Buss et al., 2004; Strelcyk and Dau, 2009). Multiple linear regression analyses showed that the combined effect of audibility, age, sensitivity to AM, and masking function slopes accounted for about 59.5% and 52.1% of the total variance in slow and fast FM detection thresholds, respectively. This is a high proportion of the variance, particularly considering the relatively rough behavioral approximation used to estimate cochlear tuning.

The clear role for place coding in slow FM is contrary to the widely accepted understanding that a time coding is used to detect FM at slow rates found in speech and music. Instead, our results provide evidence for a unitary code for two crucial features of natural sounds, AM and FM, that extends across the entire range of naturally encountered fluctuations rates. A unitary code for FM and AM at all rates may help account for the high-multicollinearity between FM and AM sensitivity observed here (Fig. 4) and in several previous studies with normal-hearing listeners (Whiteford and Oxenham, 2015; Otsuka et al., 2016; Paraouty and Lorenzi, 2017; Whiteford et al., 2017).

Implications for the perception and neural coding of complex tones

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

This study used pure tones, which are not frequently encountered in the natural environment. However, combinations of pure tones form harmonic complex tones, such as musical instrument sounds, voiced speech, and many animal vocalizations. It is known that humans perceive the pitch of harmonic complex tones in ways that are fundamentally different from other commonly studied species, such as the chinchilla (Shofner and Chaney, 2013), ferret (Walker et al., 2018), or songbird (Bregman et al., 2016). Recent work (Shofner and Chaney, 2013; Walker et al., 2018) has suggested that part of this difference can be explained by the substantially sharper cochlear tuning found in humans than in smaller mammals (Shera et al., 2002, 2010; Sumner et al., 2018; Verschooten et al., 2018). Specifically, sharper human cochlear tuning is believed to explain why humans rely primarily on low-numbered spectrally resolved harmonics (Houtsma and Smurzynski, 1990; Bernstein and Oxenham, 2003), whereas smaller mammals, such as ferrets and chinchillas, rely on the cues in the temporal envelope provided by higher spectrally unresolved harmonics (Shofner and Chaney, 2013; Walker et al., 2018).

The present study extends these previous findings by suggesting that the resolved harmonics, which are most important for human pitch perception, may be represented by their place of stimulation in a way that depends of the lower (and steeper) slope of the excitation pattern, rather than just via the temporal fine structure information encoded via the stimulus-driven spike timing (phase locking) in response to resolved harmonics. This conclusion is consistent with other studies showing that pitch perception is possible even with spectrally resolved harmonics that are too high in frequency to elicit phase locking (Oxenham et al., 2011; Lau et al., 2017). In addition, the fact that timing fidelity in the human auditory nerve is no greater than that found in smaller mammals (Verschooten et al., 2018), supports our conclusion that differences in pitch perception between humans and other mammals cannot be ascribed to

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differences in timing fidelity and phase locking, but instead may be due to differences in the sharpness of cochlear tuning.

Alternative interpretations

One alternative interpretation of our results is that hearing loss leads to a degradation in both spectral resolution and neural phase locking to temporal fine structure, and that it is the degradation in the phase locking, not cochlear filtering, that drives the relationship between spectral resolution and FM coding observed here. There are several reasons why this interpretation is unlikely. First, the literature on whether time coding degrades with SNHL, particularly for tones in quiet, is mixed. Physiological studies with non-human animals have generally found no effects (Harrison and Evans, 1979; Miller et al., 1997) or very small effects (Henry and Heinz, 2012) of SNHL on time coding, with the exception of one study (Woolf et al., 1981). Support from human studies are based on poorer behavioral performance in hearing-impaired listeners in tasks thought to use time coding (Lorenzi et al., 2006; Moore et al., 2006, 2012, Hopkins and Moore, 2007, 2011; Moore, 2014; Füllgrabe and Moore, 2017). However, many of these tasks could also be affected by poorer cochlear tuning (Oxenham et al., 2009). Binaural tasks, involving the discrimination of interaural time differences (ITDs) in the temporal fine structure of stimuli, are likely to rely on phase-locked coding. These studies have not always found a clear relationship between ITD sensitivity and hearing loss, once effects of age and audibility are accounted for (Smoski and Trahiotis, 1986; Hopkins and Moore, 2011).

A second reason why it is unlikely for the role of place coding in FM to be a byproduct of time coding degrading with SNHL is that not all the listeners in the present study had SNHL, yet

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the trends between FM and masking function slopes were maintained despite the inclusion of listeners with normal hearing.

Finally, the relationship between FM and the slopes of the masking function was specific to the low-frequency side of the masking function. Zwicker (1956) predicted over half a century ago that the steeper, low-frequency slope should play a larger role in FM-to-AM conversion. If the current findings were a spurious effect of time coding degrading with hearing loss, then the correlation should not be specific to the low-frequency slope, as the high-frequency slope is also strongly affected by hearing loss ($r = .717$, $p < .0001$). For the sake of parsimony, it seems more reasonable to interpret the similar correlations between the lower masking slope and both slow and fast FM as reflecting the same coding mechanism than to interpret them as coming from different sources with a coincidentally similar correlation.

Explaining superior FM perception at low rates within a unitary framework

A pure cochlear place-based model for FM proposes that FM is transduced to AM through cochlear filtering (Zwicker, 1956). As the frequency sweeps across the tonotopic axis, the auditory system monitors changes in the output of the cochlear filters. For a place-only model to explain FM, it would need to account for the rate-dependent trends in FM and AM sensitivity observed here (Fig. 3) and in many previous studies (Viemeister, 1979; Sheft and Yost, 1990; Moore and Sek, 1995, 1996; Lacher-Fougère and Demany, 1998; Moore and Skrodzka, 2002; Whiteford and Oxenham, 2015, 2017; Whiteford et al., 2017). One possible explanation is that the central auditory system's ability to compare changes in the output between neighboring cochlear filters is more efficient at very slow rates. This interpretation is supported by a computational modeling study showing that frequency and intensity can be represented by a

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

single code, if inter-neuronal noise correlations (Cohen and Kohn, 2011) are taken into account (Micheyl et al., 2013). Such correlations would require relatively long time windows (and hence slow modulation rates) to play a functional role. Thus, such a code would function more efficiently at slow than at fast rates, producing the observed differential effect.

Alternatively, a combined place-time code may predict better sensitivity for slow, low-carrier FM relative to the same carrier at faster rates (Fig. 3) (Paraouty et al., 2018). Place-time models extract timing information in a way that is place dependent (Loeb et al., 1983; Shamma and Klein, 2000). There are various implementations, but place-time models generally rely on an array of coincidence detectors calculating the instantaneous cross-correlation between the phase-locked responses of auditory nerve fibers innervating different cochlear locations. Again, such a correlation mechanism would require a time window over which to evaluate the correlation, and so would predict poorer performance at fast FM rates than at slow FM rates. In addition, the poor frequency tuning that occurs with hearing loss affects the traveling wave response, thereby potentially disrupting this place-time relationship (Ruggero, 2013). A combined place-time code could therefore account for the correlation between slow-rate FM and the low slope of the masking function.

Methods

Participants.

Experimental tasks were carried out by 56 participants (19 male, 37 female; average age of 66.5 years, range: 19.4-78.5 years), with an average sensitivity to tones at 1 kHz of 36.5 dB SPL, ranging from -0.7 to 68.5 dB SPL based on Task 1. The participants had no reported history of cognitive impairment. Pure-tone audiometry was assessed at octave frequencies from 250-

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

8000 Hz. Nine participants had normal hearing, defined as audiometric thresholds ≤ 20 dB hearing level (HL) at 1 kHz in both ears. The other 47 participants had varying degrees of SNHL, with audiometric thresholds at 1 kHz poorer than 20 dB HL in at least one ear and air-bone gaps < 10 dB to preclude a conductive hearing loss. Ears with SNHL ≥ 70 dB SPL from Task 1 were not included in the study. Participants with symmetric hearing (n = 37; asymmetries ≤ 10 dB at 1 kHz from Task 1) completed all monaural experimental tasks in their worse ear. Six participants had SNHL at 1 kHz in both ears, but loss in the poorer ear exceeded the study criterion; for these subjects, tasks were completed in the better ear only. One additional participant was only assessed in their better ear because loss in the poorer ear was near the study criterion (68.6 dB SPL at 1 kHz), and the subject indicated the level was uncomfortable. An additional three participants had one normal-hearing ear and one ear with SNHL at 1 kHz, and only measurements from the impaired ear were used in analyses. The final nine participants had asymmetric SNHL in both ears, defined as an asymmetry > 10 dB on Task 1. For eight of these subjects, the experimental tasks were completed for both ears separately. One participant with asymmetric hearing only completed tasks in their poorer ear due to time constraints (Table 1). However, only performance in the poorer ear was used in the analyses for all nine of these listeners. Participants provided informed consent and were given monetary compensation for their time. The Institutional Review Board of the University of Minnesota approved all experimental protocols.

Table 1. Summary of participants.

Measured Ear	**# of Participants**	**Notes**
Worse ear	38	Subjects with symmetric thresholds (asymmetry <= 10 dB) from Task 1 (n=37) or who could only be assessed in their worse ear due to time constraints (n=1).

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Better ear	7	Subjects with Task 1 thresholds in the worse ear that exceeded the study criterion (n=6) or indicated the SL for Task 2 in their worse ear was uncomfortable (n=1).
Both ears (worse ear used in analyses)	11	Task 1 asymmetry > 10 dB; n=3 had normal hearing in their better ear, and n=8 had SNHL in both ears.

Stimuli.

Stimuli were generated in Matlab (MathWorks) with a sampling rate of 48 kHz using a 24-bit Lynx Studio L22 sound card and presented over Sennheiser HD650 headphones in a sound-attenuating chamber. Tasks were measured monaurally with threshold equalizing noise (TEN) (Moore et al., 2000) presented in the contralateral ear in order to prevent audible cross-talk between the two ears. The TEN was presented continuously in each trial, with the bandwidth spanning 1 octave, geometrically centered around the test frequency. Except for tasks that involved detection of a short (20 ms) tone pip, the TEN level (defined as the level with the auditory filter's equivalent rectangular bandwidth at 1 kHz) was always 25 dB below the target level, beginning 300 ms before the onset of the first interval and ending 200 ms after the offset of the second interval. Because less noise is needed to mask very short targets, the TEN was presented 35 dB below the target level for tasks that involved detection of a short, 20-ms tone pip (Tasks 4 and 7). This noise began 200 ms before the onset of the first interval and ended 100 ms after the offset of the second interval.

To obtain a more precise estimate of sensitivity for the test frequency, pure-tone absolute thresholds were measured for each ear at 1 kHz. The target was 500 ms in duration with 10-ms raised-cosine onset and offset ramps. The reference was 500 ms of silence, and the target and the reference were separated by a 400-ms interstimulus interval (ISI). Tasks involving modulation detection were assessed for the same frequency (f_c = 1 kHz) at slow (f_m = 1 Hz) and fast (f_m = 20

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hz) rates. The target was an FM (Tasks 2 and 3) or AM (Task 4 and 5) pure tone while the reference was an unmodulated pure tone at 1 kHz. Both the target and the reference tones were 2 s in duration with 50-ms raised-cosine onset and offset ramps. In the FM tasks, the starting phase of the modulator frequency was set so that the target always began with either an increase or decrease in frequency excursion from the carrier frequency, with 50% probability determined a priori. A similar manipulation was used for the AM tasks, so that the target always began at either the beginning or middle of a sinusoidal modulator cycle and so was either increasing or decreasing in amplitude at the onset. Stimuli for the modulation tasks were presented at 65 dB SPL or 20 dB sensation level (SL), whichever was greater, based on individualized absolute thresholds at 1 kHz from Task 1.

Detection for a short (20 ms), pure-tone pip was measured with and without the presence of a 1-kHz, 500-ms pure-tone forward masker. Tone-pip frequencies were 800, 860, 920, 980, 1020, 1080, 1140, and 1200 Hz, and both the tone pip and the masker had 10-ms raised cosine onset and offset ramps. The tone pip was presented to one ear, directly following the offset of the masker, and the masker was presented to both ears to avoid potential confusion effects between the offset of the masker and the onset of the tone pip (Neff, 1986). The masker was fixed in level at either 65 dB SPL or 20 dB SL, whichever was greater, based on absolute thresholds for the 500-ms test frequency in the target ear (Task 1). The starting level of the tone pip was always 10 dB below the masker level in the masked condition. For unmasked thresholds, the starting level of the tone pip was either 40 dB SPL or 20 dB SL, whichever was greater, and the tone pip was preceded by 500 ms of silence.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Procedure.

Procedures were adapted from Whiteford et al. (2017) and are described in full below. The experiment took place across 3-6 separate sessions, with each session lasting no longer than 2 hours. All tasks were carried out using a two-interval, two-alternative forced-choice procedure with a 3-down 1-up adaptive method that tracks the 79.4% correct point of the psychometric function (Levitt, 1971). The target was presented in either the first or second interval with 50% a priori probability, and the participant's task was to click the virtual button on the computer screen (labeled "1" or "2") corresponding to the interval that contained the target. Each corresponding response button illuminated red during the presentation of the stimulus (either reference or target). Visual feedback ("Correct" or "Incorrect') was presented on the screen after each trial. All participants completed the tasks in the same order, and the tasks are described below in the order in which they were completed by the participants.

Task 1: Absolute Thresholds at 1 kHz. Participants were instructed to select the button on the computer screen that was illuminated while they heard a tone. The target was a 500-ms, 1-kHz pure tone presented to one ear, and the reference was 500 ms of silence. Three runs were measured for each ear, and the order of the presentation ear (left vs. right) was randomized across runs. Three participants were only assessed in their better ear, due to an extensive amount of hearing loss in the poorer ear according to their 1 kHz audiometric thresholds (all $\geq$ 80 dB HL). The remaining participants completed monaural absolute thresholds for both ears.

On the first trial, the target was presented at 40 dB SPL. The target changed by 8 dB for the first reversal, 4 dB for the next 2 reversals, and 2 dB for all following reversals. Absolute thresholds were determined by calculating the mean level at the final 6 reversal points. If the

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

standard deviation (SD) across the three runs was ≥ 4, then 3 additional runs were conducted for the corresponding ear, and the first three runs were regarded as practice.

Tasks 2 and 3: FM Detection. Participants were instructed to pick the tone that was "modulated" or "changing". At the beginning of each run, the target had a peak-to-peak frequency excursion ($2\Delta f$) of 5.02%. The $2\Delta f$ varied by a factor of 2 for the first two reversal points, a factor of 1.4 for the third and fourth reversal points, and a factor of 1.19 for all following reversal points. The FM difference limen (FMDL) was defined as the geometric mean of $2\Delta f$ at the final 6 reversal points.

Three runs were conducted for each modulation rate, and all three runs for slow-rate FM (f_m = 1 Hz) were completed before fast-rate FM (f_m = 20 Hz). Asymmetric participants with two qualifying ears completed six runs (three runs per ear) for each modulation rate, and the order of the presentation ear was randomized across runs. If the SD across the three runs for a given ear was ≥ 4, the participant completed an additional three runs, and only the last three runs were used in analyses.

Task 4: Detection for 20-ms Tones. Participants were instructed to select the button (labeled "1" or "2") on the computer screen that was illuminated while they heard a short, 20-ms target tone pip. The target was presented at 40 dB SPL or 20 dB SL, whichever was greater, for the first trial of each run. The level of the target changed by 8 dB for the first two reversals, 4 dB for the following two reversals, and 2 dB for all following reversals. The absolute threshold was defined as the mean target level at the final six reversal points.

Participants completed one run for each of the eight tone-pip frequencies: 800, 860, 920, 980, 1020, 1080, 1140, and 1200 Hz. The order of the tone-pip frequency conditions was

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

randomized across runs. Asymmetric participants with two qualifying ears had the order of the runs further blocked by presentation ear, so that 8 runs for the same ear had to be completed before any conditions in the opposite ear were measured. Whether the right or left ear was assessed first was randomized. One additional run was conducted for any conditions with an SD ≥ 4 dB, and only the final run for each condition was used in analyses.

Tasks 5 and 6: AM Detection. The instructions for AM detection were the same as the instructions for FM detection. The first trial of each run had a target with an AM depth of -7.96, in 20log(m) units. The target modulation depth changed by 6 dB for the first two reversals, 2 dB for the next two reversals, and 1 dB for all following reversals. The AM difference limen (AMDL) was defined as the mean modulation depth (in 20log(m)) at the last 6 reversal points.

In the same manner as the FM tasks, all three runs for slow-rate AM (f_m = 1 Hz) were completed before fast-rate AM (f_m = 20 Hz). Asymmetric participants with two qualifying ears completed six runs (three runs per ear) for each modulation rate, and the order of the presentation ear was randomized across runs. If the SD across the first three runs for a given condition were ≥ 4 dB, then three additional runs were conducted, and only the final three runs were analyzed.

Task 7: Forward Masking Patterns. The task was to determine which of two tones was followed by a short, 20-ms tone pip. Two runs were measured for each of the eight tone-pip frequencies (800, 860, 920, 980, 1020, 1080, 1140, and 1200 Hz), for a total of 16 runs, and the order of the tone-pip condition was randomized across runs. Asymmetric participants with two qualifying ears had the order of the runs further blocked by presentation ear, so that 8 runs for the same ear had to be completed before any conditions in the opposite ear were presented. The 1-kHz, 500-

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ms masker tones were fixed in frequency and level, presented binaurally at 65 dB SPL or 20 dB SL based on absolute thresholds from Task 1, whichever was greater. Within a trial, each masker was either directly followed by a 20-ms tone pip, presented monaurally to the target ear, or 20-ms of silence. The starting level of the tone pip was 10 dB below the masker level in the corresponding ear. The level of the tone pip changed by 8 dB for the first two reversals, 4 dB for the third and fourth reversals, and 2 dB for the following reversals. The masked threshold for each tone-pip frequency condition was calculated as the mean tone-pip level at the final 6 reversal points. For a given subject, if the SD of the masked threshold across the two runs was ≥ 4 dB, then the subject completed two additional runs for the corresponding tone-pip frequency. For these conditions, only the final two runs were used in analyses, and the first two runs were regarded as practice. The average across the final two runs for each tone-pip frequency was used in analyses.

Sample size. Because the strength of the relationship between FM sensitivity and forward masking slopes was unknown in listeners varying in degree of SNHL, and the number of people with SNHL at 1 kHz was expected to be limited, we set a minimum sample size requirement for SNHL subjects based on the smallest effect we would like to be able to detect. To detect a moderate correlation between masking function slopes and FM sensitivity (r = .4, alpha = .05, one-tailed test) with a power of .9, we needed a sample of n=47. We also aimed to recruit an additional 10 participants with NH at 1 kHz of similar age to the SNHL subjects. The NH sample was limited to 10 people to ensure a relatively even distribution of absolute thresholds at 1 kHz. One of these anticipated NH subjects had mild SNHL at 1 kHz in their worse ear, leading to a sample size of n=57, with n=9 NH listeners and n=48 SNHL. One SNHL subject reported a

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

history of neurological issues and was excluded from the study. Another SNHL subject had unusually poor FM sensitivity at both rates, with thresholds greater than 3 SD from the group mean. This outlier was excluded from all analyses, leading to a final sample size of n=55. Including the outlier in all analyses generally did not affect the results (Supplementary Text 2, Supplementary Table 1, and Supplementary Figs. 2-3).

Statistical analyses. The mean log-transformed thresholds ($10\log_{10}(2\Delta f\,(\%))$ and $20\log_{10}(m)$) were used in all analyses to better approximate normality, where $2\Delta f\,(\%)$ is the peak-to-peak frequency excursion (for FM) as a percentage of the carrier frequency, and m is the modulation index (for AM). All reported means ($\bar{x}$) and standard deviations (s) correspond to the log-transformed data. Confidence intervals (CIs) are 95% CIs. Pearson correlations were used to assess continuous trends; the corresponding p values were adjusted using Holm's method to correct for family-wise error rate (Holm, 1979). The p values corresponding to the correlations in Fig. 1 were corrected for 2 comparisons, 4 comparisons for Fig. 3 (all FM and AM correlations), and 8 comparisons for Figs. 4 and 5 (all FM correlations with masking function slopes). The masking function slopes and AM correlations were corrected for 4 comparisons. Paired-samples t-tests were used to assess rate-dependent differences, and effect sizes were calculated using Cohen's d_z (Lakens, 2013). The cocor package was used to calculate significant differences between correlations using Steiger's modification (Steiger, 1980; Diedenhofen and Musch, 2015). All tests were one-tailed unless otherwise stated in the results.

Bootstrap analyses were conducted to estimate the highest possible correlation detectable for each modulation task and the forward masking task, in order to ensure that correlations with these measures were not limited by test-retest reliability. For each subject and for each

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

modulation condition, performance was simulated by randomly sampling 6 runs (3 test and 3 retest) from a normal distribution based on the individual means and standard deviations from the corresponding task. An analogous procedure was conducted for each individual's masked thresholds for every tone-pip condition, with 4 runs (2 test and 2 retest) sampled from each individualized normal distribution. The average simulated runs were used to estimate the low and high frequency slopes of the masking function by calculating a linear regression between the 4 lowest and 4 highest tone-pip frequency conditions for the average test and the average retest runs (4 regressions per iteration). Simulated test-retest correlations were calculated using the simulated slopes for n=55 subjects (for forward masking) or the simulated average test and retest thresholds for each subject (for the modulated tasks). This process was repeated for 100,000 iterations. The correlations were transformed using Fisher's *r* to *z* transformation, averaged, and then transformed back to *r*, yielding an average test-retest correlation whose maximum is limited by within-subject error.

Data availability. The data that support the findings will be available in the Data Repository for U of M.

Acknowledgements

Supported by Grant R01 DC005216 from the National Institutes of Health (to A.J.O.) and an Eva O. Miller Fellowship (to K.L.W.).

Author contributions

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

KLW and AJO conceived of and designed the experiment; HAK and KLW collected the data; KLW analyzed the data; KLW and AJO wrote the paper.

Competing interests: The authors declare that no competing interests exist.

References

Attias H, Schreiner CE (1997) Temporal low-order statistics of natural sounds. Adv Neural Inf Process Syst 9:27–33.

Bernstein JG, Oxenham AJ (2003) Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? J Acoust Soc Am 113:3323–3334.

Bregman MR, Patel AD, Gentner TQ (2016) Songbirds use spectral shape, not pitch, for sound pattern recognition. Proc Natl Acad Sci 113:1666–1671.

Buss E, Hall JW, Grose JH (2004) Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. Ear Hear 25:242–250.

Chen H, Zeng F-G (2004) Frequency modulation detection in cochlear implant subjects. J Acoust Soc Am 116:2269–2277.

Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. Nat Neurosci 14:811–819.

Deal JA, Betz J, Yaffe K, Harris T, Purchase-Helzner E, Satterfield S, Pratt S, Govil N, Simonsick EM, Lin FR (2017) Hearing impairment and incident dementia and cognitive decline in older adults: The Health ABC Study. Journals Gerontol Ser A Biol Sci Med Sci 72:703–709.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Demany L, Semal C (1989) Detection thresholds for sinusoidal frequency modulation. J Acoust Soc Am 85:1295–1301.

Diedenhofen B, Musch J (2015) cocor: A comprehensive solution for the statistical comparison of correlations Olivier J, ed. PLoS One 10:e0121945.

Ewert SD, Paraouty N, Lorenzi C (2018) A two-path model of auditory modulation detection using temporal fine structure and envelope cues. Eur J Neurosci:1–14.

Füllgrabe C, Moore BCJ (2017) Evaluation of a method for determining binaural sensitivity to temporal fine structure (TFS-AF test) for older listeners with normal and impaired low-frequency hearing. Trends Hear 21:1–14.

Glasberg BR, Moore BCJ (1986) Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. J Acoust Soc Am 79:1020–1033.

Grose JH, Mamo SK (2012) Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects. Hear Res 294:49–54.

Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. Physiol Rev 90:983–1012.

Harrison R V, Evans EF (1979) Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. Arch Otorhinolaryngol 224:71–78.

He N, Mills JH, Dubno JR (2007) Frequency modulation detection: Effects of age, psychophysical method, and modulation waveform. J Acoust Soc Am 122:467–477.

Heinz MG, Colburn HS, Carney LH (2001) Evaluating auditory performance limits: i. one-

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

parameter discrimination using a computational model for the auditory nerve. Neural Comput 13:2273–2316.

Henry KS, Heinz MG (2012) Diminished temporal coding with sensorineural hearing loss emerges in background noise. Nat Neurosci 15:1362–1364.

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70.

Hopkins K, Moore BCJ (2007) Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. J Acoust Soc Am 122:1055–1068.

Hopkins K, Moore BCJ (2011) The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J Acoust Soc Am 130:334–349.

Houtsma AJM, Smurzynski J (1990) Pitch identification and discrimination for complex tones with many harmonics. J Acoust Soc Am 87:304.

Ives DT, Calcus A, Kalluri S, Strelcyk O, Sheft S, Lorenzi C (2013) Effects of noise reduction on AM and FM perception. J Assoc Res Otolaryngol 14:149–157.

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577.

Lacher-Fougère S, Demany L (1998) Modulation detection by normal and hearing-impaired listeners. Audiology 37:109–121.

Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. Front Psychol 4:1–12.

Lau BK, Mehta AH, Oxenham AJ (2017) Superoptimal perceptual integration suggests a place-

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

based representation of pitch at high frequencies. J Neurosci 37:9013–9021.

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

Liberman CM, Dodds LW, Learson DA (1986) Structure-Function Correlation in Noise-Damaged Ears: A Light and Electron-Microscopic Study. In: Basic and Applied Aspects of Noise-Induced Hearing Loss (Salvi RJ, Henderson D, Hamernik RP, Colletti V, eds), pp 163–177. Boston, MA: Springer US.

Lin FR, Albert M (2014) Hearing loss and dementia – who is listening? Aging Ment Health 18:671–673.

Lin FR, Metter EJ, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L (2011) Hearing loss and incident dementia. Arch Neurol 68:214–220.

Lin FR, Yaffe K, Xia J, Xue Q-L, Harris TB, Purchase-Helzner E, Satterfield S, Ayonayon HN, Ferrucci L, Simonsick EM (2013) Hearing loss and cognitive decline in older adults. JAMA Intern Med 173:293.

Loeb GE, White MW, Merzenich MM (1983) Spatial cross-correlation. Biol Cybern 47:149–163.

Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc Natl Acad Sci 103:18866–18869.

Micheyl C, Schrater PR, Oxenham AJ (2013) Auditory frequency and intensity discrimination explained using a cortical population rate code. PLoS Comput Biol 9:e1003336.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Miller RL, Schilling JR, Franck KR, Young ED (1997) Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory nerve fibers. J Acoust Soc Am 101:3602–3616.

Moiseff A, Konishi M (1981) Neuronal and behavioral sensitivity to binaural time differences in the owl. J Neurosci 1:40–48.

Moore BCJ (2007) Cochlear Hearing Loss: Physiological, Psychological and Technical Issues, Second Edi. John Wiley & Sons.

Moore BCJ (2014) Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss.

Moore BCJ, Glasberg BR, Hopkins K (2006) Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. Hear Res 222:16–27.

Moore BCJ, Glasberg BR, Stoev M, Füllgrabe C, Hopkins K (2012) The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). J Acoust Soc Am 131:1003.

Moore BCJ, Huss M, Vickers DA, Glasberg BR, Alcántara JI (2000) A test for the diagnosis of dead regions in the cochlea. Br J Audiol 34:205–224.

Moore BCJ, Sek A (1995) Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). J Acoust Soc Am 97:2468–2478.

Moore BCJ, Sek A (1996) Detection of frequency modulation at low modulation rates: Evidence

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for a mechanism based on phase locking. J Acoust Soc Am 100:2320–2331.

Moore BCJ, Skrodzka E (2002) Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation. J Acoust Soc Am 111:327–335.

Moore BCJ, Vickers DA, Plack CJ, Oxenham AJ (1999) Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. J Acoust Soc Am 106:2761–2778.

Neff DL (1986) Confusion effects with sinusoidal and narrow-band noise forward maskers. J Acoust Soc Am 79:1519–1529.

Nelken I, Rotman Y, Yosef OB (1999) Responses of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157.

Otsuka S, Furukawa S, Yamagishi S, Hirota K, Kashino M (2016) Relation between cochlear mechanics and performance of temporal fine structure-based tasks. J Assoc Res Otolaryngol 17:541–557.

Oxenham AJ, Micheyl C, Keebler M V (2009) Can temporal fine structure represent the fundamental frequency of unresolved harmonics? J Acoust Soc Am 125:2189–2199.

Oxenham AJ, Micheyl C, Keebler M V, Loper A, Santurette S (2011) Pitch perception beyond the traditional existence region of pitch. Proc Natl Acad Sci U S A 108:7629–7634.

Paraouty N, Ewert SD, Wallaert N, Lorenzi C (2016) Interactions between amplitude modulation and frequency modulation processing: Effects of age and hearing loss. J Acoust Soc Am 140:121–131.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Paraouty N, Lorenzi C (2017) Using individual differences to assess modulation-processing mechanisms and age effects. Hear Res 344:38–49.

Paraouty N, Stasiak A, Lorenzi C, Varnet L, Winter IM (2018) Dual coding of frequency modulation in the ventral cochlear nucleus. J Neurosci 38:4123–4137.

Robertson D, Manley GA (1974) Manipulation of frequency analysis in the cochlear ganglion of the guinea pig. J Comp Physiol 91:363–375.

Ruggero MA (2013) Cochlear delays and the traveling waves: Comments on “experimental look at cochlear mechanics.” Audiology 33:131–142.

Saberi K, Hafter ER (1995) A common neural code for frequency- and amplitude-modulated sounds. Nature 374:537–539.

Schreiner CE, Langner G (1988) Coding of temporal patterns in the central auditory nervous system. In: Auditory Function: Neurobiological Basis of Hearing (Edelman GM, Gall WE, Cowan WM, eds), pp 337–361. New York: John Wiley & Sons.

Sek A, Moore BCJ (1995) Frequency discrimination as a function of frequency, measured in several ways. J Acoust Soc Am 97:2479–2486.

Shamma S, Klein D (2000) The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. J Acoust Soc Am 107:2631–2644.

Shannon R V, Zeng F-G, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science (80-) 270:303–304.

Sheft S, Shafiro V, Lorenzi C, McMullen R, Farrell C (2012) Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

perception. Ear Hear 33:709–720.

Sheft S, Yost WA (1990) Temporal integration in amplitude modulation detection. J Acoust Soc Am 88:796–805.

Shera CA, Guinan JJ, Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci U S A 99:3318–3323.

Shera CA, Guinan JJ, Oxenham AJ (2010) Otoacoustic estimation of cochlear tuning: Validation in the chinchilla. J Assoc Res Otolaryngol 11:343–365.

Shofner WP, Chaney M (2013) Processing pitch in a nonhuman mammal (Chinchilla laniger). J Comp Psychol 127:142–153.

Siebert WM (1970) Frequency discrimination in the auditory system: Place or periodicity mechanisms? Proc IEEE 58:723–730.

Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. Nature 416:87–90.

Smoski WJ, Trahiotis C (1986) Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. J Acoust Soc Am 79:1541–1547.

Steiger JH (1980) Tests for comparing elements of a correlation matrix. Psychol Bull 87:245–251.

Strelcyk O, Dau T (2009) Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. J Acoust Soc Am 125:3328–3345.

Sumner CJ, Wells TT, Bergevin C, Sollini J, Kreft HA, Palmer AR, Oxenham AJ, Shera CA

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(2018) Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. Proc Natl Acad Sci:1–5.

Thomson RS, Auduong P, Miller AT, Gurgel RK (2017) Hearing loss as a risk factor for dementia: A systematic review. Laryngoscope Investig Otolaryngol 2:69–79.

Verschooten E, Desloovere C, Joris PX (2018) High-resolution frequency tuning but not temporal coding in the human cochlea. PLOS Biol 16:e2005164.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 66:1364–1380.

Walker KMM, Gonzalez R, Kang J, McDermott JH, King AJ (2018) Pitch perception is adapted to species-specific cochlear filtering. bioRxiv.

Wallaert N, Moore BCJ, Lorenzi C (2016) Comparing the effects of age on amplitude modulation and frequency modulation detection. J Acoust Soc Am 139:3088–3096.

Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. Neuroscience 157:484–493.

Whiteford KL, Kreft HA, Oxenham AJ (2017) Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age. J Assoc Res Otolaryngol 18:619–633.

Whiteford KL, Oxenham AJ (2015) Using individual differences to test the role of temporal and place cues in coding frequency modulation. J Acoust Soc Am 138:3093–3104.

Whiteford KL, Oxenham AJ (2017) Auditory deficits in amusia extend beyond poor pitch perception. Neuropsychologia 99:213–224.

bioRxiv preprint doi: https://doi.org/10.1101/497248; this version posted December 19, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Woolf NK, Ryan AF, Bone RC (1981) Neural phase-locking properties in the absence of cochlear outer hair cells. Hear Res 4:335–346.

Zeng F-G, Nie K, Stickney GS, Kong Y-Y, Vongphoe M, Bhargave A, Wei C, Cao K (2005) Speech recognition with amplitude and frequency modulations. Proc Natl Acad Sci 102:2293–2298.

Zurek PM, Formby C (1981) Frequency-discrimination ability of hearing-impaired listeners. J Speech Lang Hear Res 24:108.

Zwicker E (1956) Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs. Acustica 6:365–381.