

# Using individual differences to test the role of temporal and place cues in coding frequency modulation

Kelly L. Whiteford<sup>a)</sup> and Andrew J. Oxenham

*Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455, USA*

(Received 23 July 2015; revised 14 October 2015; accepted 20 October 2015; published online 18 November 2015)

The question of how frequency is coded in the peripheral auditory system remains unresolved. Previous research has suggested that slow rates of frequency modulation (FM) of a low carrier frequency may be coded via phase-locked temporal information in the auditory nerve, whereas FM at higher rates and/or high carrier frequencies may be coded via a rate-place (tonotopic) code. This hypothesis was tested in a cohort of 100 young normal-hearing listeners by comparing individual sensitivity to slow-rate (1-Hz) and fast-rate (20-Hz) FM at a carrier frequency of 500 Hz with independent measures of phase-locking (using dynamic interaural time difference, ITD, discrimination), level coding (using amplitude modulation, AM, detection), and frequency selectivity (using forward-masking patterns). All FM and AM thresholds were highly correlated with each other. However, no evidence was obtained for stronger correlations between measures thought to reflect phase-locking (e.g., slow-rate FM and ITD sensitivity), or between measures thought to reflect tonotopic coding (fast-rate FM and forward-masking patterns). The results suggest that either psychoacoustic performance in young normal-hearing listeners is not limited by peripheral coding, or that similar peripheral mechanisms limit both high- and low-rate FM coding.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4935018>]

[FJG]

Pages: 3093–3104

## I. INTRODUCTION

Periodic sounds represent an important category of natural sounds, including voiced speech, song, and many animal vocalizations. Despite their importance, there is very little consensus regarding how periodic sounds are coded in the auditory system (Plack *et al.*, 2005; Oxenham, 2013). At the most peripheral level (in the cochlea) and for the simplest periodic sounds (sinusoids), two classical theories exist. Pitch may be coded based on the *place* of maximal excitation on the cochlea, leading to changes in the rate of firing in auditory nerve fibers (rate-place code), or by the stimulus-driven *timing* of phase-locked action potentials in the auditory nerve (temporal code).

It is generally believed that low-frequency pure tones are coded by the more precise temporal code, whereas higher frequencies are coded primarily via a rate-place code. The evidence for this conjecture is indirect and comes from different sources. First, auditory-nerve phase-locking (as quantified by measures such as the synchrony index) in mammals, such as cat and guinea-pig, is strong at low frequencies but degrades rapidly at frequencies higher than about 1–2 kHz, with the exact cut-off frequency depending on the species (Rose *et al.*, 1967; Johnson, 1980; Palmer and Russell, 1986), suggesting that temporal coding is not viable at higher frequencies. Second, human behavioral pure-tone frequency discrimination (and detection of slow frequency modulation, FM) is relatively good at low frequencies, but becomes dramatically worse above about 3–4 kHz, leading to poorer difference limens (e.g., Moore, 1973; Moore and

Sek, 1995) and a reduced ability to recognize even familiar melodies (Attneave and Olson, 1971; Oxenham *et al.*, 2011). Despite the general consensus about the role of the temporal code at low frequencies, studies do not agree on the exact frequency above which the rate-place code becomes dominant, with estimates ranging from around 4 kHz (e.g., Moore, 1973) to above 8 kHz (Moore and Ernst, 2012). Third, studies have found little to no relationship between pure-tone frequency discrimination at low or high frequencies and frequency selectivity, suggesting that a rate-place code based on tonotopic representation is unlikely to limit performance (Tyler *et al.*, 1983; Moore and Peters, 1992).

Another approach to elucidating the coding of frequency has involved the detection of changes in frequency over time, known as FM. Here again, indirect evidence has been used to suggest a distinction between temporal coding and rate-place coding, depending on the conditions. Sensitivity to FM tends to be greatest at low carrier frequencies ( $f_c < 4000$  Hz) and at slow modulation rates ( $f_m \leq 10$  Hz) (Moore and Sek, 1995, 1996; Moore and Skrodzka, 2002). This pattern of results can be explained if it is assumed that low carrier frequencies are coded via a temporal code that is “sluggish,” in that it can only follow relatively slow rates of frequency change (Sek and Moore, 1995; Moore and Sek, 1996; Plack *et al.*, 2005). At higher carrier frequencies and higher modulation rates, poorer performance is explained through a reliance on rate-place coding of the temporal-envelope fluctuations induced by the FM. Although FM tones do not have any inherent envelope fluctuations (i.e., the envelope is flat), envelope cues can potentially be extracted from FM via cochlear filtering, such that FM is converted to AM, which is then detected by the fluctuations

<sup>a)</sup>Electronic mail: whit1945@umn.edu

in firing rate (rather than the timing of individual spikes) in the auditory nerve (e.g., Zwicker, 1970; Coninx, 1978a,b; Moore and Sek, 1992, 1994; Edwards and Viemeister, 1994a,b; Saberi and Hafter, 1995).

Additional support for a two-mechanism model for FM comes from a variety of behavioral studies on FM and AM detection, alone and in combination. First, there is an added benefit for quasi-trapezoidal FM detection at low carriers compared to quasi-trapezoidal AM detection, indicating that more time spent at the modulation extremes is more beneficial for detecting slow FM (i.e., where phase-locking may occur) than for detecting slow AM (Moore and Sek, 1995). Second, when a fixed amount of AM is added to FM, the added AM interferes more with the detection of fast-rate than slow-rate FM at low carrier frequencies, suggesting that slow-rate FM is coded differently from slow-rate AM. In contrast, the amount of interference of AM on FM detection at high carrier frequencies (e.g., 6000 Hz, where phase-locking is unlikely to be a strong cue) is similar at all modulation rates, suggesting similar cues for both AM and FM detection (Moore and Sek, 1996). Third, the discriminability of AM from FM decreases with increasing modulation rate, suggesting that AM and FM may use similar (and, hence, confusable) mechanisms at fast modulation rates, but separate mechanisms at slower modulation rates (Edwards and Viemeister, 1994b).

More direct evidence for the role of temporal and rate-place codes may come from correlations in performance between different tasks thought to rely on the same peripheral code. Ochi *et al.* (2014) tested the role of a phase-locking in frequency coding by correlating performance in a frequency-discrimination task with the discrimination of interaural time differences (ITDs), which are known to be represented via a temporal code. Both tasks used a bandpass-filtered tone complex centered around 1000 Hz, with a fundamental frequency ( $F_0$ ) of 100 Hz. Contrary to predictions, no positive correlation (and a slight non-significant negative correlation) was found between the monaural frequency-discrimination task and the binaural ITD task. One reason for the lack of the expected correlation may be because of the difference in the procedures used: the frequency-discrimination task involved identifying which of two intervals included changes in the stimulus frequencies, whereas the ITD task involved not only detecting an ITD, but determining the direction of ITD change from one interval to the next. In addition, the number of participants (22) was rather small for identifying correlations based on individual differences between young normal-hearing listeners, especially when compared to recent studies using individual differences paradigms (Kidd *et al.*, 2007; McDermott *et al.*, 2010). Large samples are likely to be necessary to accurately measure performance variance within the normal-hearing population.

Previous work has assessed individual differences on a variety of psychoacoustical tasks within both normal-hearing and hearing-impaired populations to reveal potential underlying coding mechanisms (e.g., Festen and Plomp, 1981, 1983; Johnson *et al.*, 1987; Watson *et al.*, 1996; Kidd *et al.*, 2007; McDermott *et al.*, 2010). Our experiment used a

similar paradigm, involving 100 young normal-hearing listeners. We aimed to minimize differences in task procedures and stimuli, with different tasks designed to tap into different underlying codes. Both diotic and dichotic AM and FM detection were tested. Phase-locked sensitivity to temporal fine structure (TFS) cues was measured using a dichotic FM disparity task, where differences in the instantaneous phase between each ear result in ITDs. Dichotic and diotic detection thresholds for slow ( $f_m = 1$  Hz) and fast ( $f_m = 20$  Hz) modulation rates were measured for both FM and AM of a 500-Hz carrier. In addition, frequency selectivity was estimated using forward-masking patterns centered around 500 Hz, along with absolute thresholds at and around 500 Hz. If slow-rate FM detection is based on phase-locking, then performance in the slow-rate (diotic) FM detection task should be strongly correlated with performance in the slow-rate dichotic FM (ITD) detection task. Similarly, if fast-rate FM detection is based on a rate-place code, then performance in the fast-rate diotic FM detection task should be correlated with both fast-rate diotic AM (representing intensity coding) and the slopes of the forward-masking patterns (representing frequency selectivity).

## II. METHODS

### A. Participants

One hundred young adults (25 male,  $M = 21.1$  yr, range: 18–32 yr) were recruited through the Research Experience Program at the University of Minnesota. All participants provided written informed consent and had normal hearing, defined as audiometric thresholds of 20 dB hearing level (HL) or better for pure tones at octave frequencies between 0.25 and 8 kHz. Participants were compensated with course credit or hourly payment for their time. The protocols were approved by the University of Minnesota Institutional Review Board.

### B. Stimuli

Stimuli were presented over open-ear headphones (Sennheiser HD650, Old Lyme, CT) in a sound-attenuating chamber. All FM and AM stimuli, diotic and dichotic, were presented at 60 dB sound pressure level (SPL). The FM and AM tasks involved either detection of FM or AM, or the detection of an interaural disparity in phase or level (dichotic FM detection and dichotic AM detection, respectively). In all cases, the carrier was a 500-Hz pure tone, 2 s in duration, including 50-ms raised-cosine onset and offset ramps. The frequency modulation difference limens (FMDLs) and amplitude modulation difference limens (AMDLS) were measured for slow ( $f_m = 1$  Hz) and fast ( $f_m = 20$  Hz) sinusoidal modulation rates. For diotic FM, the starting phase of the modulator began with either an increase or a decrease in frequency excursion from the carrier ( $\Delta f$ ), with 50% *a priori* probability. For the diotic AM detection task, the target tone randomly began at an amplitude peak or trough. The listeners' task was to identify which of two intervals contained the modulated, as opposed to the unmodulated, tone.

For the dichotic FM detection tasks, the target tone was an FM tone, with an opposite modulator starting phase in each ear. One ear was presented with an FM tone beginning with an increase in  $\Delta f$ , while the opposite ear was presented with an FM tone beginning with a decrease in  $\Delta f$ . Because the modulator starting phases are different, the two tones shift in and out of phase with each other over time, creating a moving, intracranial image when  $f_m = 1$  Hz. Figure 1 plots an example of dynamic ITDs as a function of time when  $\Delta f = 0.06\%$  and  $f_m = 1$  Hz, calculated based on the running phase difference between the signals in each ear. The reference tone was a 2-s diotic FM tone, randomly beginning with either an increase or a decrease in  $\Delta f$ . The starting instantaneous frequency for all tones was the carrier frequency of 500 Hz. The carrier, modulation rates, level, and duration were identical to those in the diotic FM tasks. An analogous design was used for the dichotic AM disparity tasks, with the target tone containing opposite modulator starting phases in each ear. One ear was presented with an AM tone beginning at an envelope peak, while the other ear was presented with an AM tone beginning at an envelope trough. The reference tone was a diotic 2-s AM tone, randomly beginning with either an envelope peak or an envelope trough.

For the forward-masking task, the forward masker was a 500-Hz pure tone, presented at 70 dB SPL for a total duration of 500 ms. The signal was 20 ms in total duration, and both the masker and the signal had 10-ms raised-cosine onset and offset ramps. The onset of the signal was contiguous with the offset of the masker, resulting in a 10-ms gap between the offset of the masker and the onset of the signal at the half-amplitude points of their respective envelopes. Thresholds were measured for signal frequencies of 400, 430, 460, 490, 510, 540, 570, and 600 Hz. The slope of masking function (signal threshold as a function of masker-signal frequency difference in octaves, calculated separately for signals below and above the masker frequency) provided an estimate of frequency selectivity.

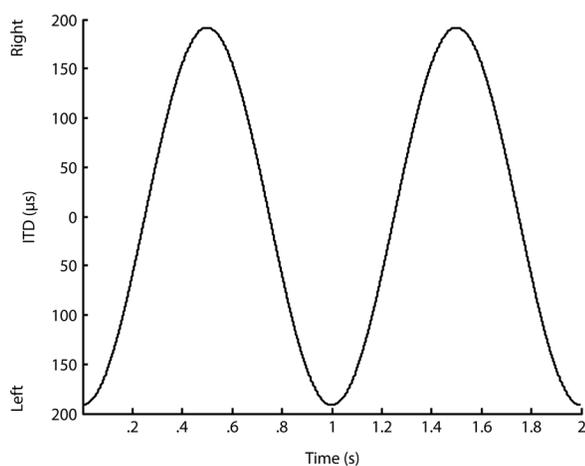


FIG. 1. Example of dynamic ITDs as a function of time when  $\Delta f = 0.06\%$  and  $f_m = 1$  Hz. The black line corresponds to the ITD at each point in time for a dichotic FM tone when  $\Delta f = 0.06\%$ , the average slow dichotic FMDL across all subjects. Note that whether the tone began as a left-lateralized percept or a right-lateralized percept depends on the starting phase of the modulator.

## C. Procedures

Participants completed ten tasks across 2–3 sessions, with a maximum duration of 2 h per session. In order to avoid fatigue, participants were instructed to take breaks as needed. All participants ran the tasks in the same order, as is typical of individual-difference paradigms (e.g., Kidd *et al.*, 2007). All tasks used a two-alternative forced-choice paradigm with a three-down, one-up adaptive procedure, converging on the 79.4% correct point of the psychometric function (Levitt, 1971). The target was randomly presented in either the first or second interval, and participants clicked a virtual button on the computer screen corresponding to the interval that they thought contained the target (i.e., “1” or “2”). Feedback was presented after each response, indicating whether the response was “correct” or “incorrect.”

All FM and AM tasks, dichotic and diotic, had a 500-ms inter-stimulus-interval (ISI). The slow ( $f_m = 1$  Hz) condition was always run before the fast ( $f_m = 20$  Hz) condition. For all FM and AM tasks, participants completed three adaptive runs. For each run, threshold was defined as the geometric mean of the tracking values at the last six reversal points. If the standard deviation across the runs was greater or equal to a specified criterion (.4 and .2 log units of the maximum frequency excursion and the modulation depth, respectively), participants completed three additional runs, and the first three runs were regarded as practice. In order to discount learning effects, only the last three runs were included in analyses. About 8% of conditions resulted in the completion of additional runs. All subsequent FM and AM tasks used this same criterion to help control for learning effects. The absolute thresholds and forward-masking patterns used a slightly different procedure to discount learning effects, as described below. The procedures for each task are described below in the order in which they were presented to subjects.

### 1. Tasks 1 and 2: Dichotic FM disparity

First, participants completed the slow-rate (1-Hz) dichotic FM disparity detection task. Participants were instructed that they would hear two tones, one at a time, and their task was to pick the tone that sounded as though it was “moving in their head.” They were reminded to look at the screen throughout the task, as they would receive visual feedback based on their response. In order to perceive lateralization and avoid confusion, the peak-to-peak frequency change must be sub-threshold but high enough for running phase to be accurately coded. Thus, each run began with a frequency excursion from the carrier ( $\Delta f$ ) of 0.2%, slightly below most FMDLs. The maximum value of the tracking variable was  $\Delta f = 1\%$ , as pilot data indicated that lateralized percepts were no longer salient with larger  $\Delta f$ s. If the maximum value was reached for more than ten consecutive trials, no threshold was recorded and listeners had to repeat three additional runs. One listener was not able to perform this task, and needed a higher starting value. This listener was able to perform the task with a starting value set to  $\Delta f = 0.6\%$ .<sup>1</sup> Initially,  $\Delta f$  varied by a factor of 2. After the first two reversals, the step size was reduced to a factor of 1.4 for the following two reversals, and was then set to the

final step size of 1.19 for the last six (measured) reversals. All subsequent FM tasks used the same series of step sizes.

Second, subjects completed the fast-rate (20-Hz) dichotic FM disparity detection task. Participants were instructed to pick the tone that had the “broader auditory image.” Again, participants were reminded to look at the feedback after each trial to help them decide how to identify the target tone. The starting value was set to  $\Delta f = 1\%$ , based on pilot data, with  $\Delta f$  never exceeding 100% throughout each run.

## 2. Tasks 3 and 4: FM detection

For both slow and fast FM detection, participants were instructed to pick the tone that was modulated and that the modulated tone will sound like it is “changing.” The initial value of the tracking variable was set to  $\Delta f = 2.51\%$  and never exceeded  $\Delta f = 100\%$ .

## 3. Task 5: Absolute threshold

Absolute thresholds were measured for all signal frequencies tested in the forward-masking task: 400, 430, 460, 490, 510, 540, 570, and 600 Hz. Participants completed one adaptive run at each signal frequency, with the signal frequency randomized between runs. The duration of the signal was the same as in the forward-masking task: 20 ms, including 10-ms onset and offset ramps (no steady state). Initially, the signal was presented at 40 dB SPL, and the initial step size was 8 dB. After two reversals, the step size was reduced to 4 dB and then to the final step size of 1 dB after two more reversals. Absolute threshold for each signal frequency was defined as the mean of the last six reversal points at the final step size. Participants were instructed to determine whether the first or second time interval, marked by lights on the virtual response box on a computer screen, “had a click in it.” The duration of each time interval was designed to be analogous to the forward-masking task. Each interval began with 500 ms of silence, followed by either a 20-ms signal (target interval) or 20 ms of silence (reference interval). The two intervals were separated by 400 ms of silence. If the standard deviation of the six reversal points within any of the runs was  $\geq 4$  dB, then one more run was completed at the corresponding signal frequency. At least 1 additional run was obtained in 23 of the 100 participants. Of the original runs, 3.4% were repeated. In the event that additional runs were needed from more than one signal frequency, the order of the additional runs was randomized.

## 4. Tasks 6 and 7: Dichotic AM disparity detection

Instructions for the slow (1-Hz) dichotic AM disparity task were identical to the slow dichotic FM disparity task. The initial modulation depth, in units of  $20 \log(m)$ , was  $-8$  dB. The step size was 6 dB for the first two reversals, and was 2 dB for the next two reversals, until the final step size of 1 dB was reached for the final six reversals. Threshold was defined as the mean depth at the last six reversal points.

Task instructions for fast (20-Hz) dichotic AM disparity were the same as the fast dichotic FM disparity task. Other

than the instructions, the procedures were identical to those used for the slow dichotic AM task.

## 5. Tasks 8 and 9: AM detection

Participants were instructed to pick the modulated (i.e., “changing”) tone. Otherwise, all procedures were the same as in the dichotic AM tasks.

## 6. Task 10: Forward-masking patterns

The 500-Hz masker was presented in both intervals of a trial, and the 20-ms signal was presented in one. Participants were instructed to pick the interval that had the “click” following the tone. The ISI was the same as in the absolute threshold task. At the beginning of each run, the signal level was 60 dB SPL. The initial step size of the adaptive procedure was 8 dB. After two reversals, the step size was decreased to 4 dB for the following two reversals, before reaching its final value of 2 dB for the final six reversals. Threshold was defined as the mean signal level at the last six reversal points.

Participants completed 2 runs for each of the 8 target frequencies, totaling 16 runs, and the order of the runs was randomized. If the standard deviation across the runs for any of the signal frequencies was  $\geq 4$  dB, then participants completed two more runs for the corresponding signal frequency. At least 1 additional run was obtained in 50 of the 100 participants. Of the original runs, 11.6% were repeated. In the event that participants had to repeat runs for two or more signal frequency conditions, the order of subsequent runs was also randomized.

## III. RESULTS

### A. Comparisons of performance in FM and AM tasks

Results in the FM and AM tasks are presented as box-plots in Fig. 2. Repeated-measures analyses of variance (ANOVAs) were conducted on the log-transformed thresholds [ $10 \log(\% \Delta f)$  and  $20 \log(m)$ ] for all (diotic and dichotic) FM and AM tasks. For FM, a  $2 \times 2$  within-subjects ANOVA revealed a main effect of modulation rate (1 Hz vs 20 Hz) [ $F(1,99) = 825$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.893$ ], a main effect of task-type (diotic vs dichotic) [ $F(1,99) = 216$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.686$ ], and a significant interaction [ $F(1,99) = 457$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.822$ ]. *Post hoc* Bonferroni-corrected *t*-tests ( $\alpha = 0.0083$ ) indicated significant differences between all comparisons except for diotic and dichotic fast FM tasks ( $p = 0.312$ ). As expected, thresholds for slow dichotic FM were substantially and significantly smaller (better) than thresholds for slow diotic FM ( $p < 0.0001$ ), indicating that slow dichotic FM disparity detection was based on the dynamic ITDs that were not available in the diotic conditions. The average threshold for slow dichotic FM is  $\Delta f = 0.06\%$ , which corresponds to a maximum instantaneous ITD of  $192 \mu\text{s}$  (see Fig. 1).

Analyses of the AM results were conducted using a  $2 \times 2$  (modulation rate vs task-type) within-subjects ANOVA. There was a main effect of modulation rate [ $F(1,99) = 127$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.562$ ], a main effect of task-type [ $F(1,99) = 354$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.782$ ], and a significant interaction [ $F(1,99) = 166$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.626$ ]. Differences between

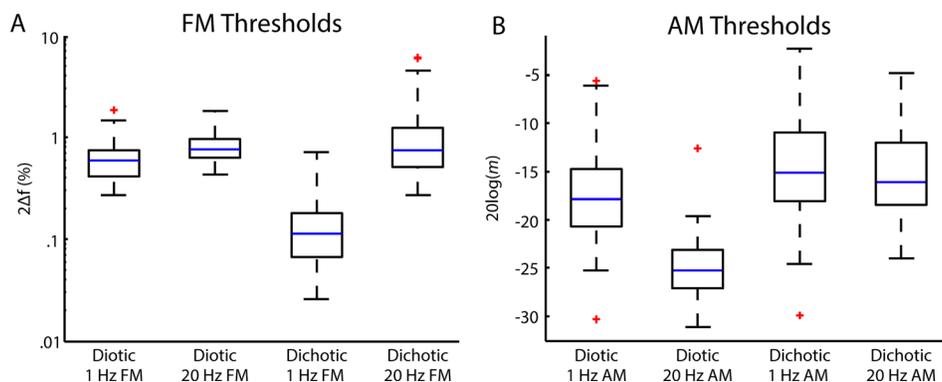


FIG. 2. (Color online) Boxplots for diotic and dichotic (A) FM detection and (B) AM detection thresholds across all participants. The two boxes closest to the y axis represent performance on diotic FM (A) and diotic AM (B) tasks. Center blue lines within each box represent the median of each group. Whiskers correspond to the lowest and highest data points within 1.5 times the lower and higher inter-quartile ranges, respectively. Small crosses represent individual data points outside the range of the whiskers, considered outliers.

AM tasks were examined using *post hoc* Bonferroni-corrected *t*-tests. All pair-wise comparisons were significant except for slow vs fast dichotic AM ( $p = 0.138$ ). Consistent with previous findings with gated carriers (Viemeister, 1979; Sheft and Yost, 1990; Moore and Sek, 1995), AMDLs were significantly better for fast AM detection compared to slow AM detection ( $p < 0.0001$ ). This effect has been ascribed to the effects of gating stimuli with low modulation rates, where the duration of the stimulus is only a small number of modulation cycles (two in the case of our 1-Hz modulation rate). In addition, slow diotic AM detection was significantly better than slow dichotic AM detection ( $p < 0.0001$ ), and fast diotic AM detection was significantly better than fast dichotic AM detection ( $p < 0.0001$ ). Thus, for AM (but not for FM), listeners were more sensitive to the detection of modulation than to the discrimination of interaural differences in modulation.

## B. Within-subjects vs between-subjects variance

As the analyses described below are correlational, it is important to examine the within-subjects vs the between-subjects variance across each of the modulation tasks. This is because correlations will be limited if the within-subjects variance is high relative to the between-subjects variance (Altman and Bland, 1983). The within-subjects variance was calculated by taking the pooled estimated variance across all three runs for all of the subjects; this is equivalent to the mean-squared error from a one-way, repeated-measures ANOVA in which run is the independent variable and threshold is the dependent variable. The square root of the within-subjects variance (i.e., the within-subjects standard deviation, SD) was compared to the between-subjects SD for each of the modulation tasks, listed in Table I. The ratio of between- vs within-subjects SD ranged from a factor of 2.63 to 1.4, indicating that the variance across subjects was greater than the variance within subjects.

To estimate the highest possible correlation our methods are capable of detecting, we calculated the average correlation based on 100 000 simulated test-retest correlations. First, six runs (three “test” and three “retest” runs) were sampled from each individual subject’s estimated distribution, based on their actual mean and standard deviation for a

given modulation task. Next, a simulated test-retest correlation was calculated using the average simulated test and retest mean for each subject. This iteration was completed 100 000 times, producing 100 000 simulated test-retest correlations. The test-retest correlations were transformed using Fisher’s *r* to *z* transformation, averaged, and then the average was transformed back to *r*. This process was completed for all modulation tasks, with the average simulated correlations ranging from  $r = 0.96$  for slow dichotic AM to  $r = 0.86$  for fast diotic FM (average across conditions was  $r = 0.92$ ). The high average simulated test-retest correlations indicate the ratio of between-subjects to within-subjects variance should be large enough for our correlations between tasks to be sensitive to individual differences between subjects.

## C. Correlational analyses of FM and AM thresholds

We expected all diotic tasks to correlate with their dichotic counterpart, as monaural processing of TFS or envelope cues should be related to performance on binaural tasks that utilize these same cues. As predicted, slow diotic FM thresholds correlated positively with slow dichotic FM thresholds ( $r = 0.42$ ,  $p < 0.0001$ ), and fast diotic FM thresholds correlated positively with fast dichotic FM thresholds ( $r = 0.54$ ,  $p < 0.0001$ ); see Fig. 3(A). Correlations for the AM data were similar to those found for the FM data: slow diotic AM correlated with slow dichotic AM ( $r = 0.57$ ,

TABLE I. Between- and within-subjects standard deviation for each modulation task. Ratio represents the ratio of the between- and within-subjects SD.

Task	Between-subjects SD	Within-subjects SD	Ratio
Dichotic AM (1 Hz)	5.61	2.13	2.63
Diotic AM (1 Hz)	4.64	2.26	2.05
Dichotic AM (20 Hz)	4.44	2.26	1.97
Dichotic FM (1 Hz)	3.38	1.71	1.98
Dichotic FM (20 Hz)	3	1.69	1.78
Diotic AM (20 Hz)	2.95	1.66	1.77
Diotic FM (1 Hz)	1.84	1.24	1.48
Diotic FM (20 Hz)	1.46	1.04	1.4

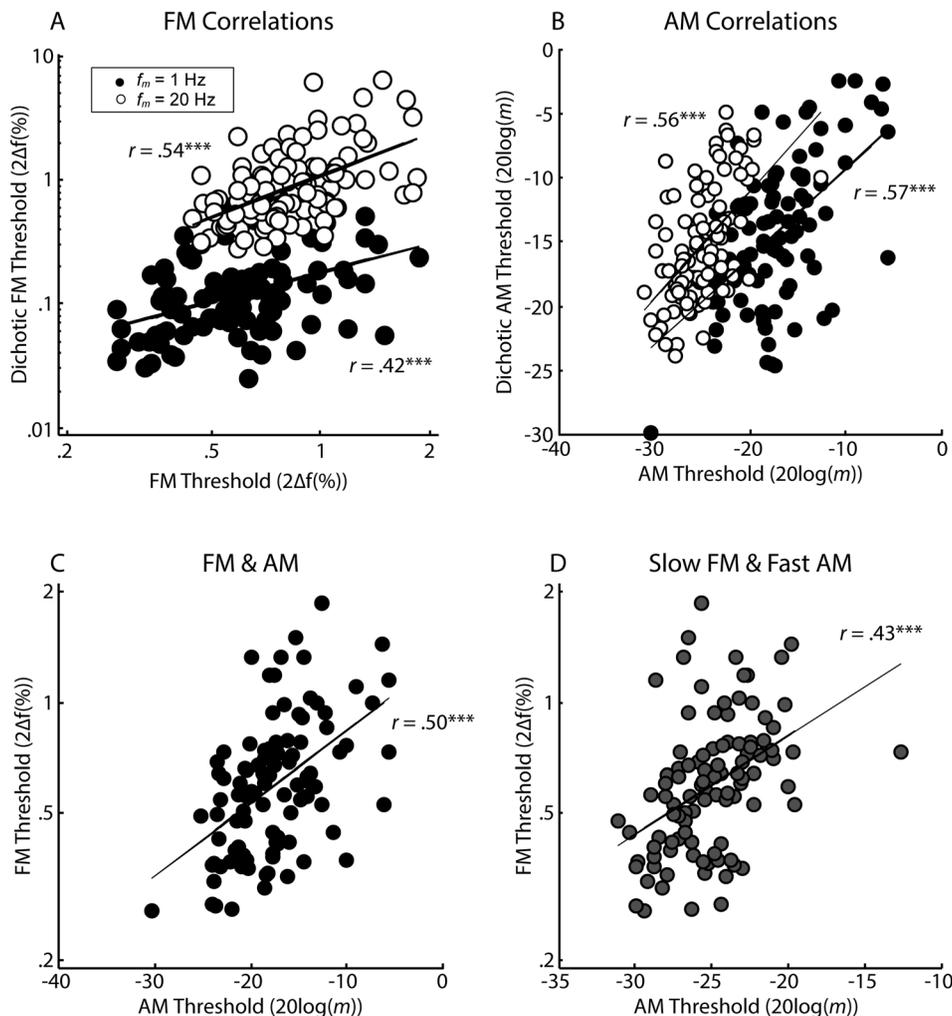


FIG. 3. Correlations between diotic and dichotic (A) FM and (B) AM. Black circles correspond to  $f_m = 1$  Hz, and white circles correspond to  $f_m = 20$  Hz. Grey circles (D) correspond to different modulation rates on the x and y axes. The black lines are the lines of best fit. For (A), both the x and y axes are plotted in peak-to-peak frequency change [ $2\Delta f(\%)$ ], where  $\Delta f$  is the frequency excursion from the carrier (in percent). (B) The x and y axes are plotted in  $20 \log(m)$ , where  $m$  is the modulation index. (C) plots the correlation between slow diotic FM detection and slow diotic AM detection, while (D) plots the correlation between slow diotic FM and fast diotic AM. “\*\*\*” indicates correlations that were highly significant ( $p < 0.0001$ ).

$p < 0.0001$ ) and fast dichotic AM correlated with fast diotic AM ( $r = 0.56$ ,  $p < 0.0001$ ); see Fig. 3(B).

Taken at face value, the strong correlation between slow diotic FM thresholds and slow dichotic FM thresholds could be interpreted as support for the hypothesis that phase-locked temporal information underlies performance in both tasks. A similarly strong correlation was also observed between the fast AM thresholds and fast FM thresholds for both diotic ( $r = 0.5$ ,  $p < 0.0001$ ) and dichotic ( $r = 0.5$ ,  $p < 0.001$ ) conditions, as would be expected if fast FM were detected via FM-to-AM conversion by cochlear filtering (e.g., Zwicker, 1970; Saberi and Hafter, 1995). Unfortunately, our results do not provide support for the dichotomy between temporal coding for slow FM and rate-place coding for fast FM, because most of the other modulation thresholds were also correlated with each other. In particular, if the correlation between diotic and dichotic slow FM thresholds reflects a common underlying (temporal) mechanism, then we would expect the correlations between thresholds for stimuli that do not share the same underlying mechanism to be lower. In fact, essentially all the modulation detection tasks were highly correlated with each other. For instance, slow FM and fast FM thresholds were correlated ( $r = 0.56$ ,  $p < 0.0001$ ), as were slow FM and slow AM ( $r = 0.5$ ,  $p < 0.0001$ ) and slow FM and fast AM ( $r = 0.43$ ,  $p < 0.0001$ ) [see Figs. 3(C) and 3(D)], despite the fact that these pairs are often regarded as being coded by

different peripheral mechanisms. Thus, our results provide no clear support for the idea that performance in slow FM detection tasks is limited by different mechanisms than performance in fast FM or AM detection tasks.

#### D. Frequency selectivity and FM detection

Mean thresholds (and standard deviations across the 100 subjects) for detection of the 20-ms signal in quiet and in the presence of the 500-Hz forward masker are shown in Fig. 4. Mean absolute thresholds for each subject were obtained by averaging thresholds across the eight signal frequencies, and mean masker effectiveness was estimated for each subject by averaging all eight forward-masked thresholds. Frequency selectivity was estimated for each subject by calculating the slope of the masking functions below and above the masker frequency separately using masker threshold as a function of the frequency separation of the masker and target in octaves. The linear regression resulted in slope estimates in units of dB/octave below and above 500 Hz. Boxplots of the lower and upper slopes of the forward-masking pattern are presented in Fig. 5. As expected, based on numerous studies of frequency selectivity (e.g., Patterson, 1976; Glasberg and Moore, 2000; Shera *et al.*, 2002), the median slopes were relatively steep and the slope of the lower side of the masking pattern

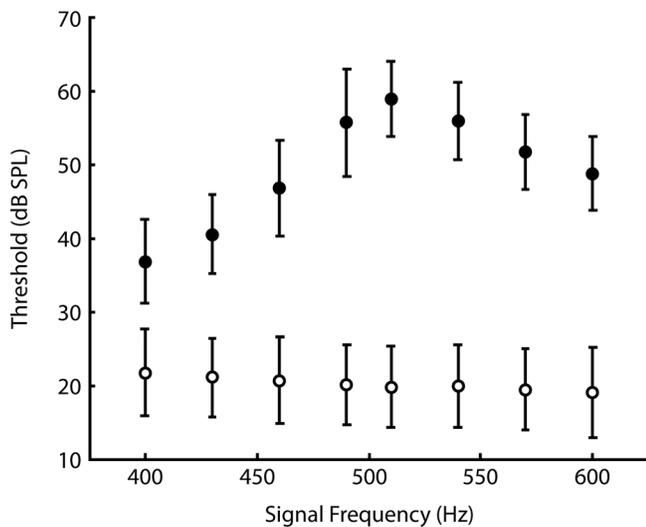


FIG. 4. Average detection thresholds as a function of signal frequency. Open circles represent the average absolute threshold for each of the signal frequencies when no masker is present. Filled circles represent the average detection threshold for each of the signal frequencies when preceded by a 500-ms, 500-Hz pure-tone masker. Error bars correspond to standard deviations across the 100 subjects.

was significantly steeper than the slope of the higher side (paired *t*-test;  $t = 39.3$ ,  $p < 0.0001$ ).

If fast-rate FM detection relies on detecting the AM induced by passing the FM stimulus through the auditory filters, then FM thresholds should be predicted by the combination of sensitivity to AM and the auditory filter slopes. Specifically, the FMDLs should approximate the smallest detectable change at the output of the characteristic frequency filter, divided by the slope of that filter (Zwicker, 1956; Moore and Glasberg, 1986; Lacher-Fougère and Demany, 1998). Predicted fast- and slow-rate FM thresholds were based on the individual subjects' fast- and slow-rate AM thresholds and their steeper masking-pattern slopes,

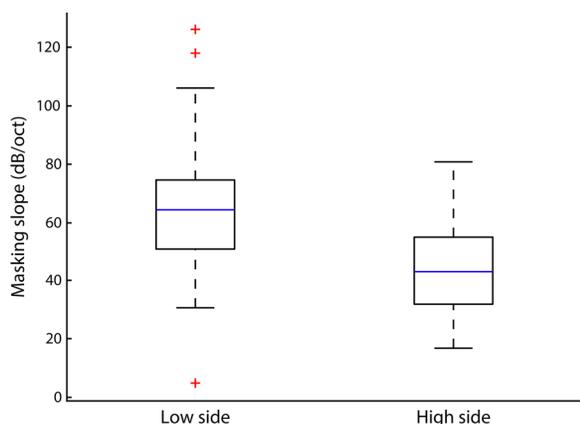


FIG. 5. (Color online) Boxplots of slopes from the forward-masking patterns. The low-side values represent the estimated slope of the masking pattern below the masker frequency. The high-side values represent the absolute value of the estimated slope of the masking pattern above the masker frequency. Center blue lines represent the median of each group. Whiskers correspond to the lowest and highest data points within 1.5 times the lower and higher inter-quartile ranges, respectively. Small crosses represent individual data points outside the range of the whiskers, considered outliers.

which was the lower slope for 88 of the 100 subjects. Correlations between the measured and predicted FMDLs were significant for both slow ( $r = 0.41$ ,  $p < 0.0001$ ) and fast ( $r = 0.38$ ,  $p < 0.0001$ ) modulation rates, although the magnitude of this effect was moderate. Again, at face value, the result appears to indicate that frequency selectivity is related to both slow and fast FM detection, but that frequency selectivity does not explain the majority of the inter-individual variance in FM detection. Again, there was no clear difference between the correlations for slow- and fast-rate FM thresholds, inconsistent with idea that fast- and slow-rate thresholds are governed by different mechanisms. Most importantly, these moderate correlations between predicted FM and measured FM are confounded by the high correlations between slow FM and slow AM ( $r = 0.5$ ) and fast FM and fast AM ( $r = 0.5$ ). Because predicted FM thresholds are calculated based on AM sensitivity divided by the steeper filter slope, and FM and AM are well correlated, a correlation between measured and predicted FM would likely be present regardless of the steepness of the filter slopes. In fact, correlations between predicted and actual FM thresholds are actually lower than just the raw correlations between FM and AM thresholds, suggesting that adding the filter slopes provides no additional information to the predictions. Thus, the correlations between measured and predicted FM are driven by the high correlations between AM and FM, rather than the individual differences in frequency selectivity.

Although the correlations between measured and predicted FM are clearly driven by the correlations between AM and FM, rather than masking-pattern slopes, the group averages between measured and predicted FMDLs may still provide useful information. A  $2 \times 2$  (threshold type vs modulation rate) within-subjects ANOVA was conducted on the log-transformed thresholds for predicted and measured FMDLs [i.e.,  $10 \log(\% \Delta f)$ ]. Results indicated a main effect of threshold type [ $F(1,99) = 492$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.833$ ], a main effect of modulation rate [ $F(1,99) = 76.9$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.437$ ], and a significant interaction [ $F(1,99) = 423$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.81$ ]. A *post hoc t*-test indicated that predicted slow FMDLs ( $M = 1.1$ ,  $SD = 2.43$ ) were significantly larger (i.e., poorer) than measured slow FMDLs ( $M = -5.27$ ,  $SD = 1.84$ ) ( $p < 0.0001$ ). Consistent with previous literature (e.g., Moore and Glasberg, 1986; Lacher-Fougère and Demany, 1998), rate-place information alone, based on the single largest change in the excitation pattern, far underestimates listeners' actual ability to detect slow FM. More surprisingly, a *t*-test showed a similar trend between predicted fast FMDLs ( $M = -2.7$ ,  $SD = 1.77$ ) and measured fast FMDLs ( $M = -3.94$ ,  $SD = 1.46$ ) ( $p < 0.0001$ ), although the difference between the predicted and measured means for fast FMDLs was smaller.

If FM relies on excitation pattern information, then FM should be related to the steepness of the auditory filter slopes. However, neither the low slope, high slope, nor the overall steepness of both filter slopes (calculated as the low slope summed with the absolute value of the high slope) were correlated with slow diotic FM (low slope:  $r = 0.17$ ; high slope:  $r = -0.14$ ; overall steepness:  $r = 0.2$ ) using a one-tailed test. In fact, each of the correlations was opposite

of the predicted direction, as steeper low slopes are positive (bigger numbers), which should be negatively related to better (smaller) FMDLs, and steeper high slopes are negative (smaller numbers), which should be positively related to better FMDLs. There was no correlation between fast diotic FM and filter slopes (low slope:  $r = 0.04$ ; high slope:  $r = -0.04$ ; overall steepness:  $r = 0.05$ ).

It is possible that the correlations between diotic FM and filter slopes are unobservable because FM thresholds are overshadowed by variability in sensitivity to AM, assuming FM is converted to AM in the cochlea. In order to account for sensitivity to AM, both slow-rate and fast-rate FMDLs and AMDLs were  $z$ -transformed so that they were on the same scale. The  $z$ -transformed AMDLs were subtracted from FMDLs at the corresponding modulation rate. These difference scores were then correlated with the  $z$ -transformed filter slopes. Correlations were conducted between the difference scores and the low slope, high slope, and overall steepness for both slow and fast FM. Although the correlations between slow difference scores and filter slopes were in the predicted direction, they were not significant (low slope:  $r = -0.02$ ,  $p = 0.42$ ; high slope:  $r = 0.12$ ,  $p = 0.12$ ; overall steepness:  $r = -0.08$ ,  $p = 0.21$ ). The correlations between fast difference scores and frequency selectivity were slightly better, but still very weak (low slope:  $r = -0.09$ ,  $p = 0.19$ ; high slope:  $r = 0.2$ ,  $p = 0.02$ ; overall steepness:  $r = -0.17$ ,  $p = 0.045$ ), with the high slope and overall steepness reaching significance without correcting for multiple comparisons. Assuming the correlations between slow-rate difference scores and the overall steepness is reflective of the true population, one would need a sample size of  $n = 427$  to reach significance using a one-tailed test. Overall, there was very little evidence for a relationship between either slow- or fast-rate FM and frequency selectivity, even when controlling for sensitivity to AM.

### E. Principal components analysis

Given the large number of measures involved in our study, we conducted a principal components analysis (PCA), using the average thresholds for all 100 participants on each of the tasks. The average absolute threshold across the signal frequencies was included in the PCA analysis, as listeners with good sensitivity would have, on average, lower absolute thresholds across the signal frequencies. Similarly, the average forward-masking threshold across the signal frequencies

was included as a measure of the overall effect of masking. Listeners with a higher overall effect of masking would have, on average, higher thresholds across all of the frequencies, regardless of slope. The overall steepness of the filter slopes from the forward-masking patterns task was included as a measure of frequency selectivity.

PCA is an important exploratory analysis to conduct because it could reveal a different structure in the dataset that is not intuitively obvious from the full correlation matrix. This is because PCA takes into account the relationship of each task with every other task when performing the dimension reduction. With 55 possible correlations, the dataset is too large to safely intuit the multivariate structure by simply inspecting the correlation matrix. In addition, PCA should isolate any common variance across conditions based, for instance, on “attentiveness,” or other non-sensory factors that may be shared by many or all of the measures. Based on our initial hypotheses, the PCA should produce separate components that reflect peripheral rate-place coding (i.e., tonotopy) and time coding (i.e., phase-locking to TFS cues). For example, if slow FM is coded via phase-locking to TFS cues, but fast FM is not, then slow diotic and dichotic FM would load onto one component. Slow and fast AM, diotic and dichotic, would load onto a second component, reflecting sensitivity to envelope cues. If fast FM is converted to AM via cochlear filtering, then fast FM would load onto a component with frequency selectivity as well as fast AM. Because some of the tasks were measured in different units [e.g., dB/oct for filter slopes vs  $20 \log(m)$  for amplitude modulation], PCA was conducted using an eigen-decomposition on the correlation matrix. Because PCA was performed on the correlation matrix, the amount of variance accounted for by each component reflects the variance accounted for when each task has a standardized variance ( $s^2 = 1$ ). This ensures that tasks measured in larger units (with, consequently, arbitrarily larger variances) do not dominate the component loadings.

The results from the PCA did not reflect our predicted results. The PCA (varimax rotated) produced three interpretable components, accounting for  $\sim 63.3\%$  of the standardized variance (Fig. 6). Component 4 only accounted for an additional 11% of the standardized variance, and was not clearly interpretable, so was not included in the analysis.

All of the FM and AM, diotic and dichotic, tasks loaded onto the first component, which accounted for 32% of the

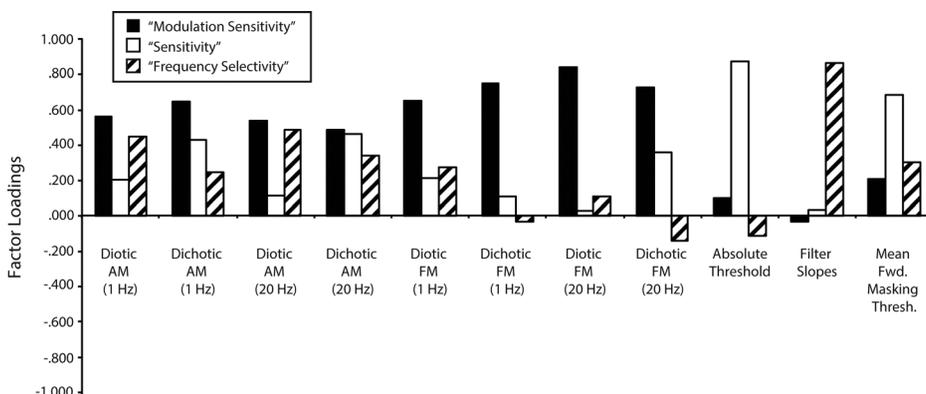


FIG. 6. PCA suggesting that three factors account for the majority of the variance. The  $x$  axis groups tasks based on the component for which they had the greatest factor loading. Solid bars correspond to the “modulation sensitivity” component, open bars correspond to the “sensitivity” component, and striped bars correspond to the “frequency selectivity” component.

standardized variance. Thus, component 1 was named the “modulation sensitivity” component, as it appears to reflect a general ability to perform AM- and FM-related tasks in both diotic and dichotic situations at both slow and fast rates. The names of components 2 and 3 were determined based on the task that loaded most onto the components. Component 2 (accounting for 17% of the standardized variance) had the highest loadings by average absolute and average forward-masked threshold, and so was termed “sensitivity,” while component 3 (accounting for 14.3% of the standardized variance) most strongly reflected filter slopes and so was termed “frequency selectivity.”

Consistent with the earlier analysis based on paired correlations, the PCA provided no evidence for separable coding mechanisms reflecting phase-locking for slow FM and rate-place coding for fast FM. Frequency selectivity appeared to be related to neither, while binaural sensitivity to temporal fine structure (as reflected in the slow dichotic FM thresholds) was equally related to diotic FM, as well as AM, at both slow and fast rates. Although the PCA reiterates the observed patterns in the correlational analyses, it provides a parsimonious description of the dataset and confirms that no other interpretable structure appears to exist within the dataset.

## IV. DISCUSSION

### A. Summary of results

The aim of this study was to use individual differences in a large cohort of young normal-hearing listeners to test the hypothesis that the coding of slow-rate FM is based on temporal (phase-locked) information, whereas the coding of fast-rate FM is based on rate-place information, through the transformation of FM to AM via peripheral auditory filtering. Strong correlations were observed between most of the modulation detection and discrimination tasks. The two main findings were not consistent with the predictions of the hypothesis. First, the correlation between a measure known to reflect timing information (dichotic FM disparity detection) was not more strongly correlated with slow FM detection than with any of the other monaural (or binaural) modulation detection tasks. Second, the measure of frequency selectivity combined with the measure of AM sensitivity did not predict fast-rate (or slow-rate) FM thresholds any better than just the measure of AM sensitivity, suggesting no clear relationship between frequency selectivity and either slow- or fast-rate FM detection thresholds. An exploratory PCA approach resulted in the same conclusions: the diotic and dichotic, slow- and fast-rate, FM and AM detection thresholds were all related to one another, but were generally unrelated to measures of absolute threshold, masked threshold, or frequency selectivity.

### B. Comparisons with previous studies

One concern is that the large number of subjects precluded extended practice in any of the tasks before measurement. It may be, therefore, that the thresholds obtained by our listeners do not reflect the sensory limits of FM or AM

detection, but rather reflect more cognitive or procedural limitations that might have been overcome by further training. To test for this possibility, we compared our thresholds with those reported in the literature from smaller, more practiced, groups of subjects.

In general, although listeners completed only 3–6 runs for each condition, average diotic FMDLs and AMDLs are comparable to those from well-trained listeners in the psychoacoustic literature. For example, the average peak-to-peak frequency change ( $2\Delta f$ ) across all 100 participants was 0.6% for slow FM and 0.81% for fast FM. Using three well-trained listeners with a 500-Hz carrier and similar slow and fast modulation rates, the average FMDLs for Moore and Sek (1996) were 0.9% for  $f_m = 2$  Hz and 1.3% for  $f_m = 20$  Hz. It is possible that our FMDLs may be slightly better than the trained listeners in the earlier study because the durations of our stimuli were twice as long (2 s, as opposed to 1 s). Demany and Semal (1989) also used a 2-s duration, trained their participants until thresholds were stable, and obtained similar average FMDLs [ $f_m = 1$  Hz:  $M = 0.732\%$ ;  $f_m = 16$  Hz:  $M = 0.902\%$ ]. Notably, our listeners were first exposed to FM and AM tones via the dichotic tasks, so they were not completely untrained with respect to exposure to FM or AM.

Our AMDLs are not as straightforward to compare with values in the literature, as most previous studies have used a different carrier frequency, modulation rate, and/or methods of measurement (e.g., constant stimuli procedures to calculate  $d'$ ). To make comparisons across studies, measures of  $d'$  from previous literature were transformed to approximate the 79% correct point. AMDLs are reported in  $20 \log(m)$ . All studies reported used a 2-interval, 2-alternative forced choice procedure. The average AMDLs across the 100 listeners from the current study were  $-17.6$  dB for slow AM and  $-25.0$  dB for fast AM. This is roughly comparable to that found by Moore and Sek (1996) using a higher carrier frequency (e.g.,  $M = -23.0$  dB for  $f_c = 1$  kHz and  $f_m = 2$  Hz). With the same 1-kHz carrier and three listeners with “extensive practice,” Moore and Sek (1995) reported average AMDLs of  $\sim -17.7$  dB for  $f_m = 2$  Hz and  $-26.0$  dB for  $f_m = 10$  Hz, very similar to our mean results.

To our knowledge, only one other study has measured dichotic FMDLs (Grose and Mamo, 2012). In this study, young listeners trained on dichotic FM until thresholds “appeared stable,” and stable performance was achieved with just an average of 1.6 practice runs. Grose and Mamo (2012) used a similar low carrier (frequency roved:  $460 \leq f_c \leq 540$ ) and slow modulation rate ( $f_m = 2$  Hz), and obtained comparable thresholds ( $2\Delta f = 0.17\%$  in Grose and Mamo vs  $2\Delta f = 0.12\%$  in the present study). There are several methodological differences that may account for better thresholds in our study: Grose and Mamo (2012) (1) roved the carrier frequency, (2) had a pure tone reference instead of an FM tone as a reference, and (3) used a three-interval task as opposed to a two-interval task.

In summary, despite the relatively small amount of practice provided to our subjects, the average FM and AM thresholds obtained in our study are very comparable to those reported in earlier studies. It seems unlikely, therefore,

that the lack of clear differences based on underlying coding mechanisms reflects generally poor performance on the part of our subjects.

### C. Analyses of subsets of data

Another way to address the potential effects of generally poor performance is to examine the results from a subset of better performers. The rationale is that the subjects with the lowest thresholds are most likely to have reached their sensory limits and so are more likely to reflect variance based on sensory limitations. To test this hypothesis, we reanalyzed the data from the “best” 30 and “worst” 30 listeners, based on their values for the first component in the PCA, and retested the idea that fast FMDLs should be predicted by a combination of fast AMDLs and masking slope. If the best 30 listeners’ results more closely reflect sensory limitations, then the correlations should be higher in that group than in the whole group. In fact, the correlations between measured and predicted FMDLs did not reach significance for either the best 30 listeners (slow FM:  $r = 0.15$ ,  $p = 0.21$ ; fast FM:  $r = 0.22$ ,  $p = 0.12$ ) or the worst 30 listeners (slow FM:  $r = 0.18$ ,  $p = 0.17$ ; fast FM:  $r = 0.09$ ,  $p = 0.32$ ). The lack of correlation can probably be explained by the reduced range of thresholds in the relevant modulation-detection tasks, but also suggests that the main findings of this study are not due solely to poor performers being limited by non-sensory factors.

### D. Variability in peripheral coding

Although the correlations between all FM and AM tasks are high, correlations between tasks thought to utilize the same peripheral coding mechanisms are not as high as one might expect given the supposed importance of phase-locking in slow FM detection and frequency selectivity in fast FM detection. One possible explanation is that the variability in peripheral coding in young, normal-hearing listeners is too small to exert a large influence on thresholds. The variability in our measure of phase-locking to TFS cues [expressed as  $2\Delta f(\%)$ ] is, in fact, more than a factor of 3 smaller than the variability observed in 12 older listeners from Grose and Mamo ( $SD = 0.516\%$  in Grose and Mamo (2012) vs  $SD = 0.163\%$  in the present study). It has been suggested that coding of TFS declines with both age (e.g., Hopkins and Moore, 2011; Grose and Mamo, 2012; Moore *et al.*, 2012) and degree of hearing loss (e.g., Hopkins and Moore, 2007, 2011; Lorenzi *et al.*, 2009), which would contribute to the increased variability of TFS coding.

In addition, it is well known that auditory filter slopes become shallower with sensorineural hearing loss (e.g., Glasberg and Moore, 1986), and the variability in bandwidth across hearing-impaired listeners is quite wide (e.g., Moore *et al.*, 1999). Although the variability in the steepness of the auditory filter slopes was quite large in our young, normal-hearing listeners (see Fig. 5), this between-subjects variability would certainly increase if the subject pool were expanded to include older and hearing-impaired listeners. Future studies including a large sample of different ages and degrees of hearing impairment may elucidate whether variability in

peripheral coding can be made large enough to outweigh other factors determining individual differences in performance.

### E. Limitations of correlational studies

Taken at face value and out of the context of the other results, the strong correlation between dichotic FM disparity detection and diotic FM detection at slow rates could have been interpreted as evidence that phase-locking to temporal fine structure dominates for slow FM detection. It was only the equally strong correlations between measures not thought to be related to phase-locking (such as fast-rate FM and AM detection) that cast doubt on this interpretation. Similar correlational analyses have become a popular method for examining questions of underlying neural coding in normal and impaired hearing, and as a function of age (e.g., Strelcyk and Dau, 2009; Ruggles *et al.*, 2011; Ochi *et al.*, 2014; Bharadwaj *et al.*, 2015). Caution is required in interpreting the results from such studies, as they rarely include measures of performance that are similar in task nature but are thought to reflect different underlying neural mechanisms. In other words, it can be important to provide measures that are *not* expected to be correlated with the others in order to demonstrate specificity of the putative mechanisms, and to ensure that the correlation does not reflect higher-level central processing that is not specific to particular underlying mechanisms. In a related domain, studies predicting FMDLs or frequency difference limens based on sensitivity to level changes and frequency selectivity need to account for possible high correlations between frequency-discrimination/detection thresholds and intensity discrimination/detection thresholds (Moore and Glasberg, 1986; Dai *et al.*, 1995).

### F. Explaining the high sensitivity to slow FM

Although our correlational approach provided no strong evidence for two distinct coding mechanisms for slow- and fast-rate FM, the fact remains that thresholds for slow-rate FM are generally lower (better) than fast-rate FM thresholds. In contrast, our slow-rate AM thresholds were considerably higher (worse) than the fast-rate AM thresholds. How can this difference be explained if both slow and fast FM detection are governed by the same underlying mechanisms?

One potential explanation lies in a recent solution to the long-standing problem of why sensitivity to intensity changes and frequency selectivity seems unable to account for sensitivity to frequency changes. Micheyl *et al.* (2013) proposed that pure-tone frequency-discrimination performance could be explained by a cortical population rate-place code, which could also explain human intensity discrimination performance. Their model relied on some correlation between spike counts of neurons with similar characteristic frequencies. This correlation resulted in a deterioration in intensity coding and a relative improvement in frequency coding, leading to reasonable predictions of thresholds in both dimensions. As spike correlations rely on a certain time window over which to count spikes, the effects of correlations between neural units will decrease with decreasing analysis duration. Thus, the relative benefit of neural correlations for

frequency coding will be observed more for long durations (or slow FM rates) than for short durations (or fast FM rates). This explanation may provide the basis for an account of the different sensitivity between slow and fast FM without the need for a separate neural code. Similarly, the decrease in frequency-discrimination abilities at high frequencies may reflect cortical coding limitations (perhaps based on the tonotopic distribution of responses), rather than peripheral limitations based on phase-locking (e.g., Oxenham *et al.*, 2011). However, further modeling work and physiological data are required to test this conjecture.

## ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R01 DC005216. We thank Jiao Xu and Lindsey Ma for assisting with data collection, and Matthew Scauzillo for assistance with creating the figures.

<sup>1</sup>Due to a programming error, five additional participants also began each run with  $\Delta f = 0.6\%$ .

- Altman, D. G., and Bland, J. M. (1983). "Measurement in medicine: The analysis of method comparison studies," *Statistician* **32**, 307–317.
- Attneave, F., and Olson, R. K. (1971). "Pitch as a medium: A new approach to psychophysical scaling," *Am. J. Psychol.* **84**, 147–166.
- Bharadwaj, X. H. M., Masud, S., Mehraei, G., Verhulst, S., and Shinn-Cunningham, B. G. (2015). "Individual differences reveal correlates of hidden hearing deficits," *J. Neurosci.* **35**, 2161–2172.
- Coninx, F. (1978a). "The detection of combined differences in frequency and intensity," *Acustica* **39**, 137–150.
- Coninx, F. (1978b). "The perception of combined frequency and amplitude modulations with clearly audible modulation depths," *Acustica* **39**, 151–154.
- Dai, H., Nguyen, Q. T., and Green, D. M. (1995). "A two-filter model for frequency discrimination," *Hear. Res.* **85**, 109–114.
- Demany, L., and Semal, C. (1989). "Detection thresholds for sinusoidal frequency modulation," *J. Acoust. Soc. Am.* **85**, 1295–1301.
- Edwards, B. W., and Viemeister, N. F. (1994a). "Modulation detection and discrimination with three-component signals," *J. Acoust. Soc. Am.* **95**, 2202–2212.
- Edwards, B. W., and Viemeister, N. F. (1994b). "Frequency modulation versus amplitude modulation discrimination: Evidence for a second frequency modulation encoding mechanism," *J. Acoust. Soc. Am.* **96**, 733–740.
- Festen, J. M., and Plomp, R. (1981). "Relations between auditory functions in normal hearing," *J. Acoust. Soc. Am.* **70**, 356–369.
- Festen, J. M., and Plomp, R. (1983). "Relations between auditory functions in impaired hearing," *J. Acoust. Soc. Am.* **73**, 652–662.
- Glasberg, B. R., and Moore, B. C. J. (1986). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.* **79**, 1020–1033.
- Glasberg, B. R., and Moore, B. C. J. (2000). "Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise," *J. Acoust. Soc. Am.* **108**, 2318–2328.
- Grose, J. H., and Mamo, S. K. (2012). "Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects," *Hear. Res.* **294**, 49–54.
- Hopkins, K., and Moore, B. C. J. (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., and Moore, B. C. J. (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.
- Johnson, D. H. (1980). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Am.* **68**, 1115–1122.
- Johnson, D. M., Watson, C. S., and Jensen, J. K. (1987). "Individual differences in auditory capabilities. I," *J. Acoust. Soc. Am.* **81**, 427–438.
- Kidd, G. R., Watson, C. S., and Gygi, B. (2007). "Individual differences in auditory abilities," *J. Acoust. Soc. Am.* **122**, 418–435.
- Lacher-Fougère, S., and Demany, L. (1998). "Modulation detection by normal and hearing-impaired listeners," *Audiology* **37**, 109–121.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., and Moore, B. C. J. (2009). "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal," *J. Acoust. Soc. Am.* **125**, 27–30.
- McDermott, J. H., Lehr, A. J., and Oxenham, A. J. (2010). "Individual differences reveal the basis of consonance," *Curr. Biol.* **20**, 1035–1041.
- Micheyl, C., Schrater, P. R., and Oxenham, A. J. (2013). "Auditory frequency and intensity discrimination explained using a cortical population rate code," *PLoS Comput. Biol.* **9**, e1003336.
- Moore, B. C. J. (1973). "Frequency difference limens for short-duration tones," *J. Acoust. Soc. Am.* **54**, 610–619.
- Moore, B. C. J., and Ernst, S. M. A. (2012). "Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code," *J. Acoust. Soc. Am.* **132**, 1542–1547.
- Moore, B. C. J., and Glasberg, B. R. (1986). "The relationship between frequency selectivity and frequency discrimination for subjects with unilateral and bilateral cochlear impairments," in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York), pp. 407–417.
- Moore, B. C. J., and Peters, R. W. (1992). "Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity," *J. Acoust. Soc. Am.* **91**, 2881–2893.
- Moore, B. C. J., and Sek, A. (1992). "Detection of combined frequency and amplitude modulation," *J. Acoust. Soc. Am.* **92**, 3119–3131.
- Moore, B. C. J., and Sek, A. (1994). "Effects of carrier frequency and background noise on the detection of mixed modulation," *J. Acoust. Soc. Am.* **96**, 741–751.
- Moore, B. C. J., and 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, B. C. J., and Sek, A. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.* **100**, 2320–2331.
- Moore, B. C. J., and 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, B. C. J., Vickers, D. A., and Mehta, A. (2012). "The effects of age on temporal fine structure sensitivity in monaural and binaural conditions," *Int. J. Audiol.* **51**, 715–721.
- Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (1999). "Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism," *J. Acoust. Soc. Am.* **106**, 2761–2778.
- Ochi, A., Yamasoba, T., and Furukawa, S. (2014). "Factors that account for inter-individual variability of lateralization performance revealed by correlations of performance among multiple psychoacoustical tasks," *Front. Neurosci.* **8**, 1–10.
- Oxenham, A. J. (2013). "Revisiting place and temporal theories of pitch," *Acoust. Sci. Technol.* **34**, 388–396.
- Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., and Santurette, S. (2011). "Pitch perception beyond the traditional existence region of pitch," *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7629–7634.
- 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.
- Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli," *J. Acoust. Soc. Am.* **59**, 640–654.
- Plack, C. J., Oxenham, A. J., Fay, R. R., and Popper, A. N. (2005). *Pitch: Neural Coding and Perception* (Springer, New York), pp. 7–56.
- 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.
- Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2011). "Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication," *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15516–15521.
- Saberi, K., and Hafter, E. R. (1995). "A common neural code for frequency- and amplitude-modulated sounds," *Nature* **374**, 537–539.

- Sek, A., and Moore, B. C. J. (1995). "Frequency discrimination as a function of frequency, measured in several ways," *J. Acoust. Soc. Am.* **97**, 2479–2486.
- Sheft, S., and Yost, W. A. (1990). "Temporal integration in amplitude modulation detection," *J. Acoust. Soc. Am.* **88**, 796–805.
- Shera, C. A., Guinan, J. J., and Oxenham, A. J. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3318–3323.
- Strelcyk, O., and Dau, T. (2009). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," *J. Acoust. Soc. Am.* **125**, 3328–3345.
- Tyler, R. S., Wood, E. J., and Fernandes, M. (1983). "Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **74**, 1190–1199.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.
- Watson, C. S., Qiu, W. W., Chamberlain, M. M., and Li, X. (1996). "Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition," *J. Acoust. Soc. Am.* **100**, 1153–1162.
- Zwicker, E. (1956). "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs" ("The elemental foundations for determining the information capacity of the auditory system"), *Acustica* **6**, 365–381.
- Zwicker, E. (1970). "Masking and psychological excitation as consequences of the ear's frequency analysis," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden, The Netherlands), pp. 376–394.