Hearing Research 290 (2012) 72–82

Contents lists available at SciVerse ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Effects of randomizing phase on the discrimination between amplitude-modulated and quasi-frequency-modulated tones

Hisaaki Tabuchi, Ewa Borucki, Bruce G. Berg*

Department of Cognitive Sciences, University of California, 2201 Social & Behavioral Sciences, Gateway Building, Irvine, CA 92697-5100, USA

ARTICLE INFO

Article history:
Received 26 September 2011
Received in revised form
10 April 2012
Accepted 23 April 2012
Available online 17 May 2012

ABSTRACT

This study investigated the bandwidth of phase sensitivity. Subjects discriminated amplitude-modulated tones (AM), and quasi-frequency-modulated tones (QFM) in a two-interval, forced-choice task. An adaptive threshold procedure was used to estimate the modulation depth needed to discriminate the stimuli as a function of carrier and modulation frequency. Non-monotonicities in threshold-bandwidth functions were often observed at higher modulation frequencies. The results are discussed in terms of two potential cues: (1) waveform envelope, (2) cubic distortion products. In order to degrade the information obtained from auditory distortions, the phase for the carrier frequency was randomly sampled from a uniform distribution, which diminished the non-monotonicities with minimal effect at lower modulation frequencies. Model simulations demonstrated that phase randomization degrades distortion product cues with only a modest effect on temporal cues. Final results show that maximum bandwidths for phase sensitivity (BW_{max}) were not proportional to carrier frequencies.

1. Introduction

Mathes and Miller (1947) were among the first to investigate phase sensitivity. They reported that an amplitude-modulated sinusoid (AM) and a quasi-frequency-modulated sinusoid (QFM) were audibly different when the stimulus bandwidth was narrow, but became indistinguishable when the ratio of modulation rate to center frequency exceeded 0.5. They proposed that auditory filters constitute the first stage of processing, an idea similar to Fletcher (1940), except that, rather than an integration of energy, "information" was transmitted by the temporal output of an auditory filter. According to their theory, AM and QFM can be discriminated only when the three sinusoids are within the frequency range of an auditory filter.

An AM/QFM discrimination task appears ideal for investigating the characteristics of a temporal processing system. First, the stimuli have the same long-term power spectra but different envelopes resulting from relative phase differences (i.e. QFM has an envelope modulation rate twice higher than that of AM and a lower modulation depth). Second, the stimuli are deterministic and may yield more precise psychophysical measurements than modulated noise. It was later recognized, however, that an interaction between the low tone of the stimulus and an internally generated cubic distortion tone (CDT) from the higher two tones might create a salient intensity cue (Goldstein, 1967a; Buunen, 1975). The primary intent of this study is to estimate the maximum bandwidth of phase sensitivity in an AM/QFM discrimination task while controlling for the potential effects of auditory distortion products on the measurements.

Only a few AM/QFM discrimination experiments have been conducted since the conjecture that distortion product cues might contaminate measurements of phase sensitivity and all have noted this potential limitation of the method. Nelson (1994) considered another proposal from Buunen (1976) that distortion products might alter the internal representation of the envelope in such a manner as to produce a CDT-induced envelope cue. He obtained psychometric functions as a function of modulation rate for AM/QFM discrimination and defined the modulation rate corresponding to $d' = 1$ as the "critical band for phase" (CB_{phs}).[1] With a 1000 Hz carrier at 40 dB SPL, the mean CB_{phs} was 118 Hz and ranged from 39 Hz to 219 Hz for 9 listeners. At 80 dB SPL, the mean was 414 Hz with a range of 285–525 Hz. Bernstein and Oxenham (2006) obtained psychometric functions for AM/QFM discrimination as a function of modulation rate, but used a lowpass noise to mask distortion tones. They also found an intensity effect for a 1500 Hz

* Corresponding author. Tel.: +1 949 824 7920.
E-mail address: bgberg@uci.edu (B.G. Berg).

[1] Nelson's (1994) measure, CB_{phs}, is on a scale of modulation rate, f_m. Given that QFM and AM with the same bandwidth have different modulation rates, current findings are reported in terms of stimulus bandwidth, $2f_m$.

0378-5955/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2012.04.021

carrier. CB_{phs} increased from 120 Hz to 210 Hz when the intensity was increased by 15 dB. Estimates of the maximum modulation rate were lower than Nelson's (1994), but this may be attributable to a procedural difference. Bernstein and Oxenham randomized the modulation frequency within each data-collecting run and this may have increased the difficulty of the task. Furthermore, the cutoff frequency for the lowpass masker was below the frequency of the CDT, so the extent to which the added noise had the intended effect is unclear. Increasing the cutoff frequency to mask the CDT would also mask the low frequency component of the stimulus, thereby affecting the envelope of the waveform. Nelson (1994) and Bernstein and Oxenham (2006) contend that the effect of level on the cutoff frequency of psychometric functions provides evidence that phase sensitivity is limited by the bandwidth of auditory filters.

Strickland and Viemeister (1997) minimized distortion product effects by using bands of AM noise and QFM noise which had the same power spectrum but different phase spectra. As an added precaution, the 800-Hz wide modulated noise was also placed in a notch of unmodulated masking noise. "Modulation rate" was defined by the width of the sidebands and the relative amplitude of the sidebands needed to discriminate AM noise from QFM noise was estimated with an adaptive procedure. Temporal modulation transfer functions (TMTFs; thresholds plotted as a function of modulation rate) were fit with a first-order Butterworth filter to obtain estimates of sensitivity (S), corresponding to the low frequency plateau of the TMTF at low modulation rates, and cutoff frequency (C_f). Estimates of C_f were essentially the same when the modulated noise was centered at 1–4 kHz, a finding confirmed by Eddins (1999) and Strickland (2000). The independence between C_f and center frequency was interpreted by Strickland and Viemeister (1997) as evidence that the rolloff of the TMTF manifests central limitations more so than limitations attributable to peripheral filtering. This implied that the limit to phase sensitivity is the maximum modulation rate that can be centrally coded rather than the bandwidth of auditory filters.

Greenwood and Joris (1996) proposed that the amplitude-modulation following response of primary auditory fibers is limited by the mechanical filtering of the basilar membrane for low carrier frequencies (spatial filter), whereas for high carrier frequencies, the upper limit is determined by synaptic processes between the hair cell and auditory nerve fiber (temporal filter). The change in the dominant process is a consequence of the logarithmic mapping of frequency along the basilar membrane and the fact that the tuning bandwidth of primary afferents increases as the critical frequency increases. Strickland (2000) generalized this idea to AM/QFM discrimination, proposing that peripheral limitations dominate for carrier frequencies less than 1000 Hz, whereas central limitations dominate for carrier frequencies greater than 1000 Hz. Her theory is consistent with most existing data, but some particulars are unclear due to the sparseness of data. For instance, Strickland's (2000) findings show that peripheral limitations are evident in the discrimination of AM/QFM noise only for very low spectral levels (i.e. 10 dB SPL) and frequency regions below 1000 Hz. When the spectral level is increased to 40 dB, the cutoff frequencies of TMTFs are the same for all frequency regions, suggesting that all limitations are central. On the other hand, Nelson (1994) shows a greater than four-fold increase in the cutoff point of psychometric functions for the discrimination of AM and QFM tones when the intensity is increased from 40 dB SPL to 100 dB SPL, suggesting a peripheral filtering effect at high intensities. One can argue that Nelson's data reflect uncontrolled distortion product effects, particularly given that a number of non-monotonic psychometric functions are reported, but Bernstein and Oxenham (2006) also found an intensity effect for smooth psychometric functions. It is uncertain whether the different effects of level are attributable to procedural differences (psychometric functions vs. TMTFs) or to stimulus differences (AM/QFM noise vs. AM/QFM tones). Although it is reasonable to expect that the two types of carriers will produce similar patterns of results across frequency regions, the assumption remains untested because the carrier frequency has not been varied in studies that employ AM/QFM tones. This study obtains TMTFs for AM/QFM tones for different carrier frequencies at a moderately high level and thus provides a test of generalization for Strickland's theory. It differs from previous studies by attempting to control rather than mask the effects of distortion products.

2. Experiment 1: non-monotonic threshold functions

Thresholds for modulation depth needed to discriminate AM tones and QFM tones were estimated as a function of carrier frequency and stimulus bandwidth.

2.1. Subjects

Six subjects, affiliates of the University of California Irvine between the ages of 19 and 32, participated in the three experiments. Listeners received monetary compensation for their participation with the exception of a member of the laboratory personnel who participated. All had normal pure-tone thresholds (ANSI, 1989) within the range of stimulus frequencies. They were trained for a minimum of 8 h prior to data collection in this experiment. All procedures were approved by the UCI Institutional Review Board.

2.2. Stimuli

AM and QFM stimuli are represented by

$$y(t) = \sin(2\pi f_C t + \theta) + m/2[\sin\{2\pi(f_C + f_m)t\}] + m/2[\sin\{2\pi(f_C - f_m)t\}]. \quad (1)$$

The stimulus defined by Eq. (1) is composed of a carrier, f_C, and two sidebands, $f_L = f_C - f_m$ and $f_H = f_C + f_m$, where f_m is the modulation frequency. The stimulus $y(t)$ is defined as AM when θ is 0 and QFM when θ is $\pi/2$. Modulation depth, m, of the sidebands is a dependent variable varying between 0 and 1. The value of m is adjusted with an adaptive procedure in order to estimate the threshold pertaining to the AM and QFM discrimination.

2.3. Modifications of the adaptive procedure

Adjusting modulation depth, m, with an adaptive procedure (Levitt, 1971) constituted a special case because an upper limit of $m = 1$ must be imposed in order to avoid over-modulation (i.e. $m > 1$). Strickland and Viemeister (1997) circumvented this problem by terminating a block of trials whenever the adaptive schedule dictated an $m > 1$. This technique is, however, insensitive to above-chance performance at high modulation rates. In our experiment, on trials where an incorrect response was obtained with $m = 1$, m remained at 1 until two correct responses were obtained. The modulation depth was then reduced by 2 dB and a "truncation count" was increased by one. An example in Fig. 1(a) shows that truncation counts were generally zero at narrow bandwidths and increased at higher bandwidths.

A second modification consisted of a non-conventional up-down schedule, in which m is decreased by 2 dB following two consecutive correct responses, and then decreased by 2 dB following each subsequent correct response. This schedule was reset and m increased by 2 dB after each incorrect response. This

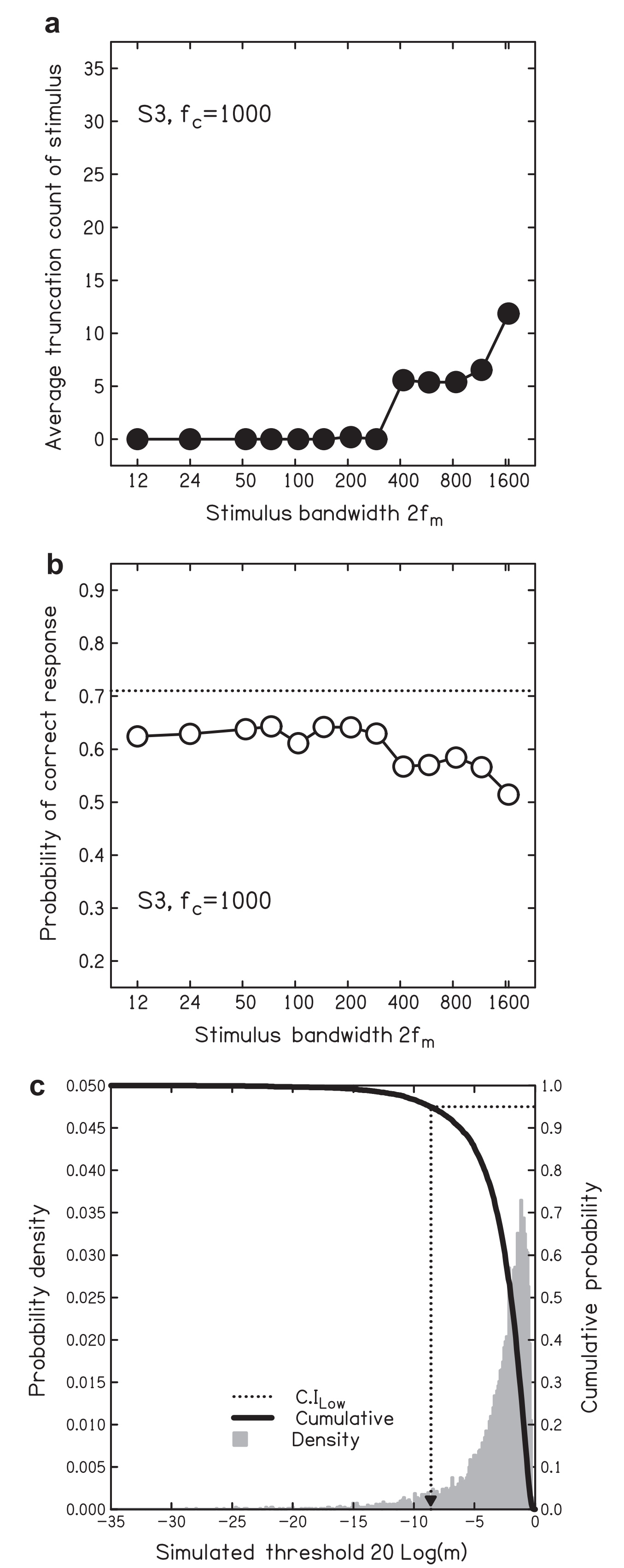

Fig. 1. Performance in the non-standard adaptive up-down procedure. (a) The average truncation count across blocks in an arbitrary condition as a function of stimulus bandwidth. (b) Probability of correct response for the same conditions as in panel (a). The dotted line at 0.71 is the probability of correct response based on the conventional up-down adaptive procedure. (c) Probability distribution of thresholds and associated cumulative probability function generated from 3000 simulations of a model that responds randomly. The arrow with the dotted line represents the one-tailed critical value, −8.62 at 5% significant level.

non-conventional procedure tracks a probability of correct response lower than 0.71, though an analytical solution to the expected probability is lacking. Empirically, the probability of a correct response was found to be approximately 0.64 when the truncation count was zero, as shown in Fig. 1(b). The intent was to lower expected thresholds in order to reduce the number of truncation counts. An evaluation of this inference was not completed because a standard two-up–one-down schedule was used in the final critical experiment.

A third modification involved the calculation of the threshold value. The occurrence of truncations when $m = 1$ violated assumptions required for averaging across reversal points in order to calculate a threshold. Instead, all stimulus values after the first four reversal points were averaged and taken as a threshold (Klein, 2001).

The chance level of performance was estimated by simulations of a model that responded randomly. The probability distribution of 3000 simulated thresholds and the corresponding cumulative function are shown in Fig. 1(c). Five percent of the thresholds are less than −8.62, and this critical value was chosen to define chance performance.

2.4. General procedure

AM and QFM were presented in random order in a 2IFC task with each of the orders equally likely. Subjects indicated which interval they believed contained the AM stimulus by pressing 1 or 2 on a keyboard, and were given accuracy feedback after every trial. Each trial consisted of two 200 ms presentation intervals with an inter-stimulus interval of 500 ms. Subjects were also shown a plot of the adaptive track after each block. Stimuli were gated by a 20-ms $\cos^2$ ramp at onset and offset. The level of the carrier was calibrated at 70 dB SPL. Sounds were generated digitally, converted to analog with an external converter (E-MU 0202 Audio/MIDI interface) at a sampling rate of 22.05 kHz, and presented diotically over headphones (Sennheiser eh350) to listeners seated in a single-wall sound attenuating chamber.

There were two independent variables; carrier frequency (500, 1000, 2000, and 4000 Hz) and modulation frequency with values between 6 and 800 Hz. Subjects typically performed 5–10 blocks of 50 trials for each condition and 4–5 conditions during each 2 h session. They completed all the conditions in quasi-random order for a given carrier frequency before moving to another carrier frequency.

2.5. Results

Thresholds are plotted as a function of stimulus bandwidth in Fig. 2, with each row representing the data for an individual and each column representing a different carrier frequency. The dotted line at −8.62 indicates chance performance obtained from simulations that emulate random responses. Thresholds are at a relative minimum in the vicinity of 100–200 Hz. Some listeners display elevated thresholds at the lowest modulation rates, an effect that is likely due to the comparatively brief stimulus duration of 200 ms (see Viemeister, 1979). A rolloff at low modulation rates precludes the fitting of a Butterworth filter to obtain estimates of S and C_f, but even for cases without a rolloff, the use of this technique is

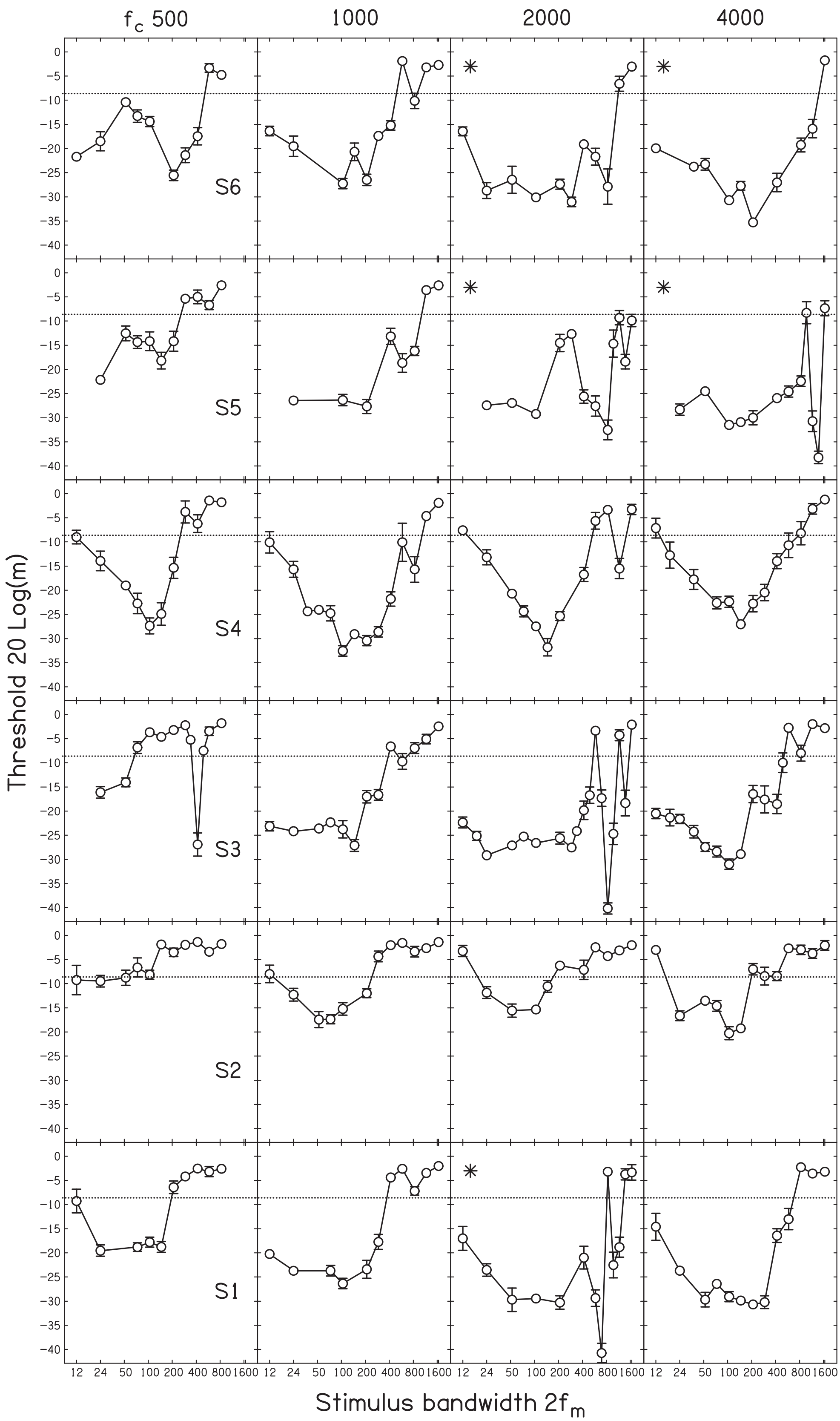

Fig. 2. Thresholds for modulation depth plotted as a function of stimulus bandwidth. Individuals and carrier frequencies are aligned in rows and columns, respectively. The dotted lines represent chance performance. The asterisk at left-top corner indicates the randomized phase and frequency conditions repeated for case studies. An error bar indicates a standard error.

compromised by the fact that not all points correspond to 71% correct. Data at the narrowest bandwidths are of secondary interest, though the range of individual differences found for this brief stimulus duration is notable. One listener (S2) shows reduced phase sensitivity for all carriers, with chance performance at all bandwidths for the 500 Hz carrier.

Non-monotonicities are found at wide stimulus bandwidths in about half of the cases. The range of individual differences is considerable. With the exception of S2, all listeners show non-monotonicities beyond a bandwidth of 400 Hz for at least one carrier frequency. The greatest effects are found for $f_C = 2000$ Hz, with S1, S3, and S5 exhibiting what can be described as a double-dip non-monotonicity. In a number of cases, thresholds are substantially lower than the plateau of the TMTF. A detailed discussion of these results follows a description of a model of distortion product effects that will provide some insights about individual differences.

This preliminary experiment exemplifies the problem of estimating phase sensitivity with an AM/QFM tone discrimination task. The threshold for S3 with $f_C = 500$ Hz and a bandwidth of 400 Hz almost certainly reflects the effects of distortion products and can be confidently disregarded as evidence of phase sensitivity at that bandwidth. In comparison, the non-monotonicities of the TMTF at high modulation rates for S5 with $f_C = 1000$ Hz and for S6 with $f_C = 2000$ Hz are less definitive. Despite some points of ambiguity, the evidence for distortion product effects at high modulation rates is persuasive. The extremely low thresholds present a clear target for the development and testing of a technique that brings distortion produce effects under experimental control.

3. Rotating quadrant model

In this section, a rotating quadrant model (RQM) is introduced that provides a plausible explanation for the non-monotonicities in TMTFs. It is assumed that the source of the effect is the interaction between the CDT and f_L. The RQM is based on three principles drawn from probe-tone, cancellation experiments: (1) the CDT interacts with f_L (Buunen, 1975), (2) the phase of the CDT is dependent on the phase of f_H (Hall, 1972b), and (3) the phase of the CDT decreases as the frequency of f_H increases and f_C remains constant (Goldstein, 1967b; Hall, 1972a; Zwicker and Fastl, 1999).

If the phase and amplitude of the CDT are represented by the vector, **CDT**, subscripted as $\mathbf{CDT_{am}}$ and $\mathbf{CDT_{qfm}}$, and the low frequency component of the stimulus is represented by the vector f_L, then the aforementioned three principles (1)–(3) can be redefined as: (I) **CDT** is added to $\boldsymbol{f_L}$, (II) the phases of $\mathbf{CDT_{am}}$ and $\mathbf{CDT_{qfm}}$ are $\pi/2$ radians out of phase, forming a right angle quadrant, and (III) the quadrant rotates as the upper sideband, f_H, increases in frequency.

The consequences of these three rules are illustrated by the representations of vector additions in Fig. 3. A fundamental assumption is that each vector sum represents the “internal level”. In state (a), $\boldsymbol{f_L} + \mathbf{CDT_{am}}$ and $\boldsymbol{f_L} + \mathbf{CDT_{qfm}}$ are identical in length, therefore representing a non-detect state. As f_H increases in frequency, $\mathbf{CDT_{am}}$ and $\mathbf{CDT_{qfm}}$ rotate synchronously as a function of f_H. State (b) illustrates that when these two vectors are rotated by $\pi/2$ radians from state (a), $\boldsymbol{f_L} + \mathbf{CDT_{qfm}}$ is greater than $\boldsymbol{f_L} + \mathbf{CDT_{am}}$, leading to a detect state. In state (c), f_H has been increased to some arbitrary frequency at which the two vectors are rotated by π radians from state (a). Since $\boldsymbol{f_L} + \mathbf{CDT_{am}}$ and $\boldsymbol{f_L} + \mathbf{CDT_{qfm}}$ are now identical, state (c) represents a non-detect state. In state (d), the quadrant has advanced by $\pi/2$ radians, and the internal intensity at f_L is now greater for AM than for QFM. One complete rotation of the quadrant will produce a threshold function with a double-dip non-monotonicity.

Without precise knowledge about the phase and intensity of CDTs for individual listeners, the RQM has little predictive power and provides only *post hoc* qualitative descriptions of threshold functions. The model is nevertheless useful in considering the range of individual differences. According to the RQM, the magnitude of a non-monotonicity depends on the amplitude of the CDT and the location of the minimum threshold depends on the phase of the CDT. Probe-tone cancellation experiments show much across-listener variability in estimates of these parameters (Buunen et al., 1974; Zwicker, 1981; Zwicker and Fastl, 1999), and so the observed range of individual differences is actually expected.

A more important use of the RQM is as a quantitative test of assumptions underlying the technique for degrading the presumed intensity cue. Simulations of the RQM show that trial-by-trial randomization of the phase of the carrier over a modest range (e.g. $-\pi/3$ to $\pi/3$) degrades intensity information, whereas the same range of phase randomization produces only a modest effect on

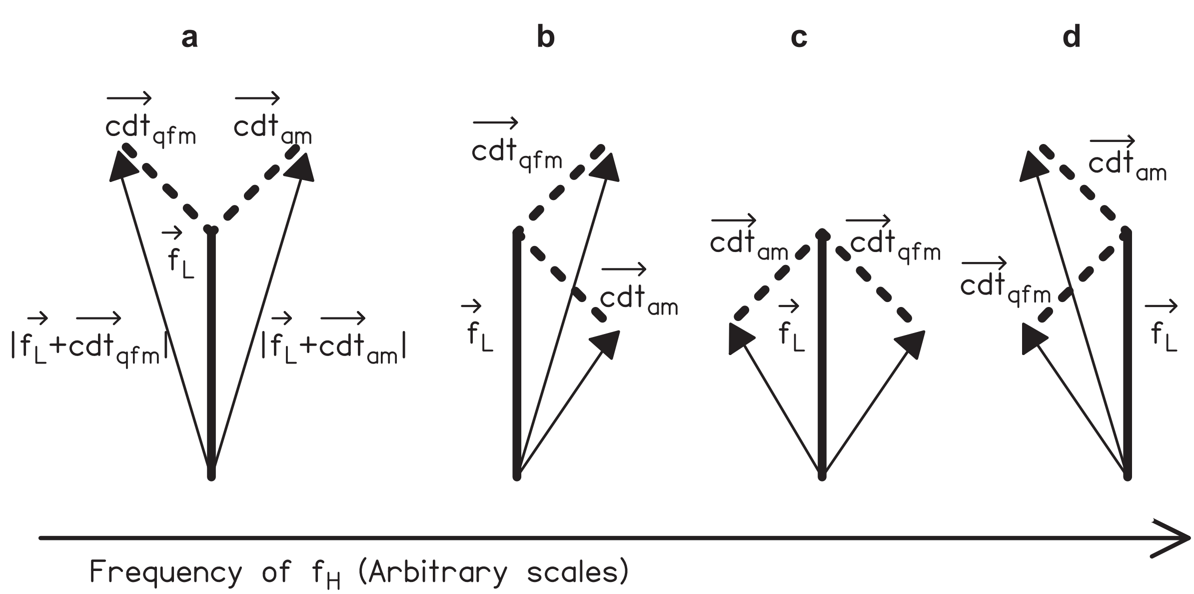

Fig. 3. Rotating quadrant model in AM/QFM discrimination. Illustrations in states (a)–(d) are aligned as a function of frequency, f_H. See text for description.

information conveyed by the envelope of the waveform, an assertion that is verified below. The level difference between AM and QFM can be expressed as $20\log(\boldsymbol{f_L} + \mathbf{CDT_{am}}/\boldsymbol{f_L} + \mathbf{CDT_{qfm}})$, where $\boldsymbol{f_L} + \mathbf{CDT} = (\boldsymbol{f_L^2} + \mathbf{CDT^2} + \mathbf{2}\boldsymbol{f_L} + \mathrm{CDT}\cos\theta)^{1/2}$ and θ is the angle between the two vectors. The maximum difference in length between $\boldsymbol{f_L} + \mathbf{CDT_{am}}$ and $\boldsymbol{f_L} + \mathbf{CDT_{qfm}}$ is found when the respective phases of the CDT vectors are $\pi/4$ and $3\pi/4$ radians (or $-3\pi/4$ and $-\pi/4$ radians). With the assumption that the intensity of the CDT is 20 dB lower than the intensity of f_H (Zwicker and Fastl, 1999), then the maximum level difference is 1.2 dB, a value slightly above a JND for a pure tone (Jesteadt et al., 1977). Fig. 4(a) shows the mean level difference across 5000 simulated trials as a function of the range of phase randomization. The mean remains above 1 dB for randomization ranges of up to $2\pi/3$ radians, but the substantial increase in the standard deviation implies the level cue becomes unreliable even for small perturbations in the carrier phase. As the range increases from $\pi/6$ radians to $\pi/2$ radians, the standard deviation of the level difference increases by more than a factor of five, representing a five-fold decrease in sensitivity under the assumptions of signal detection theory.

Temporal information is more resilient to phase randomization. Fig. 4(b) shows thresholds for two simulated decision variables as a function of the range of phase randomization. One is the max/min ratio of the Hilbert envelope (Forrest and Green, 1987), representing a depth of modulation cue. The other compares two-point power spectra of the envelope at f_m and $2f_m$, representing a modulation rate cue. For both the 'depth' and 'rate' decision variables, internal noise is added at the level of the waveform, $f(t) + e(t)$, where, $e(t)$ is a Gaussian variable. The variance of the internal noise is adjusted to yield performance in the constant-phase condition that is roughly equivalent to the thresholds of listeners and then fixed for simulations of the random phase conditions. Threshold functions for the two decision variables are essentially the same, demonstrating that the two decision variables cannot be distinguished with the current experiment. It is apparent that the information conveyed by the envelope is only slightly degraded by a moderate range of phase randomization. Simulated thresholds increase by only a few decibels when the randomization range is $\pi/2$.

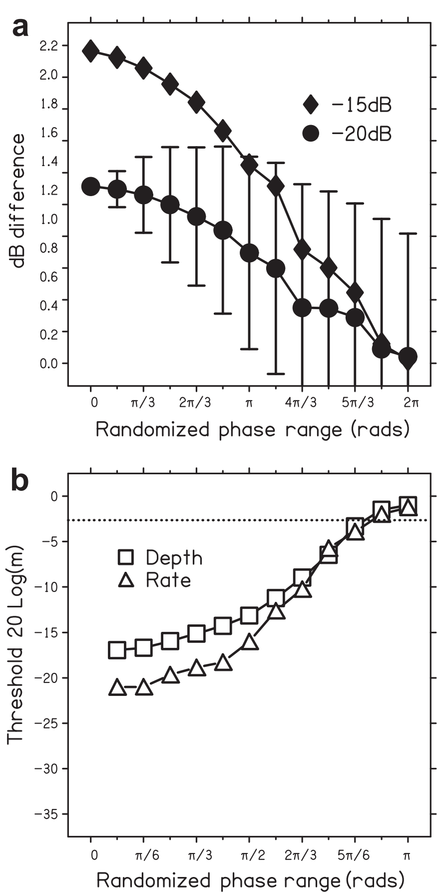

Fig. 4. The vector difference between $\mathrm{CDT_{am}}$ and $\mathrm{CDT_{qfm}}$ based on the RQM (a) and the simulated threshold from temporal decision variables (b) as a function of randomized phase range. Panel (a) plots the dB difference between the AM and QFM vectors. Circles and diamonds respectively represent cases in which the CDT is −20 dB and −15 dB below the level of f_L. Panel (b) shows simulated thresholds for the decision variable, modulation depth or modulation rate. The chance level is plotted by the dotted line at −2.65 dB.

Although the primary focus is phase randomization, the effect of randomizing the frequency of the carrier is also examined. Ostensibly, the intent is to degrade potential pitch cues (Buunen, 1975), but the level cue may also be affected by frequency randomization. If the phase of the CDT changes when the "primaries", f_H and f_C, are shifted equally in frequency, then the corresponding rotation of the RQM would degrade intensity information in the same manner as phase randomization. Although a probe-cancellation experiment dedicated to this question has not been done, data selected from Fig. 2 of Hall (1972a) which have similar values of $f_H - f_C$ suggest that the phase of the CDT shifted by $\pi/3$ radians as the pair of primaries shifted by every 100 Hz increment.

4. Experiment 2: case studies with phase and frequency randomization

Case studies are initially used to investigate the effects of randomizing the frequency and phase of the carrier and to establish appropriate ranges of randomization for degrading the level cue while having minimal effects on temporal information. Case studies are used because of the large range of distortion product effects across listeners and across carrier frequencies for individuals, but the goal is to discover a randomization range that can be generalized across listeners. A discussion of selected case studies facilitates an understanding and interpretation of the final experiment.

4.1. Stimuli

The stimuli can be described by adding two random variables, f_k and θ_k, to formula (1),

$$y(t) = \sin\{2\pi(f_C + f_k)t + (\theta + \theta_k)\} + m/2[\sin\{2\pi(f_C + f_m + f_k)t\}] + m/2[\sin\{2\pi(f_C - f_m + f_k)t\}]. \quad (2)$$

The frequency shift, f_k, was sampled from uniform distribution with a range of −10 to 10 ($\Delta f = 20$ Hz) or with a range of −40 to 40 ($\Delta f = 80$ Hz). Likewise, the phase shift, θ_k, was sampled from a uniform distribution with a range of $-\pi/12$ to $\pi/12$ ($\Delta\theta = \pi/6$) radians or with a range of $-\pi/6$ to $\pi/6$ ($\Delta\theta = \pi/3$) radians. The two random variables, f_k and θ_k, were independently sampled for AM and QFM on every trial. The resolutions of frequency and phase randomization were 1 Hz and 1°, respectively. All other aspects such as stimulus generation and procedures were identical to the first experiment. Three listeners (S1, S5, and S6) participated in the case studies.

4.2. Results

In the first experiment, the lowest thresholds are found for conditions that also display a double-dip non-monotonicity at wide bandwidths. These data are of interest because the thresholds are low enough to demonstrate that the CDT effect can be brought under experimental control and because they exemplify predictions of the RQM. Data from the first experiment are reproduced with dotted lines, whereas data from randomized conditions are represented by the solid lines in Figs. 5–7.

Without randomization, S1 in Fig. 5 displays local minima at bandwidths of 670 Hz and 940 Hz with chance performance at intermediate bandwidths. As mentioned above, this pattern is consistent with one full-rotation of the RQM (see Fig. 3). Thresholds at the minimum of the first non-monotonicity are lower than thresholds at the minimum of the second non-monotonicity. This is consistent with a common finding from cancellation experiments which show that the intensity of the CDT decreases as the frequency separation between f_C and f_H increases. The data suggest a more sophisticated RQM in which the lengths of $\mathbf{CDT_{am}}$ and $\mathbf{CDT_{qfm}}$ are a function of $f_H - f_C$.

The solid line in the upper panel of Fig. 5 shows thresholds for S1 with $\Delta f = 20$ Hz and $\Delta\theta = \pi/6$. The minimum threshold of the first non-monotonicity is unaffected by the randomization, but is more sharply "tuned" in the sense that thresholds for points adjacent to the minimum are now at chance performance. The second non-monotonicity is no longer present and thresholds for bandwidths less than 400 Hz show a maximum increase of 5 dB, with many points unaffected. The bottom panel of Fig. 5 shows the effects of increasing the phase randomization to $\Delta\theta = \pi/3$ with no frequency randomization. This extent of randomization eliminates the non-monotonicities for S1.

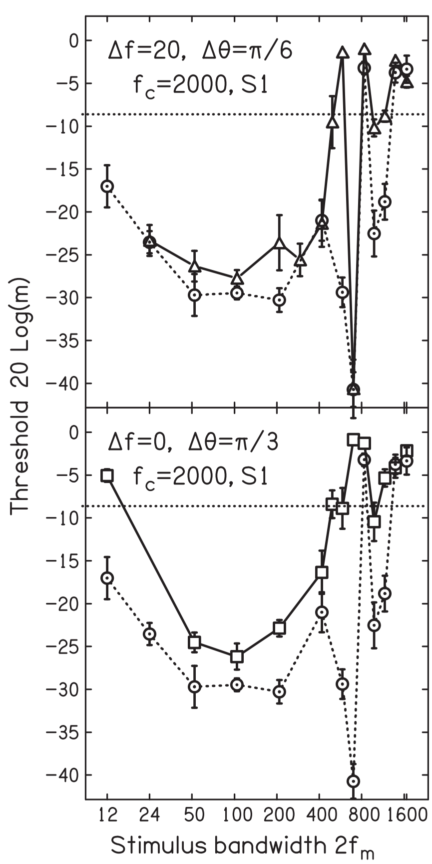

Fig. 5. Case study of the randomized conditions of Δf and $\Delta\theta$ performed by subject 1. The thresholds for randomized conditions, $\Delta f = 0$ or 20 Hz, and $\Delta\theta = \pi/6$ or $\pi/3$ radians are indicated by the solid line with a different shaped symbol. The open circles connected by the dotted line show thresholds of modulation depth for $\Delta f = 0$ and $\Delta\theta = 0$ replicated from Fig. 2. The carrier frequency and the random variables are shown at the top left corner in each panel. The dotted line consists of the critical value, −8.62 at 5% significant level.

S5 exhibits distinctive non-monotonicities, spanning a relatively broad region of bandwidths for $f_C = 2000$ Hz and exhibiting an extremely low threshold for $f_C = 4000$ Hz, as shown by the dotted lines in Fig. 6. A rare and anomalous result in which performance improves with randomization is also found for this listener. For $f_C = 2000$ Hz and a randomization of $\Delta f = 20$ Hz and $\Delta\theta = \pi/6$ (upper-left panel), the threshold at 800 Hz decreases by 5 dB in comparison to the baseline condition. The low threshold at an adjacent point, 670 Hz, establishes the consistency of this finding. There also appears to be a narrowing of the non-monotonicity which indicates that the randomization is affecting a CDT-induced cue. There are two additional cases in which performance improves with randomization (data not shown), hinting that the effect is not attributable to practice effects. A possible explanation is suggested by data from Hall's (1972b) two-tone, probe-cancellation experiment. Generally, the phase of the CDT is a linear function of the phase of f_H, a foundational assumption of the RQM, but in a few instances reported by Hall, the function relating the phase of the CDT to f_H inexplicably resembles a step function. One implication is that points near the edge of the step are unstable so that a small, random change in the phase of f_H produces a large shift in the phase of the CDT which induces a detect state. This odd effect illustrates that the effects of nonlinearities may not yet be entirely predictable, and perhaps offers a prelude to the extent of explanation that might be gained by combining the current method with a probe-tone, beat-cancellation experiment. Nonetheless, with regards to the current aim of estimating the maximum bandwidth of phase sensitivity (BW_{max}), increasing the extent of randomization to $\Delta\theta = \pi/3$ and $\Delta f = 0$ reduces the non-monotonicity to an isolated, local minimum at 800 Hz, as shown in the lower-left panel.

Results for S5 with $f_C = 4000$ Hz are shown in the right panels of Fig. 6. For bandwidths greater than 400 Hz, thresholds generally increase by about 10 dB when the randomization range is $\Delta f = 20$ and $\Delta\theta = \pi/6$, followed by another 5–10 dB increase when the randomization is increased to $\Delta f = 80$ and $\Delta\theta = \pi/3$. These findings are consistent with predictions of the RQM, demonstrating a cause-and-effect relationship between the amount of randomization and thresholds. For stimulus bandwidths less than 400 Hz, randomization has little effect.

The case study for S6 is of interest because this listener generally exhibits the best sensitivity in the first experiment. TMTFs from the first experiment, reproduced with dotted lines in Fig. 7 for $f_C = 2000$ Hz and $f_C = 4000$ Hz, have an extended low-threshold plateau followed by a relatively shallow rolloff. Results from randomization conditions suggest that the shape of the function might be the product of an overlap between the bandwidth of phase sensitivity and the region of distortion product effects. The solid line in the upper panel of Fig. 7 shows the results for $f_C = 2000$ Hz with $\Delta\theta = \pi/3$. The minimum threshold at a bandwidth of 800 Hz increases to chance performance with randomization, but the effect at 560 Hz perhaps speaks more to the point. The 6 dB increase in the threshold at this bandwidth is greater than expected if sensitivity is strictly dependent on phase sensitivity and is less than expected if sensitivity is strictly dependent on a CDT-induced intensity cue. It is plausible that both cues contribute to and that phase sensitivity remains when the intensity cue is degraded by randomization. A similar occurrence is found for $f_C = 4000$ Hz at a bandwidth of 400 Hz, shown in the lower panel. A degraded intensity cue may also underlie the threshold increase at 200 Hz, particularly because the lack of an effect at narrower bandwidths implies that phase sensitivity for this listener is largely unaffected by this range of randomization.

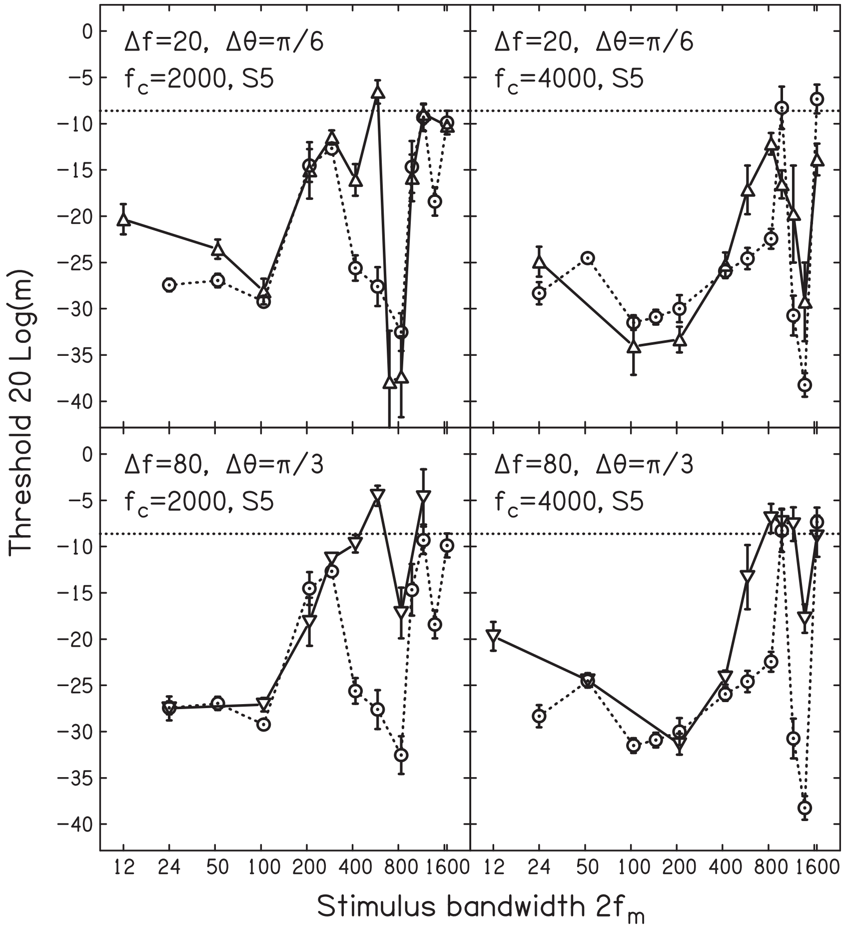

Fig. 6. Case study of the randomized conditions of Δf and $\Delta\theta$ performed by subject 5. See caption to Fig. 5.

Many of the arguments made while describing the case study results are tentative and should simply be considered as hypotheses requiring further experimentation. However, the primary premise that moderate levels of randomization affect CDT-induced level cues more than sensitivity to phase is supported by the data. For bandwidths greater than 400 Hz, it is common to see 20 dB effects with randomization, whereas at narrower bandwidths, threshold increments of more than 5 dB are rare. When non-monotonicities are not completely eliminated by randomization, they are diminished and isolated to the extent that they are easily identified and can be disregarded as evidence for phase sensitivity. Predictions for the RQM are generally supported by the data, and in cases of disagreement, the model still provides a theoretical framework for examining any discrepancies in greater detail. A definitive test of the model requires knowledge about the phase and magnitude of the CDT, information that might be obtained with a probe-tone, beat-cancellation experiment or from audiometric measurements of DPOAEs.

5. Experiment 3: discrimination of AM and QFM with a randomized phase range

In order to standardize conditions across listeners, the first experiment was repeated using phase randomization with $\Delta\theta = \pi/3$. As shown by the case studies, this range of randomization does not eliminate all non-monotonicities, but the effects are sufficiently large to identify and isolate regions of a TMTF where distortion product cues are predominant. The range of randomization does appear to balance the goals of preserving temporal information and degrading spectral information.

The adaptive schedule was modified so that m was decreased by 2 dB after two consecutive correct responses and increased by 2 dB after an incorrect response when $m < 1$. The rule when $m = 1$ was the same as the previous experiments. Simulated random responses with this schedule yielded a chance performance level of -2.65 ($\alpha = 0.05$). A threshold estimate was obtained by averaging the signal levels for all trials after the first four trials. Subjects completed 5 blocks of 50 trials for each condition. The procedure of phase randomization was the same as the one in case studies.

5.1. Results

Threshold functions are shown in Fig. 8, arranged by rows for individual listeners and by columns for carrier frequency. Although the results cannot be directly compared to the first experiment because of the change in the adaptive schedule, a number of differences are evident. The most prominent is that non-monotonicities above 400 Hz have been eliminated or greatly reduced with phase randomization. In cases where local minima remain, their locations are unchanged from the first experiment. Thresholds for bandwidths between 100 Hz and 400 Hz show a modest increase with randomization, as predicted, though a portion of this increase can be allocated to the change in the adaptive schedule. For narrow bandwidths, S4 shows the greatest

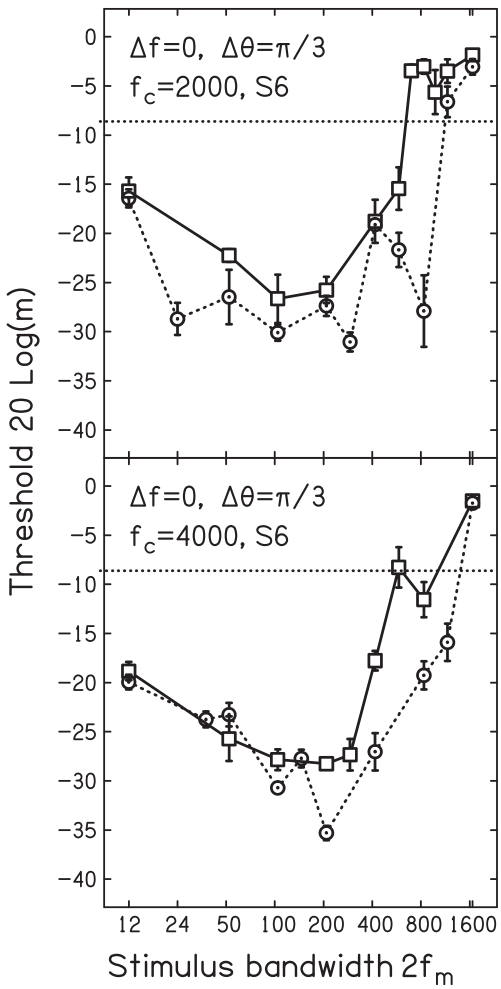

Fig. 7. Case study of the randomized conditions of Δf and $\Delta\theta$ performed by subject 6.

effects and S1 shows the most resilience to randomization. A training effect is evident for S2, who now attains better than chance performance at some bandwidths for $f_C = 500$ Hz.

Estimates of BW_{max} are represented by the single solid point in each panel of Fig. 8. Presumed distortion product effects at bandwidths greater than BW_{max} are disregarded on the basis of the narrowing of the non-monotonicity in comparison to the first experiment and greater than 10 dB increase of a local minimum and neighboring points. In most cases, the choice of the point is straightforward, but several call for an arbitrary decision. For cases where adjacent points are below but very close to the chance line, exemplified by the data for S1 with $f_C = 1000$ Hz, the narrower bandwidth is selected as BW_{max}. Although the reader may dispute some choices, any reasonable changes in the selection of the estimates have little effect on the general conclusions discussed below.

Fig. 9 summarizes data from the final experiment. BW_{max} is plotted in the left panels and f_m/f_C in the right panels. Solid points in the upper-right panel represent mean data across normal listeners from Nelson and Schroder (1995). The ratio, f_m/f_C, is the more common measure of phase sensitivity, but the range of individual differences for $f_C = 4000$ Hz is greater when the data is expressed as BW_{max}. Both measures are useful for interpreting the data. If phase sensitivity is limited by peripheral processes, then BW_{max} should increase with carrier frequency and f_m/f_C should remain constant. If central limitations predominate, then BW_{max} should remain constant and f_m/f_C should decrease with carrier frequency. BW_{max} is greater for $f_C = 1000$ Hz than for $f_C = 500$ Hz for five of the six listeners, whereas estimates of BW_{max} are relatively constant for carrier frequencies greater than 1000 Hz for all listeners except S3. Overall, the findings are consistent with Strickland's (2000) theory.

6. Discussion

With regard to the measurement of phase sensitivity, the results confirm previous findings, but presumably with more precision gained by controlling distortion product effects. For $f_C = 1000$ Hz, estimates of phase sensitivity are within the collective bounds reported in the literature (see Nelson, 1994). The pattern of results across carrier frequencies is consistent with those obtained using AM/QFM noise (Strickland and Viemeister, 1997; Strickland, 2000). The range of individual differences is also consistent with previous reports, such as Nelson's (1994) study, as discussed in the introduction. For $f_C = 1000$ Hz, Nelson and Schroder (1995) report estimates of f_m/f_C ranging from 0.1 to 0.6 for normal-hearing listeners, a study distinguished by its large sample size. The substantial range of estimates across different experiments, as summarized by Nelson (1994), might be attributable to the small number of listeners used in the experiments.

Foremost, the current results expound individual differences, an issue that has not previously been addressed in the context of AM/QFM discrimination tasks, even though a number of idiosyncratic results are found in the literature. Several listeners in Nelson's (1994) study, for instance, exhibit highly irregular psychometric functions, and Strickland and Viemeister (1997; Fig. 5) report that thresholds for two of four listeners are unaffected when the level of a notched-noise masker is increased by 60 dB. Viewed through the lens of Strickland's (2000) theory, individual differences in phase sensitivity engender some interesting theoretical issues. The range of estimates of BW_{max} across listeners in this study is about the same at all carrier frequencies, which implies that differences exist across individuals for both central and peripheral processes. What is unknown is the extent to which limitations in peripheral and central processes are correlated. Do changes in peripheral filtering, through aging or environmentally induced impairment, necessitate a corresponding change in central processing, or is central processing independent of and robust to changes in peripheral filtering? Data from this "first study" do not approach this question, but with refinements, the method of phase randomization may provide estimates of temporal sensitivity with enough precision to support fine-grained interpretations.

Increased precision might be gained in a number of ways. Pilot data show that thresholds at lower carrier frequencies are resilient to a phase randomization range as great as $\pi/2$ radians, as predicted by the simulation illustrated in Fig. 4(b). This implies that the range of randomization used to degrade distortion product cues can be increased for some listeners. A promising methodology for degrading phase information while preserving spectral information is currently being tested (Borucki and Berg, 2010), and concurrent application with the current method may yield a clearer disassociation of the relative contributions of spectral and temporal cues. With increasing precision in measurements of phase sensitivity, individual differences might ultimately prove to be highly informative instead of being viewed simply as a nuisance variable.

A unique contribution of this work is the demonstration that distortion product effects can be systematically investigated with a discrimination task. Much of our knowledge about the perceptual effects of distortion products has been obtained with probe-tone, cancellation experiments, which are cumbersome and require highly trained, knowledgeable listeners. Although much research has been done on distortion products since the first audiometric measurements of otoacoustic emissions (Kemp, 1978), little work

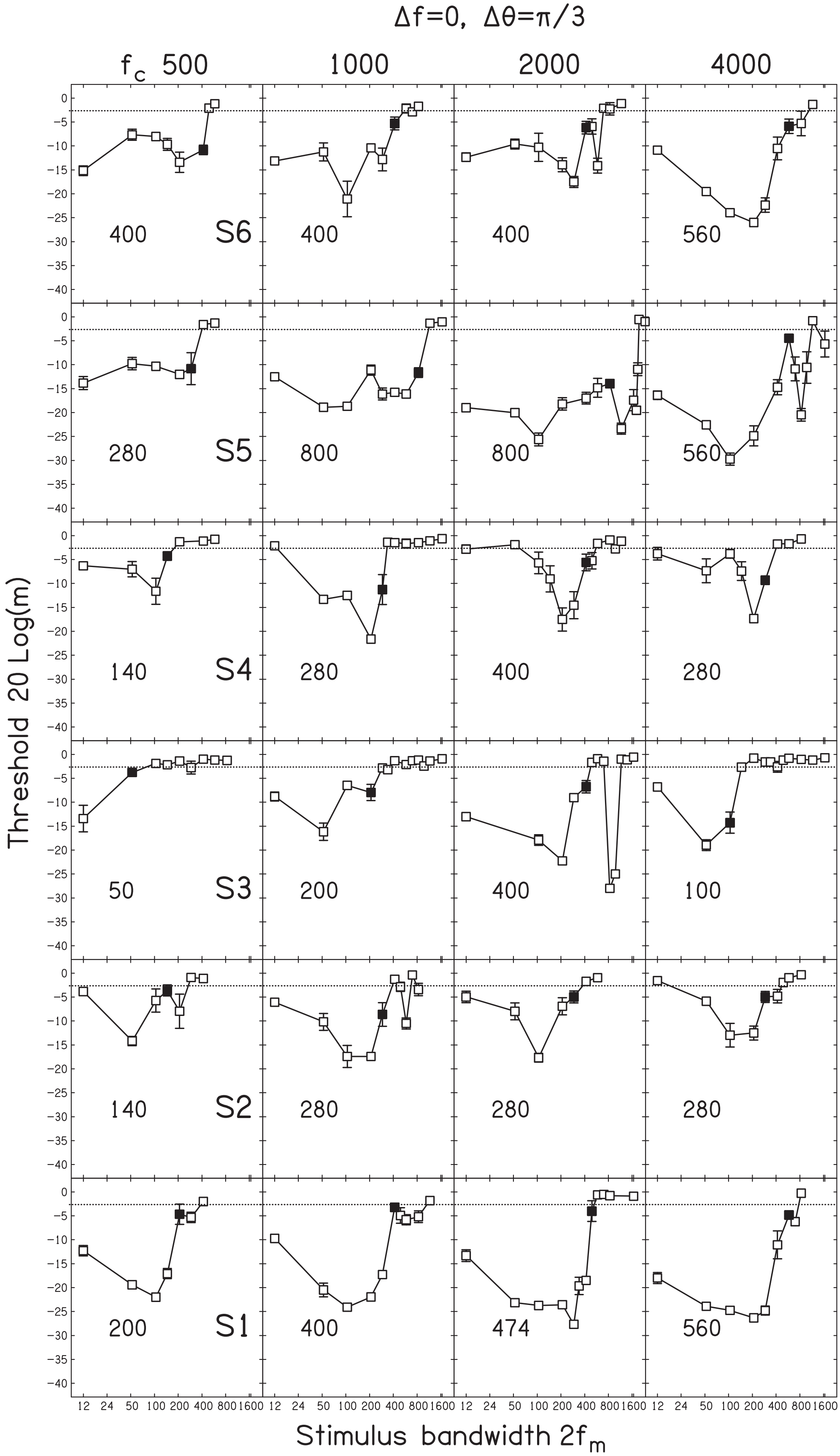

Fig. 8. Thresholds for modulation depth plotted as a function of stimulus bandwidth. Individuals and carrier frequencies are aligned in rows and columns, respectively. The open squares connected with the solid line represent the thresholds with randomization of $\Delta f = 0$ and $\Delta\theta = \pi/3$ radians. Estimated maximum critical bandwidths for phase sensitivity BW_{max}, shown as filled squares, are indicated at the bottom left corner in each panel. The dotted line at −2.65 represents chance performances.

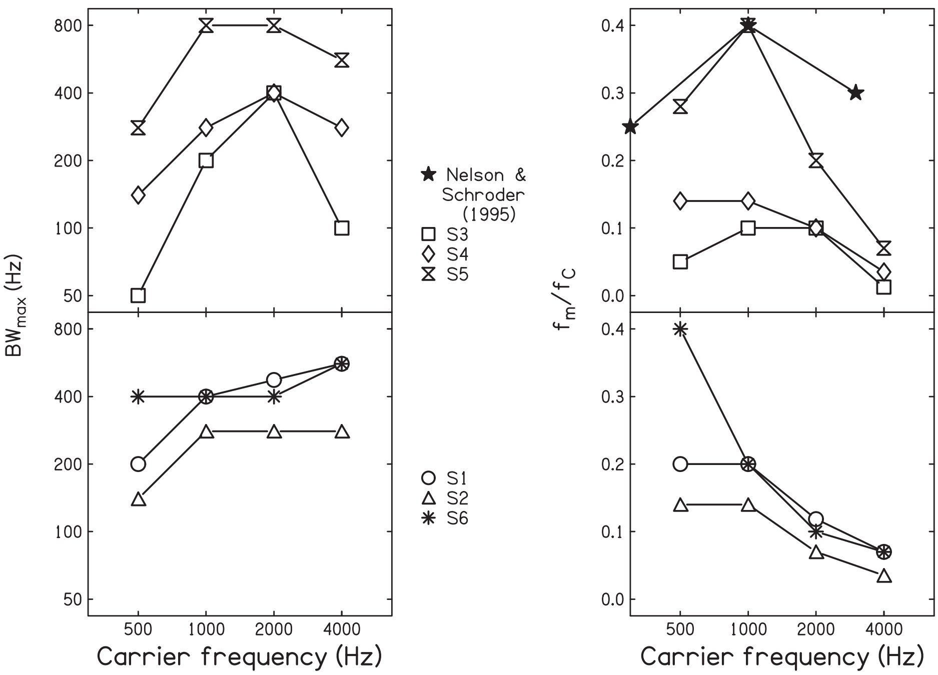

Fig. 9. Estimates of BW_{max} and f_m/f_C as a function of carrier frequency. For clarity, results for S3, S4, and S5 are shown in the top panel and results for S1, S2, and S6 are shown in the bottom panels.

has been done on the perceptual effects of distortion products in this century. The phase randomization technique holds the promise of a new and informative line of research.

Acknowledgments

The authors thank Dr. Virginia Richards, Dr. Mathew Turner, Allison Shim, and reviewers for helpful comments. This work was supported by a grant from the National Science Foundation (BCS-07464003).

References

ANSI, 1989. ANSI S3.6-1989. American National Standard Specifications for Audiometers. American National Standards Institute, New York.
Bernstein, J.G., Oxenham, A.J., 2006. The relationship between frequency selectivity and pitch discrimination: sensorineural hearing loss. J. Acoust. Soc. Am. 120, 3929–3945.
Borucki, E., Berg, B.G., 2010. Disassociating spectral and temporal influences in an AM/QFM (amplitude and quasi-frequency modulated) discrimination task. J. Acoust. Soc. Am. 127, 1988.
Buunen, T.J.F., 1975. Two hypothesis on monaural phase effects. Acustica 34, 98–105.
Buunen, T.J.F., 1976. On the perception of phase differences in acoustic signals. Ph.D. dissertation, Technische Hogeschool Deft.
Buunen, T.J.F., Festen, J.M., Bilsen, F.A., van den Brink, G., 1974. Phase effects in a three-component signal. J. Acoust. Soc. Am. 55, 297–303.
Eddins, D.A., 1999. Amplitude-modulation detection at low- and high-audio frequencies. J. Acoust. Soc. Am. 105, 829–837.
Fletcher, H., 1940. Auditory patterns. Rev. Mod. Phys. 12, 47–65.
Forrest, T.G., Green, D.M., 1987. Detection of partially filled gaps in noise and the temporal modulation transfer function. J. Acoust. Soc. Am. 82, 1933–1943.
Goldstein, J.L., 1967a. Auditory spectral filtering and monaural phase perception. J. Acoust. Soc. Am. 41, 458–479.
Goldstein, J.L., 1967b. Auditory nonlinearity. J. Acoust. Soc. Am. 41, 676–689.
Greenwood, D.D., Joris, P.X., 1996. Mechanical and "temporal" filtering as codeterminants of the response by cat primary fibers to amplitude-modulated signals. J. Acoust. Soc. Am. 99, 1029–1039.
Hall, J.L., 1972a. Auditory distortion products f_2-f_1 and $2f_1-f_2$. J. Acoust. Soc. Am. 51, 1863–1871.
Hall, J.L., 1972b. Monaural phase effect: cancellation and reinforcement of distortion products f_2-f_1 and $2f_1-f_2$. J. Acoust. Soc. Am. 51, 1872–1881.
Jesteadt, W., Wier, C.C., Green, D.M., 1977. Intensity discrimination as a function of frequency and sensation level. J. Acoust. Soc. Am. 61, 169–177.
Kemp, D.T., 1978. Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am. 64, 1386–1391.
Klein, S.A., 2001. Measuring, estimating and understanding the psychometric function: a commentary. Percept. Psychophys. 63, 1421–1455.
Levitt, H., 1971. Transformed up-down methods in psychophysics. J. Acoust. Soc. Am. 49, 467–477.
Mathes, R.C., Miller, R.L., 1947. Phase effects in monaural perception. J. Acoust. Soc. Am. 19, 780–797.
Nelson, D.A., 1994. Level-dependent critical bandwidth for phase discrimination. J. Acoust. Soc. Am. 95, 1514–1524.
Nelson, D.A., Schroder, A.C., 1995. Critical bandwidth for phase discrimination in hearing impaired listeners. J. Acoust. Soc. Am. 98, 1969–1976.
Strickland, E.A., 2000. The effects of frequency region and level on the temporal modulation transfer function. J. Acoust. Soc. Am. 107, 942–952.
Strickland, E.A., Viemeister, N.F., 1997. The effects of frequency region and bandwidth on the temporal modulation transfer function. J. Acoust. Soc. Am. 102, 1799–1810.
Viemeister, N.F., 1979. Temporal modulation transfer functions based upon modulation thresholds. J. Acoust. Soc. Am. 66, 1364–1380.
Zwicker, E., 1981. Dependence of level and phase of the $(2f_1-f_2)$-cancellation tone on frequency range, frequency difference, level of primaries, and subject. J. Acoust. Soc. Am. 70, 1277–1288.
Zwicker, E., Fastl, H., 1999. Psychoacoustics: Facts and Models. Springer, Berlin.