Level dependence in behavioral measurements of auditory-filter phase characteristics

Yi Shen[a] and Jennifer J. Lentz
Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405-7000

(Received 31 October 2008; revised 11 August 2009; accepted 13 August 2009)

Two masking experiments were conducted to behaviorally estimate auditory-filter phase curvatures at different stimulus levels. Maskers were harmonic complexes consisting of equal-amplitude tones and phase spectra with varied curvatures. In Experiment 1, sinusoidal signal thresholds were measured at 2 and 4 kHz at fixed masker levels ranging from 50 to 90 dB sound pressure level (SPL). In Experiment 2, the masker level that just masked a sinusoidal signal at 2 and 4 kHz was measured at fixed signal levels of 25, 38, and 50 dB SPL. For both experiments, the estimated phase curvature approached zero (became less negative) with increasing stimulus level. This shift could suggest that the off-frequency phase characteristic of the auditory filter has an increasingly greater role on the estimated auditory-filter phase curvature at higher stimulus levels. This explanation is supported through the use of psychophysical modeling.

[DOI: 10.1121/1.3224709]

PACS number(s): 43.66.Nm, 43.66.Dc, 43.66.Ba [BCM] Pages: 2501–2510

I. INTRODUCTION

Historically, masking experiments have been successful tools in describing the frequency selectivity of the auditory system (e.g., Fletcher, 1940) and the magnitude response of auditory filters (Patterson *et al.*, 1982), but more recently, there also has been a focus on investigating the phase response of the auditory system using behavioral methods (e.g., Lentz and Leek, 2001; Oxenham and Dau, 2001b). These experiments employ harmonic complexes as maskers rather than the noise stimuli used in many classic masking experiments. Smith *et al.* (1986) and Kohlrausch and Sander (1995) were among the first to introduce the idea of estimating the phase characteristics of the auditory system using harmonic stimuli. Both studies showed that the phase relationships of the components of a harmonic masker can greatly influence the detection threshold of a pure-tone signal, even for stimuli with identical power spectra. Two specific phase settings in the maskers, namely, the positive Schroeder phase (+Schr) and the negative Schroeder phase (−Schr) [described by Schroeder (1970)], can lead to threshold differences as large as 20 dB. The +Schr stimulus (characterized by a downward linear frequency sweep) is the time reverse of −Schr stimulus and is typically a less effective masker than −Schr stimulus. This difference in thresholds for stimuli with identical power spectra but different phase spectra is called the masker phase effect.

It is thought that the masker phase effect arises from interactions between the frequency glide of the auditory-filter impulse response and the frequency glides of the complex maskers (Kohlrausch and Sander, 1995). In particular, the impulse response of the auditory filter has a low-to-high frequency glide. The rate of the glide, which is related to the phase curvature, interacts with that of the masker. Kohlrausch and Sander (1995) argued that when the phase curvature of the cochlea counteracts that of a masker, the internal representation of the masker has all frequency components in phase. As a consequence, a highly modulated temporal envelope is formed. When a tone is added to this internally modulated masker, the resulting internal representation will contain valleys that have a higher signal-to-noise ratio than the peaks. The higher signal-to-noise ratio in the valleys provides a better chance to detect a target signal and could lead to a lower detection threshold than for a less-peaked waveform having no distinct peaks and valleys. Psychophysical masking period patterns (MPPs) and physiological experiments support this interpretation by demonstrating that the +Schr stimulus has a more peaked representation than the −Schr stimulus (Kohlrausch and Sander, 1995; Recio and Rhode, 2000; Summers *et al.*, 2003).

Closer evaluation of the relationships between masker phase and masked threshold can be made using a procedure developed by Lentz and Leek (2001) by introducing a scalar C into the original Schroeder-phase formula:

$$\theta_n = C\pi n(n-1)/N, \qquad (1)$$

where θ_n is the phase of the nth harmonic and N is the total number of components. Scalar C is proportional to the phase curvature and hence is inversely related to the rate of the frequency glide in the complex. This modification allows a systematic manipulation of the phase settings ranging between negative Schroeder phase ($C=-1$) and positive Schroeder phase ($C=1$) without altering the spectral composition of the stimulus. Using this equation, the C value at which the masking efficiency reaches a minimum ($C_{\min}$) can be measured. This $C_{\min}$ value corresponds to the phase setting in the masker that best mirrors the cochlear phase response and therefore provides an indirect measure of the auditory-filter phase curvature. $C_{\min}$ is not necessarily the positive Schroeder-phase setting, but all estimates of $C_{\min}$ in

---

[a] Author to whom correspondence should be addressed. Electronic address: shen2@indiana.edu

humans have been positive in sign (Lentz and Leek, 2001; Oxenham and Dau, 2001b, 2004).

Although the cochlear nonlinearity has been implicated in the size of the masker phase effect, little psychophysical work exists which assesses whether the behaviorally estimated phase curvature is level dependent. Most studies have evaluated level dependence in Schroeder-phase masking using only +Schr and −Schr stimuli as maskers. Carlyon and Datta (1997a) showed that for masking of long-duration tones, the difference in masking efficiency between +Schr and −Schr stimuli tends to increase with increasing stimulus level. Carlyon and Datta (1997b) also indicated that MPPs of −Schr stimuli do not typically change with level, but the degree to which the MPPs for +Schr stimuli vary with time tends to increase with increasing stimulus level and decreases at very high levels (see also, Summers, 2000).

This non-monotonic dependence on level can be understood as a trade-off between two separate effects. On the one hand, increases in level cause modulated sounds to become less effective as maskers than unmodulated sounds (Bacon *et al.*, 1997). Because the +Schr stimulus has a greater internal modulation than the −Schr stimulus, it would become a (relatively) less effective masker as stimulus level increases. Consistent with this idea is the finding that the difference in masked thresholds provided by the +Schr and −Schr stimuli tends to increase with increasing stimulus level (Carlyon and Datta, 1997a).

On the other hand, level increases cause the temporal modulation of +Schr and −Schr stimuli to become more similar. This effect has been demonstrated physiologically by Summers *et al.* (2003), who measured basilar-membrane (BM) responses to Schroeder-phase complexes in guinea pig using laser velocimetry. Their results show that while the envelopes of BM responses to −Schr stimuli do not vary greatly with level, the BM response envelopes of +Schr stimuli show a reduction in temporal modulation depth as level increases. At low levels, the +Schr BM responses are significantly peakier than those of −Schr stimuli and are similar to BM impulse responses derived using broadband noises. In contrast, at high levels, the difference in modulation depth between +Schr and −Schr stimuli is much less obvious. Summers *et al.* (2003) argued that these results could have arisen due to an auditory-filter phase curvature that is not constant along the frequency axis. This phase curvature has a large negative curvature at the characteristic frequency (CF) and a reduced curvature (nearer to zero) at lower frequencies. At low stimulus levels, then, a single location along the BM responds to a narrow frequency region around the CF, and a positive Schroeder-phase curvature could compensate the curvature of the auditory-filter phase response thereby leading to a peaked BM response. As level increases, however, the response region of this same location extends to lower frequencies where the +Schr setting does not mirror the auditory-filter phase response. As a result, the frequency components in +Schr stimuli do not arrive at the observation location synchronously, resulting in a less-peaked response envelope.

Predictions can be drawn from this "off-frequency phase influence" hypothesis about the behaviorally estimated auditory-filter phase curvature. Specifically, when measuring signal detection thresholds as a function of masker phase curvature (by systematically varying C), it is expected that $C_{\min}$ would be more positive at lower levels than at higher levels. To our knowledge, Oxenham and Dau (2001b) provided the only experiment that systematically explored whether changes in stimulus level lead to changes in the behaviorally estimated auditory-filter phase curvature. Three masker levels (40, 60, and 85 dB) were tested at each of three signal frequencies (250, 1000, and 4000 Hz). They found that $C_{\min}$ did not change significantly with stimulus level, which was in contradiction with the hypothesis proposed by Summers *et al.* (2003). The present study uses a pair of experiments to expand upon the work of Oxenham and Dau (2001b). These experiments are intended to determine whether the estimate of auditory-filter phase curvature varies with stimulus level by using a larger sample of masker levels and by testing different C values at various masker and signal levels.

II. EXPERIMENT 1: EFFECTS OF MASKER LEVEL ON ESTIMATED PHASE CURVATURE

A. Methods

1. Stimuli

Thresholds were measured for a sinusoidal signal in the presence of a simultaneous masker. Both the signal and the masker were 300 ms in duration. They were gated together with 30-ms raised-cosine onset and offset ramps. Two signal frequencies, f_s, were tested: 2000 and 4000 Hz. These signal frequencies were selected to improve chances of pinpointing minima in the functions, which are expected to be $C=1.0$ and below for the chosen masker fundamental frequency and masker bandwidth in the present experiment.[1] On every stimulus presentation, the starting phase of the signal was selected randomly from a distribution of $0-2\pi$ radians. The random starting phase of the signal was chosen so that our thresholds can be compared to those of other studies in which scalar C was varied (e.g., Oxenham and Dau, 2001b).

The masker was a harmonic tone complex with a fundamental frequency of 100 Hz and frequency components ranging between $0.4f_s$ and $1.6f_s$. The phases of the components were selected according to a modification of Schroeder's phase equation (Lentz and Leek, 2001). For each signal frequency, the masker was presented at fixed overall levels of 50-, 60-, 70-, 80-, and 90-dB sound pressure level (SPL). These levels are about 14 dB higher than the masker component level for $f_s=2$ kHz and 17 dB higher for $f_s=4$ kHz. An additional high-pass broadband noise (cutoff frequency $=1.8f_s$) was presented at a total power of 50-dB SPL simultaneous with the complex masker in order to limit off-frequency listening to frequencies beyond that of the masker. Although this masker may not have been at a high enough level to limit off-frequency listening, spot checks on two of the subjects with a high-pass masker presented at 80-dB SPL indicated no difference in threshold values for the 50- and 80-dB SPL masker levels. Thus, it is expected that listeners were not listening off frequency. For each f_s and stimulus level, signal detection thresholds were measured for C values ranging from −1.0 to +1.5. At least seven C values were tested at each masker level; these C values were chosen based on pilot data and varied at the different masker levels.

The stimuli were generated digitally and presented using a 24-bit Tucker-Davis-Technologies Real-Time processor (TDT RP2.1; sampling rate=48 828 Hz), a programmable attenuator (TDT PA5), and a headphone buffer (TDT HB6). Stimuli were then presented monaurally via Sennheiser HD250 II Linear headphones. The experiment was conducted in a double-walled sound-attenuating booth.

2. Procedure

An adaptive three-interval, three-alternative forced-choice procedure was used in conjunction with a 2-down, 1-up tracking rule to estimate the 70.7%-correct point on the psychometric function (Levitt, 1971). The masker stimulus was presented in all three intervals, with the signal stimulus being added to any one of the three intervals with equal probability. Within each trial, the three intervals were indicated by LED lights and separated by 500-ms silent pauses. The participants responded to the stimuli via a button box and were given correct-answer feedback through the LED lights. Each track consisted of eight reversals. The track began with a signal level that was equal to that of the masker. The initial step size was 5 dB, which was reduced to 2 dB after the first two reversals. Threshold was defined as the mean of the signal levels at the final six reversals.

The experiment was divided into two separate sections with all listeners being tested at 2000 Hz before being tested at 4000 Hz. For each signal frequency, the presentation order of the masker levels was randomly chosen with the constraint that the measurements at high masker levels (80- and 90-dB SPL) were not run in adjacent blocks. Once the masker level was selected, all the C values were tested in random order before the next masker level was chosen. After all masker levels were tested once for a given signal frequency, a new random order of masker levels was selected and the process repeated. Thresholds were measured at least four times at each masker level for each signal frequency. A final threshold was based on the average of these four threshold measurements. When the standard deviation across these four threshold estimates exceeded 8 dB, two more threshold estimates were included in the mean threshold. This happened for only two threshold estimates for one of the subjects (NH3). Measurements were conducted in 1.5-h sessions spanning seven to eight visits with no more than one session per day for each subject.

3. Subjects

Four subjects (two male) participated. One was the first author (NH4), and the other three were paid on an hourly basis. The subjects' ages ranged from 24 to 28 years. The pure tone thresholds of all subjects were 10-dB hearing level or better for audiometric frequencies between 250 and 6000 Hz, and the ear with better audiometric thresholds was tested. Subjects NH1 and NH3 had no previous experience in psychoacoustic experiments and received about 1 h of training before data collection started.

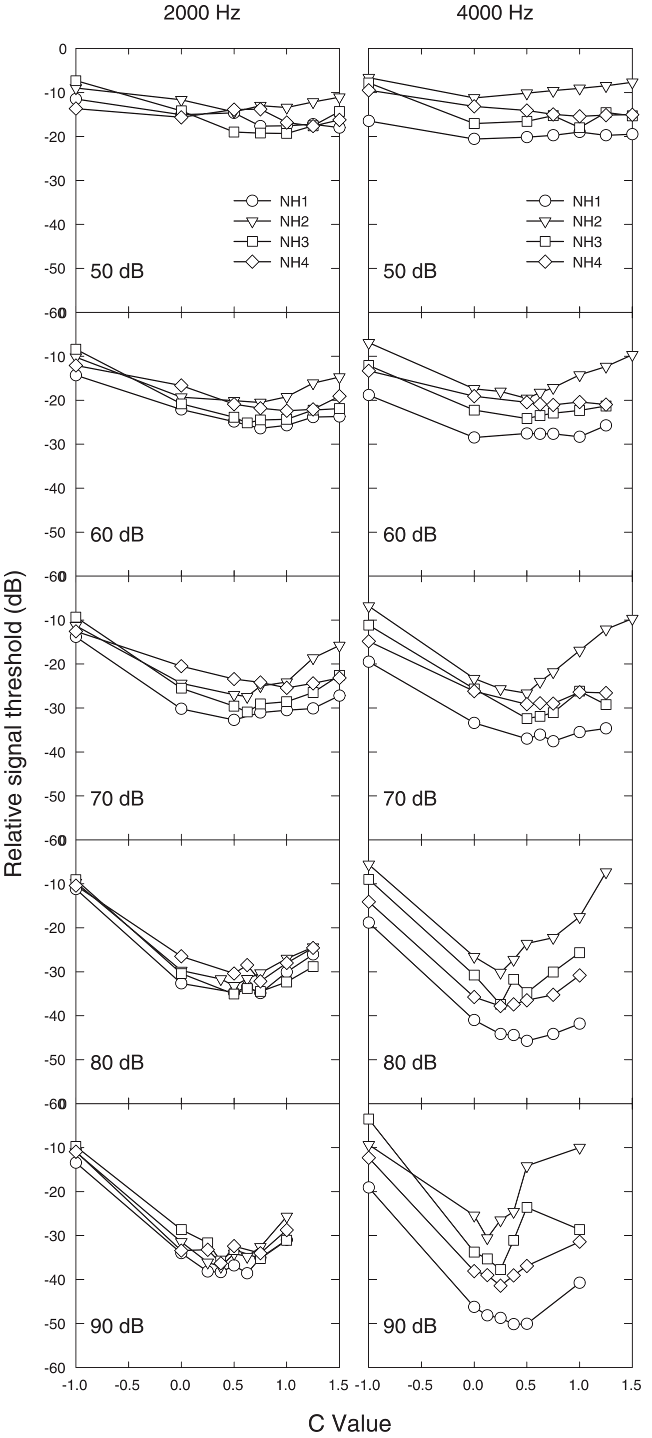

FIG. 1. Individual masked thresholds, relative to the overall masker level, are plotted as a function of scalar C, for f_s=2000 Hz (left) and 4000 Hz (right). Results for overall masker levels of 50-, 60-, 70-, 80-, and 90-dB SPL are shown in separate panels.

B. Results

The individual data for f_s=2000 and 4000 Hz are shown in the left and right panels of Fig. 1, respectively. Masked thresholds, expressed as the signal level relative to the overall masker level in decibels, are plotted as a function of the scalar C. Results for masker levels of 50-, 60-, 70-, 80-, and 90-dB SPL are shown in separate panels, and different symbols indicate data obtained from the four individual listeners.

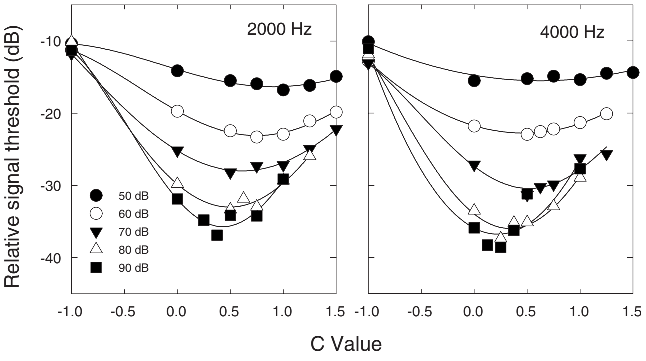

FIG. 2. Mean masked thresholds across subjects, relative to the overall masker level, are plotted for $f_s=2000$ Hz (left) and 4000 Hz (right). Results for the five overall masker levels are denoted with different symbols. Solid lines indicate data fits with a sinusoidal function $[y=y_0+a\sin((2\pi x/b)+c)]$, with y_0, a, b, and c as free parameters.

For both signal frequencies, all individuals show a similar pattern in the functions relating masked threshold to C. In general, as scalar C increases, threshold decreases and then increases forming a dip centered around a positive C value. At 2000 Hz, standard errors across the four listeners are typically below 2.5 dB. Greater variability in masked thresholds is present across listeners at 4000 Hz (standard errors increase with increasing stimulus level and reach 7-dB at 90-dB SPL). Although Fig. 1 reveals large variability across listeners, such variability is commonly reported in masking studies using Schroeder-phase maskers. For very similar stimuli, the data of Oxenham and Dau (2001b) have standard errors of 5 or 6 dB (across four listeners) at 4000 Hz (see error bars in Fig. 6 of their paper).

The mean experimental data are shown in Fig. 2, which plots thresholds obtained at different masker levels as different symbols. At each masker level, the average threshold as a function of scalar C resembles the individual data, where the threshold curves form dips around positive C values. We will call the C value where the threshold reaches its minimum $C_{\min}$ and name the threshold difference between $C=-1$ and $C=C_{\min}$ the depth of threshold. Across different stimulus levels, we see that although the thresholds at $C=-1$ ($-$Schr) are quite stable, the thresholds at $C=C_{\min}$ vary greatly with level. As a consequence, the depth of threshold increases as the masker level increases.

The value of $C_{\min}$ tends to shift toward lower C values with increases in the masker level. $C_{\min}$ was estimated using one of the methods suggested by Oxenham and Dau (2001b). The mean data in each condition were fitted with a sinusoidal function, $y=y_0+a\sin((2\pi x/b)+c)$, where y_0, a, b, and c are four free parameters. The best-fitting functions (in a least-squares sense) are plotted in Fig. 2 as the solid lines. The minima of these functions at the different masker levels are taken as estimated values of $C_{\min}$. For $f_s=2000$ Hz, the minimum tends to be near 0.8 at 60-dB SPL and shifts to about 0.4 at 90-dB SPL. For $f_s=4000$ Hz, the minimum is near 0.5 at 60-dB SPL and is between 0.2 and 0.3 at 90-dB SPL. For both signal frequencies, there is a change of a factor of 2 in the minimum between 60- and 90-dB SPL. Although there is some variability in $C_{\min}$ across listeners (note Fig. 1), all listeners show decreasing shifts in $C_{\min}$ with increasing stimulus level.

The thresholds reported in Fig. 2 obtained at 4000 Hz can be compared directly with those reported by Oxenham and Dau (2001b) in their Fig. 6, as we used the same fundamental frequency and range of frequencies for the 4000-Hz signal. Our average thresholds are approximately 15 dB higher than those in Oxenham and Dau (2001b). However, Oxenham and Dau pointed out an error in their figure, which is that the thresholds are actually plotted relative to overall masker level and not (as is stated in the caption of their Fig. 6) relative to masker component level (Oxenham and Dau, personal communication). Taking this correction into account, the thresholds of Fig. 2 are quite similar to those of Oxenham and Dau (2001b). Oxenham and Dau (2001b) reported 4000-Hz thresholds in terms of signal-to-noise ratio between -10 and -15 dB at $C=-1$ for all stimulus levels. Our thresholds for this same condition range between -10 and -13 dB. Thresholds are also similar between the two studies for C values greater than zero. For example, data of Oxenham and Dau (2001b) at 85-dB SPL are between -30 and -35 dB for $C\geq 0$, and Fig. 1 shows threshold levels between about -30 and -40 dB at 80 and 90-dB SPL.

The biggest discrepancy between our data and those of Oxenham and Dau (2001b) is that our data show decreases in $C_{\min}$ with increases in stimulus level whereas the data of Oxenham and Dau (2001b) do not. It is not obvious what causes this discrepancy between the two experiments. Some possible reasons are as follows: Oxenham and Dau (2001b) tested the 4000-Hz signal in the presence of a low-pass masker to limit the effect of low-frequency distortion at high stimulus levels, whereas the current experiment did not include a low-frequency noise masker. To assess whether the presence of the low-pass masker might have influenced the pattern of results, thresholds for the signal frequency of 2000 Hz were measured again for NH4 at two masker levels (60- and 90-dB SPL) in the presence of a low-pass masker.[2] These thresholds did not differ from the previous thresholds, and the $C_{\min}$ did not change. This result suggests the shift of $C_{\min}$ in our experiment is not a consequence of low-frequency distortion at high stimulus levels. It is also possible that the present experiment had a better chance of detecting level dependence because the C values were chosen specifically to optimize $C_{\min}$ estimates and the masker level was altered in 10-dB steps. Oxenham and Dau (2001b) only tested $C\leq 1.0$ and may not have been able to pinpoint a minimum in some of the functions. Note that in Fig. 2, some of the functions, especially those at low levels, reveal minima only because C values greater than 1.0 were tested.

C. Estimating the phase curvature of the auditory filters

In this experimental paradigm, lower masked thresholds are thought to reflect more peaked internal waveforms, providing listeners the opportunity to "listen-in-the-valleys" (Buus, 1985) more than when internal waveforms are not as peaked. Because it has been argued that these highly peaked waveforms reflect the greatest interaction between the

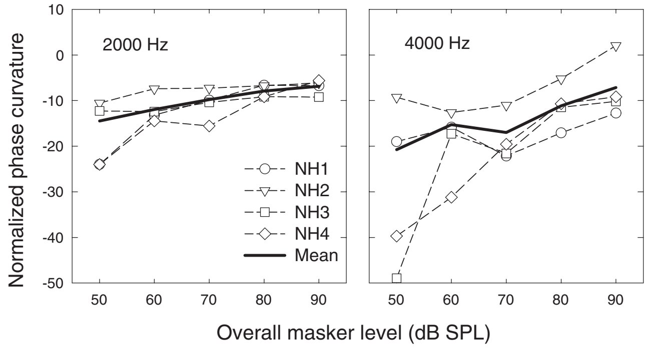

FIG. 3. The estimated phase curvature for f_s=2000 (left) and 4000 Hz (right), transformed into dimensionless units by multiplying by $f_s^2/2\pi$. The open symbols and the solid lines represent the estimated phase curvature based on curve fits to the individual data and the mean data, respectively.

auditory-filter phase curvature and the cochlear phase curvature, the C value that provides the lowest detection threshold ($C_{\min}$) might be considered an indirect measure of the auditory-filter phase curvature (Kohlrausch and Sander, 1995). Estimates of the phase curvature are based on an assumption that the phase curvature within the auditory-filter passband is constant and has the same magnitude and opposite sign to the phase curvature of the complex masker.

Using the estimates of $C_{\min}$ from the mean data, the phase curvature of the auditory filter was calculated from

$$\frac{d^2\theta}{df^2} = -C_{\min}\frac{2\pi}{Nf_0^2}, \tag{2}$$

where N is the total number of components and f_0 is the fundamental frequency of the masker. This procedure also was carried out on the individual data to assess whether individual differences in the phase curvature as a function of the masker level are present. $C_{\min}$ was estimated by fitting a sinusoidal function to the individual data in the same manner as was done for the mean data. It should be noted that at lower masker levels, the threshold varies little with the scalar C, and therefore the estimate of the phase curvature has greater error.

For ease in comparing across frequencies, absolute phase curvature values were normalized by multiplying the estimated curvature by $f_s^2/2\pi$ (Shera, 2001). The resulting quantities are dimensionless. Figure 3 shows the normalized phase curvature as a function of overall masker level for f_s =2000 and 4000 Hz.

The magnitude of the estimated auditory-filter phase curvature based on the mean data tends to decrease with increasing masker level (i.e., it approaches zero) for both signal frequencies. The estimated curvature also plateaus above 70-dB SPL at 2000 Hz for three of the four listeners, but no plateau region is observable at 4000 Hz. At low masker levels, the curvature at 4000 Hz appears to be more negative than at 2000 Hz. However, there is a great deal of variability across individual subjects, and the curvature does not differ markedly for the two signal frequencies. These curvature estimates agree well with the results from other studies where auditory-filter phase curvatures were estimated either at a fixed masker level (Oxenham and Dau, 2001b) or for a fixed signal level (Lentz and Leek, 2001). Using a 75-dB SPL masker, Oxenham and Dau (2001b) reported normalized auditory-filter phase curvature estimates of about -10 and -17 for 2000 and 4000 Hz, respectively. Lentz and Leek (2001) used a fixed-signal level of 40-dB SPL and found curvatures of about -8 and -25 at 2000 and 4000 Hz. These previous studies provide a range of curvature estimates that are consistent with our average estimates of -10 and -16 at 2000 and 4000 Hz for a 70-dB SPL masker.

To confirm that stimulus level influences the estimated auditory-filter phase curvature, a repeated-measures analysis of variance treating masker level and signal frequency as within-subject factors revealed a significant effect of level $[F(4,12)=10.99,\ p<0.005]$ but not signal frequency $[F(1,3)=6.02,\ p=0.09]$. There was no significant interaction between masker level and signal frequency $[F(4,12)=0.80,\ p=0.55]$, suggesting that the changes in phase curvature with stimulus level do not vary across these frequencies. Figure 3 also shows individual differences in the rate of change of curvature with masker level. For example, at 2000 Hz, NH1 and NH4 show larger shifts in curvature with masker level than NH2 and NH3. At 4000 Hz, NH1 shows a smaller shift of curvature than the other subjects. Despite these different rates, the estimated magnitude of the auditory-filter phase curvature for all subjects decreases with increasing masker level.

One must be cautious, however, in interpreting these results as providing support for level-dependent changes in the phase characteristic of the auditory filter. First, this observation is based on the estimation of $C_{\min}$ and an assumption that the phase curvature is constant across the auditory-filter passband. Pinpointing $C_{\min}$ can be difficult due to a clustering of low threshold values near $C_{\min}$ (see Figs. 1 and 2, especially at lower masker levels), and the assumption of constant phase curvature may not hold. Second, if the off-frequency phase response of the auditory filter significantly influences the response of the BM to the complex masker as suggested by Summers *et al.* (2003), a level-dependent shift in $C_{\min}$ is expected regardless of whether the auditory-filter curvature at the CF is varying with level. Finally, the level-dependent shifts in curvature reported here do not replicate the findings of Oxenham and Dau (2001b) who did not observe a level-dependent shift in $C_{\min}$ even though they used a very similar experimental design. Given that the source of the difference in experimental results is unknown, Experiment 2 tests whether the results of Experiment 1 generalize using a different experimental paradigm.

III. EXPERIMENT 2: EFFECTS OF SIGNAL LEVEL ON ESTIMATED PHASE CURVATURE

A. Methods

The parameters of the harmonic complex stimuli used in the present experiment were identical to those used in the first experiment. The overall masker level that just masked a signal at different signal levels was measured for C values ranging from -1.0 to 1.5. The signal levels were 25-, 38-, and 50-dB SPL for f_s=2000 Hz, and were 25- and 38-dB SPL for f_s=4000 Hz. Based on pilot listening, the 50-dB

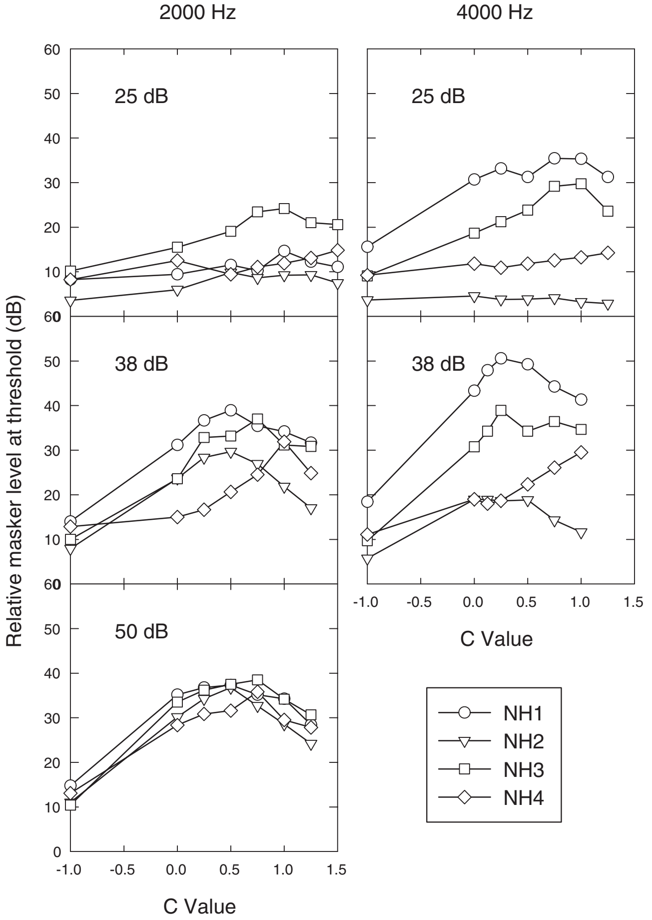

FIG. 4. Individual overall masker levels at threshold, relative to the signal level, are plotted as a function of C for $f_s=2000$ Hz (left) and 4000 Hz (right). Results from the measurements for signal levels of 25-, 38-, and 50-dB SPL are shown in separate panels.

SPL signal level was excluded from the 4000-Hz condition to prevent presenting excessively high masker levels (greater than 100-dB SPL). The lowest signal level, 25-dB SPL, was at least 10 dB above the absolute threshold for all subjects. The methods for estimating the masker level needed just to mask the signal were the same as in the first experiment, with the following exceptions. The overall masker level was increased after two consecutive correct responses and decreased after one incorrect response, and the total number of reversals for each track was 10. The initial step size was 8 dB, which was reduced to 5 dB after the first two reversals, and reduced to 2 dB after another two reversals. The estimated threshold was the mean of the masker levels at the final six reversals. A final threshold was based on the average of 4 repetitions. All four subjects from the first experiment participated and the same ears were tested.

B. Results

The masker levels that just masked the signal are plotted in Fig. 4 as a function of C. The individual data for $f_s=2000$ and 4000 Hz are shown in the left and right panels, respectively. Lower masker levels at threshold indicate more effective maskers. The trends observed in this experiment are generally similar to those found in Experiment 1. First, as the scalar C increases, the masker level usually reaches a peak value at a positive C value. However, for subject NH4, this maximum cannot be observed for the function relating masker level to C at low signal levels for both signal frequencies. Second, masker levels at threshold are typically lowest at $C=-1$, indicating that the −Schr stimulus is the most effective masker. Third, the relative masker levels at threshold are generally higher at higher signal levels, suggesting that, relatively speaking, the higher-level maskers are not as effective as the lower-level maskers.

The differences in masked threshold across individuals tend to be much larger than those reported in Experiment 1 and are again greater at 4000 Hz than at 2000 Hz. These individual differences typically occur across all C values tested and are consistent with the individual variability present in Experiment 1. For example, Fig. 4 shows that at 4000 Hz, NH1 had the highest masked thresholds, reflecting little susceptibility to masking, whereas NH2 had much lower thresholds, reflecting greater susceptibility to masking. These two listeners also experienced the most (NH2) and least (NH1) masking in Experiment 1. This across-observer variability is not likely to be due to variability within a subject because for each subject, the standard deviation of the four repetitions revealed relatively small within-subject variability. The similarity across experiments suggests that intrinsic observer-dependent sources are responsible for the large across-observer variability. Large individual differences have also been observed by Lentz and Leek (2001) who measured the levels of Schroeder-phase stimuli required to mask a 40-dB SPL tone at 2000 and 4000 Hz. At 2000 Hz, one subject had thresholds that were consistently 20–25 dB higher than the other subjects, and at 4000 Hz, a different subject had thresholds that were 10–15 dB higher than the others.

As in Experiment 1, the points of minimum masking ($C_{\min}$) corresponding to the peak masker levels were estimated in order to investigate the level dependence of the phase curvature of the auditory filter. Due to the large individual differences, this analysis was based on the individual data only. The same procedure described for Experiment 1 was performed to estimate $C_{\min}$. Briefly, results were fitted with a sinusoidal function. The C at the maximum of the function is estimated as $C_{\min}$. The conditions in which the fitted function failed to reach a maximum in the range of $-1<C<1.5$ was excluded from further analysis. The phase curvatures derived from these estimated $C_{\min}$ are shown in Fig. 5 as functions of signal level for $f_s=2000$ and 4000 Hz. A repeated-measures analysis of variance was conducted based on the relatively complete estimation data from NH1, NH2, and NH3 at 2000 Hz, treating signal level as a within-subject factor. The analysis revealed a significant effect of signal level on the estimated phase curvatures [$F(2,4)=37.16$, $p=0.003$]. This demonstrates that the psychophysically estimated phase curvature shifts toward zero as signal level increases, and replicates the results of Experiment 1 using a different paradigm.

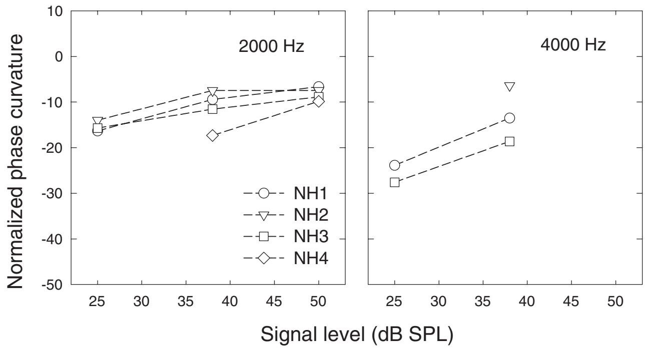

FIG. 5. The estimated phase curvature for f_s=2000 (left) and 4000 Hz (right), transformed into dimensionless units by multiplying by $f_s^2/2\pi$. The estimated phase curvature was based on curve fits to the individual data in Experiment 2.

IV. DISCUSSION

A. Relationship to physiological findings

In the present study, we investigated whether the behaviorally estimated auditory-filter phase curvature was level dependent using a Schroeder-phase masking paradigm. Two experiments revealed a decrease in the magnitude of the phase curvature of the least-efficient masker with increasing stimulus level. One possible interpretation of these results could be that the auditory-filter phase curvature is level dependent. However, this interpretation contrasts physiological measurements of BM vibration (de Boer and Nuttall, 1997; Recio *et al.*, 1998) and auditory nerve response (Carney *et al.*, 1999) in which level-invariant frequency glides have been observed in the impulse responses. Given the contradiction, we must consider the possibility that the phase curvature estimated from the psychoacoustic experiments is not the phase curvature of the auditory filter *per se*. It is likely to be influenced by the near-zero phase curvature in the frequency region away from the CF (Summers *et al.*, 2003).

The auditory-filter curvature directly relates to the instantaneous frequency (IF) trajectory of the auditory-filter impulse response. Specifically, the IF trajectory is the inverse function of the filter's group delay (e.g., Shera, 2001), and the phase curvature is defined as the (negated) slope of the group delay. Because the IF trajectory of the auditory-filter impulse response typically exhibits a low-to-high frequency glide due to the dispersion of the cochlear traveling wave, the auditory-filter curvature is almost always negative in sign. The rate of the glide in the IF trajectory varies greatly as the IF approaches the CF of the filter, causing the auditory-filter curvature to be frequency dependent. This implies that behavioral measurements of the curvature using Schroeder complexes, which assume that the auditory-filter curvature is roughly constant within the passband of the auditory filter, might be based on an invalid assumption.

If the Schroeder-phase masking experiment "pinpoints" the auditory-filter phase curvature of a specific frequency, our experimental results would directly suggest a level-dependent auditory-filter phase curvature. The discrepancy between the behavioral data and the physiological findings would be unresolvable. Further studies would be needed to identify whether the slight changes in the auditory-filter phase curvature that were undetectable physiologically could give rise to significant perceptual consequences. On the other hand, if the "off-frequency phase influence" hypothesis (Summers *et al.*, 2003) is correct, behavioral curvature estimation would be based on the integration of the auditory-filter curvature over frequency. In this case, one might measure a level-dependent psychophysical curvature estimate even though a level-invariant auditory-filter curvature is present. This will be described below.

The behavioral curvature estimation could reflect an interaction between the level-dependent magnitude response of the auditory filter and the distribution of the auditory-filter phase curvature along the frequency axis. At low levels, the shape of the auditory filter is relatively narrow, and the frequency components nearest the signal frequency provide the most masking. These components fall into the portion of the phase response that changes the most with frequency and has a negative phase curvature. In contrast at high levels, the low-frequency skirt of the auditory filter tends to broaden leading to a greater masking contribution from low-frequency components than at lower stimulus levels. The filter's phase response to these low-frequency components is linear and has a curvature near zero. If those low-frequency components have a large influence on the response at the output of the filter, the estimated phase curvature would reflect a different phase curvature than the curvature near the center of the filter (Oxenham and Ewert, 2005). In this way, the behavioral curvature estimates could shift with level even if the filter has a level-invariant auditory-filter curvature.

In terms of the impulse response of the auditory filter, the hypothesis can also be described in the time domain. As stimulus level increases, the envelope of the impulse response becomes increasingly asymmetric, with an increased emphasis on the earlier portion of the impulse response where the IF is lower and the slope of the frequency glide is steeper. As a consequence, the estimated psychophysical curvature has a magnitude that decreases with increasing stimulus level.

B. Model demonstration

To demonstrate the viability of the hypothesis that the estimation of the behavioral curvature reflects an interaction between a level-dependent magnitude response and a level-invariant phase response, model predictions of two artificial auditory-filter models (Models A and B) are compared. Model A is an auditory filter with a level-varying magnitude response but a constant curvature in the phase response (as typically assumed by psychophysical estimates of Schroeder-phase masking). Model B has the same magnitude response as Model A, but a frequency-variant phase curvature. In both filter models, the magnitude responses are identical to those of the corresponding level-dependent gammachirp filter (Irino and Patterson, 2001), but the phase responses are forced to be level invariant. Because the phase responses of the filters are artificially manipulated, these test filters are not intended to reflect cochlear processing, but are being used to demonstrate that a level-invariant auditory-filter phase response can lead to level-varying estimates of the phase cur-

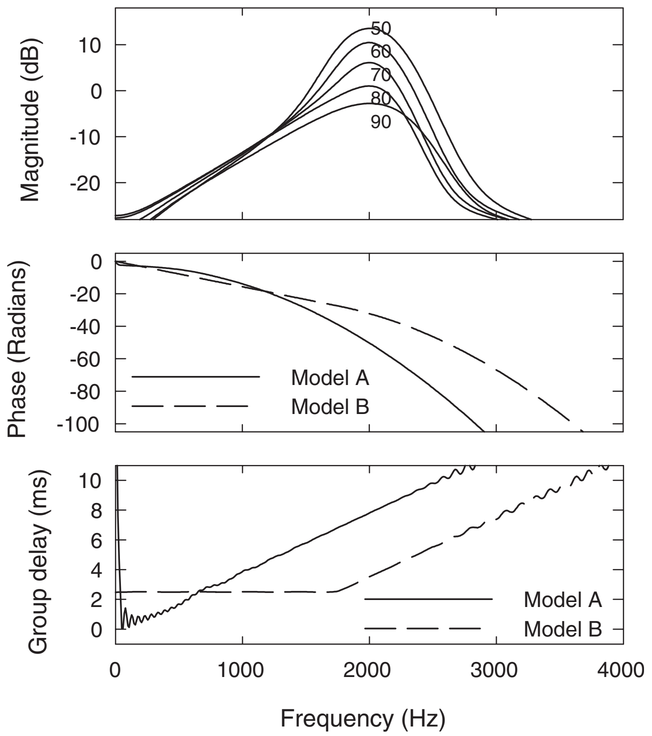

FIG. 6. Magnitude (top), phase (mid), and group delay (bottom) of two artificial auditory-filter models, Models A and B. These two filters have the same magnitude response adopted from a level-dependent gammachirp filter centered at 2000 Hz, illustrated in the top panel for the five masker levels tested in Experiment 1. A single level-invariant phase response (and therefore group delay) of the filters is used for each of the stimulus levels, with Model A having a constant curvature and Model B having a curvature of zero below 1750 Hz and a non-zero curvature elsewhere.

vature. The model simulation will demonstrate that when a constant curvature across the auditory-filter passband is assumed, no shift in the psychophysical curvature estimates would be observed even for a level-varying magnitude response. In contrast, a shift of the behaviorally measured curvature similar to what has been observed in Experiments 1 and 2 could be achieved by having a level-invariant but frequency-variant phase curvature.

Following the approach of Oxenham and Dau (2001a), gammachirp filters with the magnitude responses described by Irino and Patterson (2001) and artificial phase characteristics are presented. The magnitude responses, phase responses, and group delays of these filters are shown in Fig. 6. The two models have the same magnitude response, which was adopted from the compressive gammachirp filters at various stimulus levels.[3] The center frequency of the auditory filter was fixed at 2 kHz. Model A has a constant auditory-filter curvature equivalent to $C=-1$ (corresponding to a dimensionless curvature of -16); hence it has a linear group delay function in the frequency domain. In contrast, Model B has zero curvature at low frequencies up to 1.75 kHz and a constant curvature equivalent to $C=-1$ at all other frequencies. The group delay is therefore a piece-wise linear function (see the bottom panel of Fig. 6). Note that 1.75 kHz is approximately 1 equivalent rectangular bandwidth (ERB) below the filter center frequency (2 kHz).

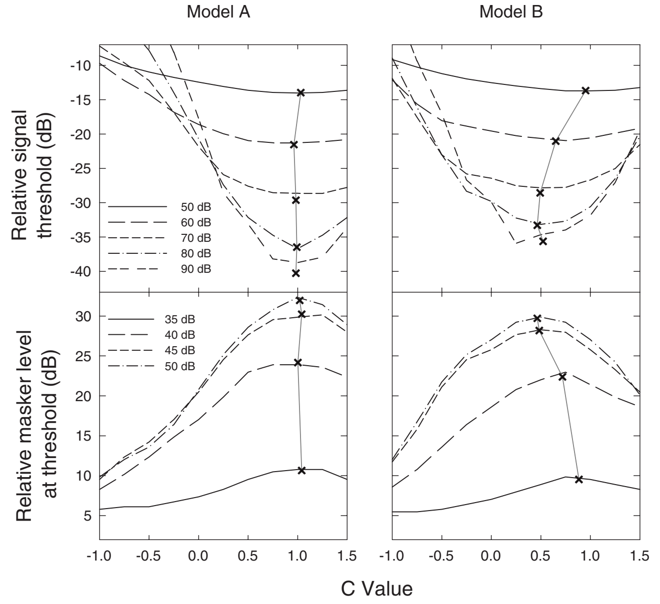

FIG. 7. The predictions from Model A (left) and Model B (right) for 2000-Hz conditions in Experiment 1 (upper panels) and Experiment 2 (lower panels). Different stimulus levels are indicated by different line styles. The predicted thresholds at each stimulus level were fit using a sinusoidal function to estimate the C value that led to the least effective masker. The minima (Experiment 1) and maxima (Experiment 2) of these fitted functions are marked with Xes and connected with gray lines.

In order to provide model predictions of the experimental data, stimuli were the maskers used for testing at 2000 Hz. Maskers alone and maskers with a 2000-Hz signal were both passed through a single auditory filter (Model A or Model B)[4] in cascade with a model of hair cell and auditory nerve fibers (Meddis, 1986) forming internal representations of the stimuli. All model parameters (for either Model A or Model B), except the magnitude response, were fixed across all experimental conditions, with the different magnitude responses used at each level. An internal power ratio between the signal-plus-masker and the masker alone was calculated from the mean-square firing rates produced by the model.[5] The signal level at threshold was determined when the power ratio exceeded a certain criterion; this criterion was fixed across all conditions. Results from Models A and B are shown in the left and right panels of Fig. 7, respectively. The simulations of Experiments 1 and 2 are shown separately in the upper and the lower panels. In the simulation of Experiment 2, signal levels of 35-, 40-, 45-, and 50-dB SPL were used instead of the ones in the actual experiment. This was due to the limited dynamic range of the Meddis (1986) model, which could not provide reliable results at low stimulus levels. The curve-fitting procedure described in Experiments 1 and 2 was performed to estimate $C_{\min}$ from the predicted thresholds. For both experiments, predicted thresholds at each stimulus level were curve fitted using a sinusoidal function. The minima (Experiment 1) and maxima (Experiment 2) of these fitted functions are indicated in Fig. 7 with Xes.

Both models illustrate the same general trends that are present in the experimental data, but each model also reveals certain limitations. First, Models A and B accurately predict a substantial change in threshold with increasing C and all curves have pronounced minima (Experiment 1) or maxima (Experiment 2). Second, the two models show that a level-

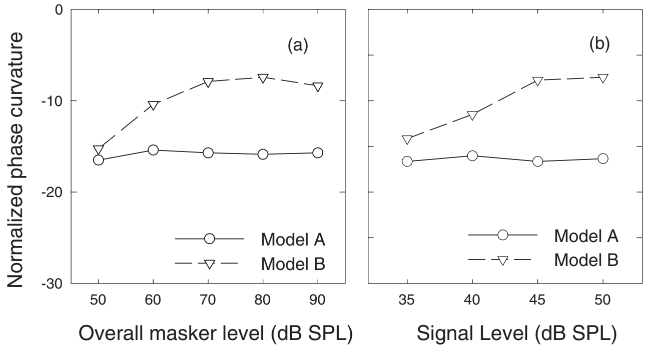

FIG. 8. The normalized magnitude of the phase curvature obtained for 2000-Hz conditions in (a) Experiment 1 and (b) Experiment 2. Results from Models A and B are shown as open circles and triangles, respectively.

dependent masker phase effect is present—as stimulus level increases the difference in masked threshold between $C=-1$ and $C_{\min}$ also increases. For Model A, $C_{\min}$ is always located at or close to $C=1.0$, the same value used to generate the curvature of the auditory-filter phase response. The clear correspondence between the two curvatures provides robust evidence that psychophysical experiments can be used to measure the auditory-filter phase curvature. However, Model A does not reveal a level-dependent shift in $C_{\min}$, as was observed in the data for both Experiments 1 and 2. In contrast, Model B leads to a level-dependent estimate of $C_{\min}$. The values of $C_{\min}$ range between 0.5 at the highest masker levels and 1.0 at the lowest masker levels. Notably, this factor of 2 change in $C_{\min}$ is quite similar to that observed in the experimental data. In this case, the model with a level-invariant and frequency-varying auditory-filter phase curvature predicts a level-dependent shift in $C_{\min}$. At the lowest stimulus level tested, the model prediction of $C_{\min}=1$ is consistent with the curvature used for the auditory-filter phase response at the CF. However, the model does not accurately predict the curvature of the phase response for the higher masker levels (e.g., 80- and 90-dB SPL). For these highest levels, the estimated phase curvature is likely due to greater weighting of low-frequency components at the output of the auditory filter. This greater weighting leads to an estimated curvature that is somewhere between zero (the phase curvature of the low-frequency tail of the auditory filter) and the curvature associated with $C=-1$ (the phase curvature in the auditory-filter passband). These modeling results demonstrate that the off-frequency phase response can influence the behaviorally estimated auditory-filter phase curvature, as suggested by Summers *et al.* (2003). Given this result, one must be cautious in interpreting psychophysical measured curvature as providing estimates of the true auditory-filter curvature, especially at higher stimulus levels.

To compare the resulting estimated curvatures of the two models more closely, the $C_{\min}$ values, converted to normalized curvature values, are plotted in Fig. 8. Figure 8 illustrates that Model A predicts a curvature that is more negative than Model B for all stimulus levels, and this predicted curvature is in good agreement with the auditory-filter curvature used for Model A (-16). The curvature magnitudes predicted by Model B also decrease with increasing level. At the lowest stimulus levels, the psychophysical curvature estimates are close to the value used in the auditory-filter model. When comparing the model curvature predictions with those of Experiments 1 and 2, it is apparent that Model B provides a better approximation to the experimental data of Experiment 1 (Fig. 3) and Experiment 2 (Fig. 5). Specifically, Model B leads to a fairly constant prediction of curvature for stimulus levels above 70-dB SPL (e.g. Experiment 1), which is comparable to the plateau of the estimated phase curvature observed in the data at 2000 Hz. The simplified phase curvature used here is likely to be similar but not necessarily identical to the phase curvature of the auditory filter. Further work characterizing the phase characteristics of the auditory-filter phase curvature may be necessary to yield better predictions of the data. Regardless, this model demonstrates that level-dependent psychophysical phase curvature estimates might reflect a level-independent auditory-filter phase curvature combined with level-dependent changes in the magnitude response with level.

These modeling results demonstrate that 1) if the auditory-filter curvature is level- and frequency-invariant (as in Model A), no shift in psychophysical curvature estimates would be observed with increasing stimulus level. 2) With the same modeling framework but a frequency-dependent auditory-filter curvature (as in Model B), shifts in the psychophysical curvature estimates similar to the ones observed in experimental data are apparent despite a level-invariant auditory-filter phase curvature. Simulations from Model B illustrate that the auditory-filter curvature at frequencies below 1 ERB from the auditory-filter center frequency could contribute to the psychophysical curvature estimates at high levels. Additional modeling shows that when the critical frequency dividing the two segments of the auditory-filter curvature (zero-curvature and constant-curvature segments) in Model B was set to 1.5 kHz (about 2 ERBs below the center frequency), the shifts in the psychophysical curvature with the increasing level could still be measured. It seems that the psychophysical curvature consists of integration of the auditory-filter curvature in these models.

V. SUMMARY

The present study provides estimates of the auditory-filter phase curvature at different stimulus levels using modified Schroeder-phase maskers. Two experiments, one using a fixed masker level and another using a fixed signal level, indicate that the estimated magnitude of the phase curvature decreases as stimulus level increases. This result demonstrates the existence of level dependence in the behaviorally measured phase curvature of the auditory filter, even though such results are not consistent with physiological measurements of the phase response at a single cochlear location. In an effort to resolve the contradictions between psychophysical and physiological results, a plausible mechanism underlying the psychophysical curvature is suggested following the "off-frequency phase influence" hypothesis by Summers *et al.* (2003), in which psychophysically estimated curvature reflects an interaction between the magnitude and the phase response of the auditory filter. Psychoacoustic modeling displays evidence of the viability of this mechanism. To develop more accurate future behavioral techniques for estimating the auditory-filter phase curvature, the role of the magnitude response has to be carefully considered.

ACKNOWLEDGMENTS

We would like to thank Dr. Armin Kohlrausch and Dr. Andrew Oxenham for their constructive reviews and helpful suggestions. Dr. Toshio Irino provided the computer program for the implementation of the gammachirp filter. Data collection was assisted by Melissa Papesh and Susie Valentine, and Dr. Robert Withnell provided many fruitful discussions that have contributed to the development of this work.

[1] Rather than testing the same signal frequencies as Oxenham and Dau (2001b), 2-kHz signals were selected instead of 1-kHz or 250-Hz tones because of the following considerations: (1) Auditory-filter bandwidths are wider for higher frequencies, allowing a greater number of components to interact within a single auditory filter. These interactions will produce greater changes in threshold at low levels and will allow more accurate estimates of auditory-filter phase curvature. (2) At frequencies below 1 kHz, a significant portion of the masker power could be presented to cochlear locations where the logarithmic frequency mapping does not hold. This might confound the interpretation of the results, and such a confound will not be as pronounced for frequencies above 1 kHz. (3) The threshold minima for signal frequencies below 1 kHz are likely to exceed $C=1$ for most listeners for the chosen masker fundamental frequency and masker bandwidth in the experiment, therefore potentially making the estimation of the minimum difficult.

[2] The low-pass noise used in these spot checks had a cut-off frequency of 600 Hz and a spectrum level of 41-dB SPL for the 90-dB conditions and 11-dB SPL for the 60-dB conditions. These levels were chosen to be about 15 dB lower than the average spectrum level of the complex masker, following Oxenham and Dau (2001b).

[3] Model parameters were identical to those in the original paper (Irino and Patterson, 2001): $n=4$, $b_1=2.02$, $b_2=1.14$, $c_1=-3.70$, $c_2=0.979$, and $f_{\mathrm{rat}}=0.573+0.0101P_s$. The compressive gammachirp filter depends on the stimulus level through the parameter P_s, which is the probe tone level in the notched-noise experiment used for model fitting (Irino and Patterson, 2001). In the present study, P_s was set to be 20 dB below the masker level. This is a rough approximation equating the sound energy within the passband of the filter between our experiments and the notched noise paradigm.

[4] It is worth pointing out that Models A and B are functionally linear. Because the filter's magnitude and phase response are separately specified, the filters could be conveniently realized via linear FIR filters. The possibility of using linear filters to predict cochlear responses to Schroeder-phase stimuli has been studied by Summers *et al.* (2003) in detail. They measured a series of "indirect" impulse responses (de Boer and Nuttall, 1997) of the cochlea at various levels. Responses were predicted by convolving the acoustic Schroeder-phase stimuli with the indirect impulse responses. These modeled responses were compared with the experimentally measured responses. Results showed that realistic predictions could be achieved by this series of linear models, which gives justification in applying linear-filter models here.

[5] The durations of the stimuli used here were 300 ms. The first 100 ms and the last 50 ms of the model output was excluded from the calculation of the mean-square ratio of the firing rates to minimize the effect of the onset and offset transients.

Bacon, S. P., Lee, J., Peterson, D. N., and Rainey, D. (**1997**). "Masking by modulated and unmodulated noise: Effects of bandwidth, modulation rate, signal frequency, and masker level," J. Acoust. Soc. Am. **101**, 1600–1610.

Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Carlyon, R. P., and Datta, A. J. (**1997a**). "Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system," J. Acoust. Soc. Am. **101**, 3636–3647.

Carlyon, R. P., and Datta, A. J. (**1997b**). "Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components," J. Acoust. Soc. Am. **101**, 3648–3657.

Carney, L. H., McDuffy, M. J., and Shekhter, I. (**1999**). "Frequency glides in the impulse responses of auditory-nerve fibers," J. Acoust. Soc. Am. **105**, 2384–2391.

de Boer, E., and Nuttall, A. L. (**1997**). "“The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions," J. Acoust. Soc. Am. **101**, 3583–3592.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–65.

Irino, T., and Patterson, R. D. (**2001**). "A compressive gammachirp auditory filter for both physiological and psychophysical data," J. Acoust. Soc. Am. **109**, 2008–2022.

Kohlrausch, A., and Sander, A. (**1995**). "Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets," J. Acoust. Soc. Am. **97**, 1817–1829.

Lentz, J. J., and Leek, M. R. (**2001**). "Psychophysical estimates of cochlear phase response: Masking by harmonic complexes," J. Assoc. Res. Otolaryngol. **2**, 408–422.

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

Meddis, R. (**1986**). "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711.

Oxenham, A. J., and Dau, T. (**2001a**). "Reconciling frequency selectivity and phase effects in masking," J. Acoust. Soc. Am. **110**, 1525–1538.

Oxenham, A. J., and Dau, T. (**2001b**). "Towards a measure of auditory-filter phase response," J. Acoust. Soc. Am. **110**, 3169–3178.

Oxenham, A. J., and Dau, T. (**2004**). "Masker phase effects in normal-hearing and hearing-impaired listeners: Evidence for peripheral compression at low signal frequencies," J. Acoust. Soc. Am. **116**, 2248–2257.

Oxenham, A. J., and Ewert, S. D. (**2005**). "Estimates of auditory filter phase response at and below characteristic frequency," J. Acoust. Soc. Am. **117**, 1713–1716.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788–1803.

Recio, A., and Rhode, W. S. (**2000**). "Basilar membrane responses to broadband stimuli," J. Acoust. Soc. Am. **108**, 2281–2298.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). "Basilar-membrane responses to clicks at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **103**, 1972–1989.

Schroeder, M. (**1970**). "Synthesis of low peak factor signals and binary sequences with low autocorrelation," IEEE Trans. Inf. Theory **16**, 85–89.

Shera, C. A. (**2001**). "Frequency glides in click responses of the basilar membrane and auditory nerve: Their scaling behavior and origin in traveling-wave dispersion," J. Acoust. Soc. Am. **109**, 2023–2034.

Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (**1986**). "Phase effects in masking related to dispersion in the inner ear," J. Acoust. Soc. Am. **80**, 1631–1637.

Summers, V. (**2000**). "Effects of hearing impairment and presentation level on masking period patterns for Schroeder-phase harmonic complexes," J. Acoust. Soc. Am. **108**, 2307–2317.

Summers, V., de Boer, E., and Nuttall, A. L. (**2003**). "Basilar-membrane responses to multicomponent (Schroeder-phase) signals: Understanding intensity effects," J. Acoust. Soc. Am. **114**, 294–306.