Estimates of auditory filter phase response at and below characteristic frequency (L)

Andrew J. Oxenham[a)]
Research Laboratory of Electronics, and Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Stephan D. Ewert[b)]
AG Medizinische Physik, Carl-von-Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

(Received 2 August 2004; revised 2 January 2005; accepted 7 January 2005)

Animal studies in basal cochlear regions have shown that basilar-membrane phase curvature (or rate of change of group delay with frequency) is negative around characteristic frequency (CF), but near zero well below CF. This study examined whether psychophysical masking experiments in humans show the same difference between on- and off-CF phase curvature. Masked thresholds were measured for a 2-kHz signal in the presence of harmonic tone complex maskers with a fundamental frequency of 100 Hz, band-limited between 200 and 1400 Hz (off-frequency masker) or between 1400 and 2600 Hz (on-frequency masker). The results from four normal-hearing listeners are consistent with predictions from animal physiological data: negative phase curvature is found for the on-frequency masker, whereas the phase curvature for the off-frequency masker is near zero. The method and results provide a strong test for the temporal response of computational models of human cochlear filtering. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.1863012]

PACS numbers: 43.66.Ba, 43.66.Dc [JHG] Pages: 1713–1716

I. INTRODUCTION

The amount of masking produced by a harmonic tone complex depends strongly on the phase relationships between the individual masker components (Smith *et al.*, 1986; Kohlrausch and Sander, 1995). The most commonly used stimuli in such experiments have been Schroeder-phase complexes (Schroeder, 1970), which have a group delay that changes linearly with frequency, producing a constant phase curvature and a time waveform that can be described as a repeating linear frequency glide. A modification of the original Schroeder equation, provided by Lentz and Leek (2001), which describes the phase of each component, is given by

$$\theta_n = C\pi n(n-1)/N, \tag{1}$$

where n is the component number, N is the total number of components, and C is a multiplicative constant. Positive (m_+) and negative (m_-) Schroeder-phase stimuli correspond to $C=1$ and $C=-1$, respectively, and a sine-phase (or cosine-phase) complex can be generated using $C=0$. Positive and negative Schroeder-phase complexes have very similar time waveforms,[1] but can produce differences in masked threshold that exceed 20 dB. The current explanation is that the positive phase curvature of the m_+ waveform interacts with the negative curvature of the basilar-membrane filter to produce a waveform with near-zero phase curvature, such as a sine-phase complex. The resulting highly modulated temporal envelope allows for signal detection at points in time with low signal-to-masker ratios. In contrast, the negative phase curvature of the m_- waveform interacts with the basilar membrane to produce a waveform that still has a relatively flat temporal envelope, resulting in higher signal-to-masker ratios at threshold [Kohlrausch and Sander (1995); see also Oxenham and Dau (2001a), for the possible role of peripheral compression in determining thresholds in simultaneous masking]. The phase curvature of the complex is given by

$$\frac{d^2\theta}{df^2} = C\frac{2\pi}{Nf_0^2}, \tag{2}$$

where N is the total number of components in the complex, and f_0 is the fundamental frequency (Kohlrausch and Sander, 1995).

Using these complexes, it has been possible to derive estimates of the phase curvature of cochlear filtering at characteristic frequencies (CFs) between 125 and 8000 Hz (Oxenham and Dau, 2001b). The reasoning is that the masker phase curvature producing the lowest signal threshold is the one that produces the most highly modulated temporal envelope after cochlear filtering (Kohlrausch and Sander, 1995), which tends to occur when the filtered waveform has a phase curvature of zero. Thus, minimum threshold will be reached when the phase curvature of the masker is equal and opposite to the phase curvature of the cochlear filter at the CF corresponding to the signal frequency (Oxenham and Dau, 2001b). Physiological studies of basilar-membrane responses to Schroeder-phase stimuli have confirmed the basic pattern of expected effects, such as more modulated responses for stimuli with positive phase curvature (Recio and Rhode, 2000; Summers *et al.*, 2003).

Studies so far have used harmonic complexes with constant phase curvature (Smith *et al.*, 1986; Kohlrausch and

[a)]Electronic mail: oxenham@mit.edu
[b)]Present address: Center for Applied Hearing Research, Acoustic Technology, Ørsted·DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark.

Sander, 1995; Carlyon and Datta, 1997; Summers and Leek, 1998; Oxenham and Dau, 2001a, b). By deriving the phase curvature of cochlear filtering from such maskers it is assumed that cochlear filtering can also be approximated as having constant phase curvature. As discussed by Oxenham and Dau (2001b), this assumption is supported by physiological data for frequencies around CF (Shera, 2001). However, at frequencies well below CF, phase curvature tends to zero. In fact the physiological data from animals using CFs of 1 kHz and above suggest that the phase response can be roughly divided into two regions, with negative curvature around CF and zero curvature for frequencies half an octave or more below CF (Shera, 2001). The situation appears to be different in apical regions of the cochlea (at CFs lower than about 1000 Hz), where the phase curvature at CF tends to zero (Oxenham and Dau, 2001b), or possibly even becomes positive (Carney *et al.*, 1999).

The change in phase curvature with frequency observed at high CFs provides an opportunity to further test the hypothesis that changes in masked threshold as a function of masker phase curvature reflect the phase response of cochlear filtering. Specifically, a masker with components around the signal frequency should produce minimum masking when the masker phase curvature is positive and "cancels out" the negative phase curvature of the on-frequency cochlear phase response. In contrast, a masker with components only well below the signal frequency should produce minimum masking when the masker phase curvature is zero, reflecting the fact that the off-frequency cochlear phase curvature is thought to be near zero. This prediction forms the basis of the experiment reported here.

II. METHODS

A. Stimuli

Masked thresholds of a 2-kHz sinusoid were measured in the presence of a harmonic tone complex masker with a fundamental frequency (F0) of 100 Hz. The masker bandwidth extended from 1400 to 2600 Hz for the on-frequency condition, and from 200 to 1400 Hz for the off-frequency condition. Thus, the masker had the same bandwidth (1200 Hz) and the same total number of sinusoidal components (13) in both conditions. In the on-frequency condition, all masker components had equal amplitudes and the overall (rms) masker level was 73 dB SPL. In the off-frequency condition, the masker spectrum was shaped with a -6 dB/oct slope, such that each component was attenuated relative to the adjacent lower component, to reduce the influence of possible edge effects produced by the highest-frequency masker component (Kohlrausch and Houtsma, 1992). The overall (rms) off-frequency masker level of 88 dB SPL was 15 dB higher than in the on-frequency condition, so as to produce similar signal thresholds in both conditions. The signal had a total duration of 260 ms, including 30-ms raised-cosine onset and offset ramps, and was temporally centered within the masker, which had a total duration of 320 ms, including 10-ms raised-cosine onset and offset ramps. The starting phase of the signal was randomized from trial to trial. The starting phases of the masker components were selected according to the modification of Schroeder's (1970) equation, as shown in Eq. (1) (Lentz and Leek, 2001), with C values ranging from -1 to 1 in steps of 0.25.

The stimuli (masker and signal) were presented to the left ear of each subject. A contralateral Gaussian noise, bandpass filtered between 1400 and 2800 Hz, was gated synchronously with each masker and presented to the right ear at a spectrum level of 20 dB SPL/Hz. This was to prevent the detection of the signal by the right ear via acoustic or electric crosstalk. In the on-frequency conditions, a pink noise, bandpass filtered between 25 and 1000 Hz, was presented to the left (signal) ear at a level of 44 dB SPL per $\frac{1}{3}$ octave band. This level was chosen to be 25 dB below the level of the masker within the $\frac{1}{3}$ octave around the signal frequency. The purpose of the pink noise was to mask possible distortion products generated by the signal and masker.

Stimuli were generated digitally and played out via a RME DIGI96/8 PAD soundcard and an external SEKD ADSP 2496 PRO digital-to-analog converter at 24-bit resolution and a sampling rate of 32 kHz. The stimuli were then passed through a Behringer HA4600 headphone buffer and presented over Sennheiser HD 580 headphones to listeners seated in a double-walled sound-attenuating booth.

B. Procedure

An adaptive three-interval three-alternative forced-choice procedure was used in conjunction with a two-down one-up tracking rule to estimate the 70.7% correct point on the psychometric function (Levitt, 1971). The masker was presented on each interval and the signal was presented randomly in one of the three intervals with equal probability. Each interval in a trial was separated by an interstimulus interval (ISI) of 500 ms. The intervals were marked on a computer monitor and feedback was provided after each trial. Listeners responded via the computer keyboard or mouse. The initial step size was 8 dB, which was reduced to 4 dB after the first two reversals and then to 2 dB after the next two reversals. Threshold was defined as the mean of the remaining six reversals. At least four threshold estimates were made for each condition. In the very rare case that standard deviation across the four runs was greater than 4 dB, another two estimates were made and the mean of all six was recorded. The conditions were run using a randomized blocked design, with all conditions being presented once before embarking on a repetition of the conditions. The order of presentation of the conditions was selected randomly for each listener and each repetition. Measurements were made in 2-h sessions, including many short breaks. No more than one session per listener was completed in any one day.

C. Listeners

Four listeners (three male, one female—ML) participated. Two were the authors; the other two were students at the University of Oldenburg. The ages of the subjects ranged from 23 to 32 years. All had audiometric thresholds of 15 dB HL or lower at octave frequencies between 250 and 8000 Hz. Subjects AO, SE, and KB had considerable experience in psychoacoustic detection tasks and were familiarized with

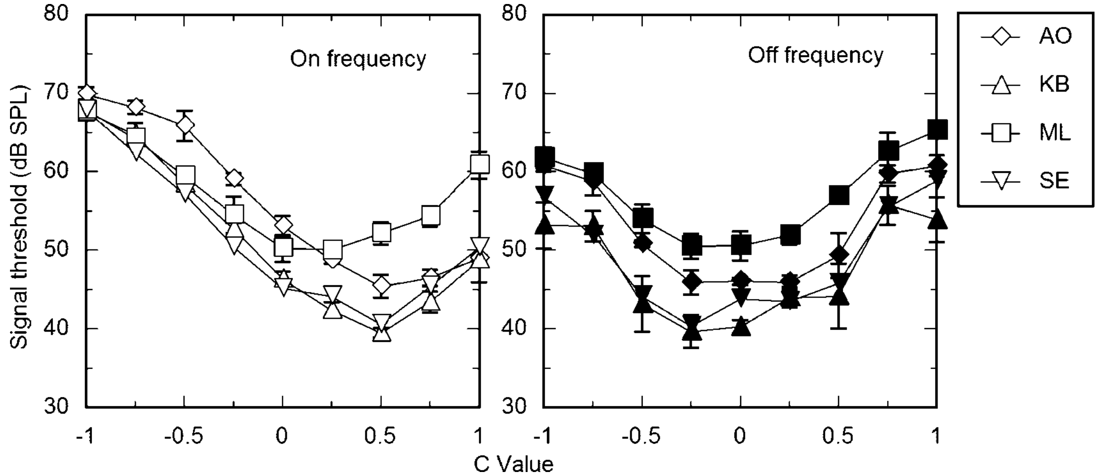

FIG. 1. Individual signal thresholds in the presence of the on-frequency (left panel) and off-frequency (right panel) maskers. Error bars denote ± 1 standard deviation across the individual repetitions. In line with predictions, masking minima are reached for a positive phase curvature for on-frequency masker, but are closer to zero for the off-frequency masker.

the present experiment before data were collected. Subject ML had no previous experience in psychoacoustic experiments and received about 2 h of training before data collection started. Subjects KB and ML were compensated for their services on an hourly basis.

III. RESULTS AND DISCUSSION

The individual results are shown in Fig. 1. The left panel shows results with the on-frequency masker; the right panel shows results with the off-frequency masker. Different symbols represent thresholds from different subjects, as shown in the legend, and error bars denote ± 1 standard deviation of the individual estimates. Despite differences in individual thresholds of 10 dB or more, the patterns of results are generally similar across the four subjects. In particular, all four subjects show a minimum masked threshold at a positive C value for the on-frequency masker: AO, KB, and SE show a minimum at around $C=0.5$, while ML shows similar thresholds at $C=0$ and $C=0.25$, with a slightly lower value at 0.25. All four subjects show a minimum at a lower C value for the off-frequency masker: individual minima are either at $C=0$ or $C=-0.25$.

The differences between the on- and off-frequency conditions are seen clearly in a direct comparison of the mean results (Fig. 2). The two curves are fitted sinusoidal functions, as described in Oxenham and Dau (2001b). The minima of the fitted functions were 0.42 and -0.05 for the on- and off-frequency conditions, respectively. The masker curvature at these minima in rad/Hz^2 [as calculated from Eq. (2)] can be normalized into dimensionless units by multiplying by $f_s^2/2\pi$, where f_s is the signal frequency (Oxenham and Dau, 2001b; Shera, 2001). The normalized masker curvature at the minimum of the masking function for the present on-frequency data is 12.9, implying a normalized cochlear filter phase curvature of -12.9, which is in good agreement with previous estimates at 2000 Hz, using different subjects and different masker F0s and bandwidths, which range between -8 and -16 (Lentz and Leek, 2001; Oxenham and Dau, 2001b). The novel condition tested here involves the off-frequency masker. Here, the estimated curvature using the same signal frequency is close to zero, in clear contradiction to an assumption that curvature remains constant with frequency at a given CF.

Overall, the data support the hypothesis outlined in the introduction:

(1) Masking patterns for the off-frequency masker produce a minimum at around zero phase curvature, suggesting no phase curvature in the cochlear filter response to stimuli well below CF.
(2) Masking patterns for the on-frequency masker produce a minimum at a positive masker phase curvature, presumably counteracting the negative phase curvature of cochlear filtering around CF.

A phase versus frequency plot of these data (not shown) is consistent with physiological data of basilar-membrane motion at the basal end of the cochlear, in showing a nearly straight-line section for frequencies half-an-octave or more below CF and a more curved negative-going section for frequencies around CF (de Boer and Nuttall, 1997; Ruggero *et al.*, 1997; Rhode and Recio, 2000; Shera, 2001). The situ-

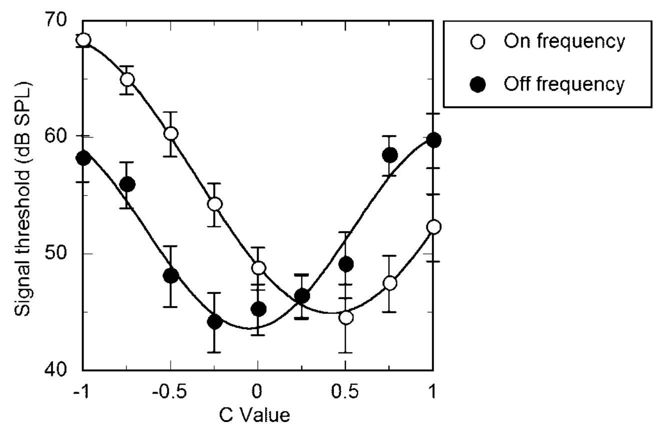

FIG. 2. Mean signal thresholds, along with fitted sinusoidal functions. Open symbols represent thresholds from the on-frequency condition; filled symbols represent thresholds from the off-frequency condition. Error bars denote ± 1 standard error of the mean across the four subjects. Minima of the fitted functions are reached for C values of 0.42 and -0.05 for the on- and off-frequency conditions, respectively.

ation at signal frequencies well below the one tested here (2000 Hz) is likely to be different, in that the phase curvature tends to zero even in the on-frequency case (Oxenham and Dau, 2001b).

One assumption of our technique is that the masker phase curvature that produces the most modulated temporal envelope is independent of the relative component amplitudes. This assumption is important because component amplitudes can be varied by experimental manipulations as well as by outer-, middle-, and inner-ear filtering. Numerical simulations have confirmed this assumption in all conditions tested so far when the Hilbert envelope is used. However, the assumption does not always hold when measures such as the crest factor (Strickland and Viemeister, 1996) are used directly on the stimulus waveform, rather than the envelope. Although it is universally assumed that signal detection in these situations is based on envelope, rather than raw waveform, properties, we cannot completely rule out some influence of temporal factors not represented in the Hilbert envelope. Empirical tests are required to settle this question, although informal testing has not yet yielded any situations where the phase curvature producing minimum threshold is affected by the relative amplitudes of the masker components.

One potential limitation of the present technique is that it does not allow a detailed view of how the curvature changes with frequency; we have only shown the estimated curvature for two fairly broad sample frequency regions. By using maskers with nonconstant phase curvature, it may be possible to map out changes in curvature with frequency in more detail. However, it may be that inherent measurement errors would swamp more fine-grained effects than those found here. A method for using envelope modulation in the auditory nerve to map out cochlear phase responses has been proposed recently (van der Heijden and Joris, 2003). By using the modulation at the beat frequencies between unequally spaced sinusoids, it was possible to derive a much more detailed phase response even in cochlear regions where the auditory nerve fibers no longer phase-lock to the temporal fine structure. While our technique and that of van der Heijden and Joris (2003) both use aspects of envelope modulation, theirs has the advantage of providing a much more detailed function of phase versus frequency. Unfortunately, there does not seem to be any obvious way of translating their technique into a task suitable for behavioral experiments.

In summary, masked thresholds for a sinusoidal signal in a harmonic tone complex vary with masker phase curvature and masker frequency region as would be predicted from animal physiological studies of basal cochlear mechanics: on-frequency maskers indicate negative phase curvature, while maskers well below the signal frequency indicate near-zero phase curvature. The results suggest that it is possible—to a limited extent—to map out changes in human cochlear phase response with frequency. Such results should provide important tests for computational models of human cochlear filtering.

ACKNOWLEDGMENTS

This work was supported primarily by the National Institutes of Health (Grant No. R01 DC 03909). It was begun while the second author was a visiting scientist at MIT's Research Laboratory of Electronics, and was completed while the first author was a fellow at the Hanse Institute for Advanced Study, Delmenhorst, Germany. We thank Torsten Dau, Van Summers, and an anonymous reviewer for helpful comments on previous versions of this manuscript.

[1] If cosines are used as the basis functions, the negative sign in the equation leads to a simple time reversal of the overall waveform. If sines are used as the basis functions (as was done in this study), the negative sign leads to a time-reversal and inverting of the overall waveform. This is because the sine and cosine functions are dot-symmetric and symmetric about zero, respectively.

Carlyon, R. P., and Datta, A. J. (**1997**). ''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.

Kohlrausch, A., and Houtsma, A. J. M. (**1992**). ''Pitch related to spectral edges of broadband signals,'' Philos. Trans. R. Soc. London, Ser. B **336**, 375–382.

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.

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.

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

Rhode, W. S., and Recio, A. (**2000**). ''Study of mechanical motions in the basal region of the chinchilla cochlea,'' J. Acoust. Soc. Am. **107**, 3317–3332.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Schroeder, M. R. (**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.

Strickland, E. A., and Viemeister, N. F. (**1996**). ''Cues for discrimination of envelopes,'' J. Acoust. Soc. Am. **99**, 3638–3646.

Summers, V., and Leek, M. R. (**1998**). ''Masking of tones and speech by Schroeder-phase harmonic complexes in normally hearing and hearing-impaired listeners,'' Hear. Res. **118**, 139–150.

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.

van der Heijden, M., and Joris, P. X. (**2003**). ''Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies,'' J. Neurosci. **23**, 9194–9198.