Comparison of level discrimination, increment detection, and comodulation masking release in the audio- and envelope-frequency domains

Paul C. Nelson
Department of Biomedical and Chemical Engineering and Institute for Sensory Research, Syracuse University, Syracuse, New York 13244

Stephan D. Ewert
Centre for Applied Hearing Research, Technical University of Denmark, Lyngby, Denmark

Laurel H. Carney
Department of Biomedical and Chemical Engineering and Institute for Sensory Research, Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, New York 13244

Torsten Dau
Centre for Applied Hearing Research, Technical University of Denmark, Lyngby, Denmark

(Received 3 May 2006; revised 20 December 2006; accepted 10 January 2007)

In general, the temporal structure of stimuli must be considered to account for certain observations made in detection and masking experiments in the audio-frequency domain. Two such phenomena are (1) a heightened sensitivity to amplitude increments with a temporal fringe compared to gated level discrimination performance and (2) lower tone-in-noise detection thresholds using a modulated masker compared to those using an unmodulated masker. In the current study, translations of these two experiments were carried out to test the hypothesis that analogous cues might be used in the envelope-frequency domain. Pure-tone carrier amplitude-modulation (AM) depth-discrimination thresholds were found to be similar using both traditional gated stimuli and using a temporally modulated fringe for a fixed standard depth (m_s=0.25) and a range of AM frequencies (4–64 Hz). In a second experiment, masked sinusoidal AM detection thresholds were compared in conditions with and without slow and regular fluctuations imposed on the instantaneous masker AM depth. Release from masking was obtained only for very slow masker fluctuations (less than 2 Hz). A physiologically motivated model that effectively acts as a first-order envelope change detector accounted for several, but not all, of the key aspects of the data. [DOI: 10.1121/1.2535868]

PACS number(s): 43.66.Mk, 43.66.Dc [JHG] Pages: 2168–2181

I. INTRODUCTION

When discussing temporal cues in sounds, it is necessary to differentiate features that change on different time scales. On a relatively short time scale, a sound's pressure waveform fluctuates about zero; these variations are referred to as fine structure and are determined by the instantaneous audio frequency of the stimulus. In contrast, a signal's temporal envelope changes on a longer time scale and is always positive; the envelope is a description of the slow variations in overall level that define the instantaneous amplitude. It is useful to describe complex temporal amplitude modulations in the envelope-frequency domain. Virtually all natural sounds have complex audio-frequency domain spectra and complex envelope-frequency domain spectra.

A variety of fundamental experimental paradigms originally used in audio-frequency psychoacoustics have recently been translated into their envelope-frequency equivalents. In the process, certain parallels have emerged between the effective signal processing that is inferred to take place in the two domains. Masked tone-detection experiments that compare the amount of masking with different spectral relationships between the signal and masker indicate perceptual frequency tuning in both audio frequency (e.g., Wegel and Lane, 1924) and envelope frequency (e.g., Houtgast, 1989). Also, the "asymmetry of masking" has been observed in both domains [e.g., Moore *et al.*, 1998; Derleth and Dau, 2000 (audio frequency); Ewert *et al.*, 2002 (envelope frequency)], with tones proving to be relatively ineffective maskers of noise signals compared to the masking effect on tonal signals produced by a noise masker. Wojtczak and Viemeister (2005) showed nonsimultaneous (forward) masking in the envelope-frequency domain that has direct counterparts in audio-frequency psychophysics (e.g., Lüscher and Zwislocki, 1949). The ability to resolve tone level is also broadly similar across domains, with an approximately 1–2-dB increase in the standard level (SPL in audio frequency, $20 \log m$ in envelope frequency) required to reliably discriminate the levels of two supra-threshold sounds [e.g., Florentine *et al.*, 1987 (audio); Ewert and Dau, 2004 (envelope)]. These qualitative similarities suggest that the processes fundamental to perception in the two domains may be conceptualized in a single framework, despite the fact that the underlying mechanisms may be quite different.

 0001-4966/2007/121(4)/2168/14/$23.00

In this study, the hypothesis that two other robust audio-frequency phenomena would be observed in the envelope-frequency domain was tested. These phenomena are (1) a heightened sensitivity to increments with a continuous carrier (or a temporal fringe) relative to gated-carrier level discrimination performance and (2) lower thresholds in a tone-in-noise detection task with a temporally amplitude-modulated (AM) masker than in conditions with a random (unmodulated) masker. The second observation has been termed comodulation masking release (CMR) in the audio-frequency domain because it is most robust when several frequency channels are simultaneously and coherently modulated. Both audio-frequency observations can be at least partially attributed to AM-related cues (e.g., Gallun and Hafter, 2006; Schooneveldt and Moore, 1989). Therefore, upon transposition into the envelope-frequency domain, the analogous cues in such tasks would be related to the second-order envelope (Lorenzi *et al.*, 2001), or "venelope" (Ewert *et al.*, 2002).

The current understanding of venelope perception can be summarized as follows: Sinusoidal modulation of the depth of a first-order AM carrier is detectable under some conditions (e.g., Lorenzi *et al.*, 2001), but the perceptual salience of venelope components is generally found to be weaker than that of first-order envelope fluctuations (Ewert *et al.*, 2002). Venelope fluctuations can interact with envelope detection and vice versa (e.g., Moore *et al.*, 1999; Ewert *et al.*, 2002); this finding can be qualitatively accounted for with several physiologically realistic nonlinearities that effectively transform second-order envelope components into first-order envelope components (Shofner *et al.*, 1996). Alternatively, venelope cues could take the form of temporal variations in the output of envelope-frequency-tuned modulation filters (Ewert *et al.*, 2002; Füllgrabe *et al.*, 2005; Füllgrabe and Lorenzi, 2005), which could interfere at some later stage of processing with the output of the modulation filter tuned to the signal frequency.

In the audio-frequency domain, listeners are more sensitive to level differences presented as continuous-carrier level increments than presented as gated tones with different SPLs (e.g., Campbell and Lasky, 1967; Viemeister and Bacon, 1988; Bacon and Viemeister, 1994). An energy-based detection model cannot explain the difference in thresholds in the two conditions. Instead, the temporal structure of the standard, target, and interstimulus intervals must be taken into account. This finding can be considered from several perspectives. One possibility is that the memory requirements of the system are higher in gated-carrier level discrimination than for increment detection, where listeners could potentially perform the task without comparing across intervals even in a two-alternative forced-choice task (Harris, 1963). This explanation is less than satisfactory because, near threshold, there is certainly an element of comparison across intervals even in the continuous-carrier task: the listener must decide which interval sounded the most like it contained an increment or "bump." Also, the relatively short intervals between stimuli probably render memory noise negligible with respect to sensation noise in most two-or three-interval paradigms (Durlach and Braida, 1969).

Another related explanation holds that the improved sensitivity results because the system could be making decisions by detecting changes in the increment task (Macmillan, 1971; Hafter *et al.*, 1997). Onsets and offsets of gated stimuli result in excitation of the putative modulation filterbank (e.g., Dau *et al.*, 1997) in both standard and target intervals that depends on the shape and duration of the ramps applied to the carrier (Gallun and Hafter, 2006). Increment detection paradigms using a continuous pedestal, on the other hand, cause a change in the signal envelope only in the target interval. As a result, transient onset (and offset) responses are present in only one of the observation intervals. Also, gated paradigms require the listener to identify the interval containing the more intense sound, while continuous-carrier level discrimination can be performed without knowing the direction of the change in level (Hafter *et al.*, 1997).

Physiologically, absolute firing-rate changes in single auditory-nerve fiber (ANF) responses to increases in SPL do not depend on the temporal position of the increment with respect to the onset of the pedestal (e.g., Smith and Zwislocki, 1975; Smith and Brachman, 1982; Smith *et al.*, 1985). Instead, the *relative* increase in instantaneous rate increases as the delay between pedestal onset and increment is lengthened, because the response to the pedestal decreases with time. If it is assumed that a constant fractional increase in the response is required to reach threshold, these findings can also qualitatively account for the perceptual gated-continuous difference.

Fringe effects in level discrimination have provided a long-standing challenge for pure power-spectrum models of audio-frequency processing; tone-in-noise detection tasks that compare masking by modulated and unmodulated maskers have emerged more recently as challenges to such long-term energy-based models (e.g., Schooneveldt and Moore, 1989; Verhey *et al.*, 1999). This article focuses on a simple class of CMR experiments: those that use a single noise masker centered on the signal frequency. This class of CMR paradigms yields the most significant and robust release from masking when the masker is broadband and fully modulated (Verhey *et al.*, 2003). Several cues could potentially underlie a release from masking (i.e., lower thresholds with a modulated masker compared to unmodulated masker conditions). A "dip-listening" model suggests that the listeners are able to selectively attend to the periods of the stimulus with low masker amplitudes, thus improving the local signal-to-noise ratio (SNR; Buus, 1985). Another possible cue, which is mainly based on the processing in the peripheral channel tuned to the signal frequency (within-channel processing), is the overall smoothing of the masker fluctuations upon addition of the signal (Schooneveldt and Moore, 1989; Verhey *et al.*, 1999). Across-peripheral-channel comparisons of target-interval differences might also be used if the bandwidth of the masker is sufficiently broad (e.g., Hall *et al.*, 1984). All of these mechanisms have been used to understand CMR in the audio-frequency domain.

Based on these empirical audio-frequency observations, the current study presents envelope-frequency-domain versions of the experiments that led to them. Listeners' access to venelope cues should determine whether differential effects

will be observed in (1) continuous- and gated-carrier AM depth discrimination and (2) sinusoidal AM (SAM) detection in the presence of a noise AM masker with and without slow and regular fluctuations in overall modulation depth. The remainder of this article is divided into three main sections. Two lines of psychophysical experiments are described and discussed in the first two sections. The third part focuses on interpretation of the findings with the help of a physiologically motivated computational model.

II. EXPERIMENT I. LEVEL DISCRIMINATION AND INCREMENT DETECTION IN THE ENVELOPE-FREQUENCY DOMAIN

The goal of the first set of experiments was to determine whether continuous-carrier AM-depth-discrimination thresholds were lower than traditional gated-carrier thresholds. Two reasonable hypotheses lead to predictions of a difference in performance between the two paradigms. First, adaptation at the output of modulation-tuned channels could have the effect of masking across-interval depth differences, since neural responses often exhibit more variability at higher response amplitudes (e.g., Young and Barta, 1986; see Sec. IV of the current study). This would result in poorer performance in the gated conditions. Alternatively, because energy increment detection in the audio-frequency domain is at least partly associated with modulation detection (Wojtczak and Viemeister, 1999) and coded along the modulation dimension (Gallun and Hafter, 2006), a corresponding task in the envelope domain may provide another cue along an additional dimension, the hypothetical venelope dimension, which might lead to lower thresholds in the continuous-carrier condition relative to the gated case.

A. Methods

1. Listeners

The experiments were carried out at the Centre for Applied Hearing Research at the Technical University of Denmark (DTU). All of the listeners participated voluntarily and had pure-tone detection thresholds less than 20 dB HL at octave frequencies between 125 and 8000 Hz. Their ages ranged from 23 to 39 years. Three of the four subjects in the main experiments had significant experience in related psychoacoustic testing; two of the authors (PCN, TD) were part of this group. Four additional listeners were recruited to participate in the extension of experiment I (Sec. II B 2) because of the relatively large across-subject variability.

2. Apparatus and stimuli

Subjects listened diotically via Sennheiser HD 580 circumaural headphones in a double-walled, sound-attenuating booth. Stimulus generation and presentation were carried out in MATLAB using the AFC software package developed at the University of Oldenburg and at DTU. A 48-kHz sampling rate was used to digitally generate stimuli. The carrier was a 70-dB SPL, 5.5-kHz pure tone. Sinusoidal AM was applied for the entire 500-ms duration, and a 50-ms raised-cosine window was applied at the onset and offset of observation-interval stimuli. Inter-observation-interval durations (between possible target interval presentations) were also 500 ms in duration.

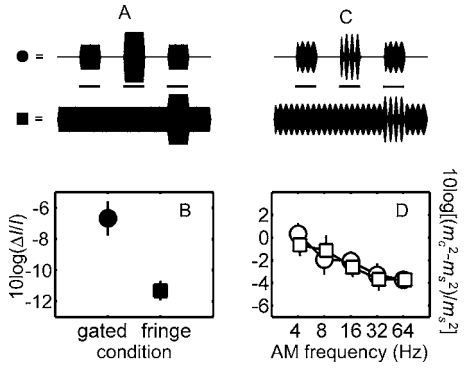

FIG. 1. Comparison of the gated-continuous difference in the audio (left) and envelope (right) frequency domains. Schematic illustrations of stimulus waveforms in the two experiments are shown in (a) and (c) (horizontal bars between the stimuli indicate the timing of the 500-ms observation intervals). (b) Audio-frequency level discrimination thresholds measured with gated (closed circles) and quasi-continuous pedestals (closed squares). (d) AM depth discrimination thresholds for a modulated standard (m_s=0.25) obtained with traditional gated intervals (open circles), and with quasi-continuous modulation presented before, between, and after the observation intervals (open squares). For both (b) and (d): f_c=5500 Hz; standard SPL =70 dB. Each symbol is the average threshold for four listeners; error bars indicate ±1 standard deviation of the individual mean thresholds.

For comparison to the envelope-domain results, thresholds for the audio-frequency versions of pure-tone gated level discrimination and quasi-continuous increment detection (i.e., with a temporal fringe) were also measured in the same subjects. Signal duration, inter-observation-interval, gating parameters, SPL, and carrier frequency were the same as in the envelope-frequency experiments. The corresponding example stimulus waveforms are shown in Fig. 1(a).

The modulating waveforms for the AM depth discrimination paradigm in the gated conditions were identical to those described in Ewert and Dau (2004). The observation-interval stimuli are described by the following equation:

$$s(t) = \sin(2\pi f_c t)[1 + m_s\sqrt{1 + m_{\text{inc}}}\sin(2\pi f_m t)],$$

where f_c is the carrier frequency (5500 Hz), m_s is the standard modulation depth, m_{inc} is the relative depth increment (zero in the standard intervals), and f_m is the modulation frequency. The comparison (target interval) depth can be related to the standard depth and depth increment simply as $m_c = m_s\sqrt{1+m_{\text{inc}}}$. Using a notation more in line with audio-frequency level discrimination literature, m_{inc} can also be thought of as the Weber fraction, i.e., $m_{\text{inc}} = (m_c^2 - m_s^2)/m_s^2$. Whereas the earlier study (Ewert and Dau, 2004) focused on the effects of standard-interval modulation depth for a fixed-frequency (16-Hz) sinusoidal AM, the current experiments used a fixed standard depth (m_s) of −12 dB (in $20\log m$; linear m=0.25) and varied two other parameters. Here, the influences of modulation frequency (f_m=4, 8, 16, 32, and 64 Hz) and gating choices were examined. The traditional AM depth-discrimination stimuli (e.g., Wakefield and Viemeister, 1990; Lee and Bacon, 1997; Ewert and Dau, 2004) are referred to as "gated" and the envelope-domain equivalent of increment

detection as "quasi-continuous" or "fringe" conditions.

The critical difference between the gated and fringe conditions was confined to the stimulus presented between observation intervals in the three-interval paradigm. In gated conditions, a silent interval separated three modulated tones (the two standard intervals contained tones with a modulation depth m_s of 0.25; the target interval was a tone with some AM depth higher than m_s). In contrast, the quasi-continuous conditions were comprised of a 500-ms modulated tone (m_s=0.25) that was present in the two inter-observation intervals and also before the first interval and after the third and final interval. Example stimulus waveforms for the gated (top) and fringe conditions (bottom) are shown in Fig. 1(c). Stimulus amplitudes in all three intervals were gated with 50-ms ramps, regardless of the gating mode (this was also the case for the audio-frequency level discrimination stimuli).

3. Procedure

Listeners were trained until four consecutive threshold estimates in each condition showed no evidence of learning. Two additional threshold estimates were obtained if the standard deviation of the four estimates was greater than 3 dB (this happened once in all of the experiments described here). Average data are presented as the mean and standard deviation of the subjects' final depth-discrimination threshold estimates.

A three-interval, three-alternative forced-choice paradigm with visual correct-answer feedback was used along with a two-down, one-up adaptive tracking procedure (Levitt, 1971). This combination of parameters yields convergence on the 70.7% point of the psychometric function and a threshold estimate that corresponds to a d' of unity. The listeners' task was to identify the observation interval containing the higher signal AM depth. Observation-interval timing was identified with visual cues presented synchronously with the standard and target interval stimuli on the computer monitor. The stimulus parameter that was varied in the tracking procedure was the fractional AM depth increment in dB ($10 \log m_{\text{inc}}$). The initial 4-dB signal-interval step size was halved after each of the first two track reversals occurring after consecutive correct responses. Six reversals were required after the final 1-dB step size was reached; threshold for a given track was taken as the mean signal level corresponding to the target-interval AM depth used at those six points. The order of stimulus presentation was randomized across parameters (gating mode and AM frequency) for each listener.

The audio-frequency level-discrimination experimental procedures were essentially identical to those used to measure AM depth-discrimination sensitivity. The tracking variable used was also similar. The Weber fraction in dB ($10 \log \Delta I/I$) was adjusted until the target interval was just noticeably different from the two standard observation intervals.

B. Results

1. Discrimination thresholds with gated and fringe presentation modes

The magnitude of the audio-frequency gated-continuous difference was measured first; mean level discrimination results are shown in Fig. 1(b). Enhanced sensitivity to increment (fringe) conditions has been demonstrated in previous studies; the average difference for the listeners in the current study was 4–5 dB. The magnitude of the effect in our listeners was in line with the average 4.6-dB difference found at SPLs above 35 dB by Viemeister and Bacon (1988), who used a continuous 1000-Hz carrier and 200-ms observation intervals. Absolute discrimination thresholds in the gated conditions in the current study ($10 \log \Delta I/I$=−6.5 dB) are slightly better than the thresholds for matching carrier frequencies and standard levels reported by Florentine *et al.* (1987); this may be attributable to differences in presentation mode (monaural in Florentine *et al.* versus diotic in the current study).

Average modulation-depth discrimination thresholds are shown in Fig. 1(d) for a range of modulation frequencies. The most relevant aspect of the data for the purposes of the current study is the similarity in performance for the gated (circles) and fringe conditions (squares), which is in contrast to the findings in the audio-frequency domain. Performance was broadly consistent across listeners, as suggested by the size of the standard deviation bars (<1.6 dB). Listener L4 was slightly more sensitive in the fringe conditions, while L2 exhibited lower thresholds in the gated conditions. Because these individual differences were similar in magnitude and stable across AM frequency for both listeners, they effectively cancelled out in the mean data.

Mean AM-depth Weber fractions [in $10 \log((m_c^2 - m_s^2)/m_s^2)$] dropped from approximately 0 dB at 4 Hz to −4 dB at 32 and 64 Hz. These values are equivalent to target-interval depths at threshold ranging from m_c=0.35 to 0.30 for a modulated standard (m_s=0.25) and are consistent with previous studies that have found decreases in threshold at higher AM frequencies with a fixed-duration stimulus (i.e., Lee and Bacon, 1997). Thresholds in the gated condition at 16 Hz (−2.1 dB) were within 1 dB of those reported by Ewert and Dau (2004), who used a 16-Hz signal with a standard depth of m_s=0.225, among others, imposed on a 65-dB SPL, 4-kHz carrier.

2. Gated and fringe AM detection thresholds and comparison with "static" level discrimination performance

Previous studies have reported an enhancement of SAM detection thresholds at low modulation rates ($\leqslant \sim 10$ Hz) when a temporal fringe was used instead of gating the carriers (Viemeister, 1979; Yost and Sheft, 1997). The current finding of identical discrimination thresholds in the first experiment (with m_s=0.25) appears incompatible with these earlier findings. In order to determine whether the listeners in the current study also exhibited a gated-continuous difference for AM detection, an extension of the first experiment was added: thresholds in several related envelope-processing

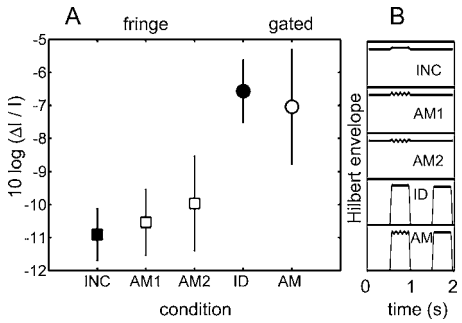

FIG. 2. (a) Mean audio-frequency level discrimination thresholds (solid symbols) and envelope-frequency detection thresholds (open symbols) under different gating conditions. Squares represent performance with a quasi-continuous carrier; circles correspond to thresholds with gated carriers. Conditions AM1 and AM2 are distinguishable based on the presence (AM1) or absence (AM2) of a dc component in the target interval. (b) Schematic illustrations of the stimulus envelopes used in each condition. Conditions INC and ID are re-plotted from Fig. 1(b). Error bars represent across-listener standard deviations (as in Fig. 1).

tasks were directly compared in an effort to better map out the differences between AM discrimination and detection, and to establish relationships between "dynamic" AM cues and "static" audio-frequency level discrimination. The question addressed is how gated and continuous level discrimination and AM detection thresholds are related. Based on the results of the first experiment, one hypothesis is that AM detection should *not* depend on the choice of gating parameters if the same mechanism underlies AM detection and AM depth discrimination.

To test this hypothesis, the dynamic AM stimuli were matched to the static level-increment stimuli in their carrier frequency (5.5 kHz), standard level (70 dB SPL), overall duration (500 ms), and onset/offset ramp duration (50 ms). Restricting the ramp duration parameter to be the same resulted in an AM stimulus that contained five cycles of 10-Hz SAM signal (50 ms on+50 ms off=100 ms period). Different combinations of gated versus fringe presentation modes and static versus dynamic fluctuations in the target interval resulted in the five test stimuli used in the extension of experiment I. Schematics of the envelope waveforms in each condition are shown in Fig. 2(b). For clarity, a two-interval task is represented (a 3AFC task was used in the actual experiment), with the target interval onset beginning at 0.5 s and the standard interval beginning at 1.5 s. Conditions with temporal fringes were identified as INC (static increment), AM1 (dynamic AM increment, with a dc offset), and AM2 (dynamic AM increment, no dc offset). Gated stimuli were labeled ID (static gated intensity discrimination) and AM (gated AM detection).

Mean data across eight listeners are shown in Fig. 2(a). Performance is defined in terms of a Weber fraction, where ΔI is determined by the difference between the maximum and minimum values of the envelope in the target interval for the AM conditions (open symbols), and by the difference in peak intensities across the standard and target intervals in the audio-frequency level-discrimination conditions (closed symbols).

One main result is that thresholds were similar for all three fringe conditions INC, AM1, and AM2 [Fig. 2(a)], and for both gated conditions ID and AM (t test p values for these comparisons were all greater than 0.13). This finding suggests that the system was not more sensitive when several dynamic temporal envelope fluctuations were presented than when a fixed energy increment with a single onset and offset was used. The similarity between thresholds in AM1 and AM2 suggests that the listeners were probably not using an overall level cue in condition AM1. The finding that the INC and AM1 and AM2 stimuli produce similar thresholds supports the hypothesis that increment detection is linked to modulation detection (and not primarily based on the detection of an energy change).

Another comparison to make in Fig. 2(a) is across the open symbols (conditions with SAM in the target interval). Listeners were more sensitive in a low-f_m AM-detection task with a temporal fringe (AM2) than with gated carriers (AM). This is in line with noise-carrier studies of Viemeister (1979) and Sheft and Yost (1990), and with the tone-carrier experiments of Yost and Sheft (1997), but seemingly at odds with the result from the depth-discrimination task, where gating parameters had no effect for discrimination of a suprathreshold AM depth. Converting the thresholds to $20 \log m$, the difference between thresholds in the AM2 condition (−32 dB) and the gated AM condition (−26.5 dB) amounts to about 5.5 dB.

C. Discussion

1. Adaptation and change detection

The similarity between gated and quasi-continuous AM depth discrimination thresholds can be interpreted in terms of the adaptation mechanisms that have been used to qualitatively explain the audio-frequency asymmetry in performance seen in gated-carrier level discrimination and continuous-carrier increment detection (see the Introduction). If an increased amount of adaptation in gated conditions underlies gated-continuous differences, then the current results suggest one of at least two conclusions in the envelope-frequency domain. Either there is little or no adaptation at the output of modulation-tuned channels or, if there is adaptation, then the response to an increment in AM depth must decrease with the same time course as the adaptation, so that the relative response increment remains constant as a function of time.

There is some peripheral physiology that initially appears consistent with a transformation supporting the latter interpretation. Smith *et al.* (1985) reported a decrease in the AM response modulation of ANFs as a function of time: the response modulation depth decreased with short-term adaptation (i.e., the effect lasted for approximately 10 ms). In contrast with the current study, Smith *et al.* (1985) used stimuli with high AM frequencies (150–600 Hz) and imposed the modulation on gated carriers with short (2.5 ms) rise-fall times. For the lower fluctuation rates (4–64 Hz) and slow (50 ms) ramp functions used here, it is unlikely that the small effect observed in peripheral physiology could have an impact on the observed similarity between gated and continuous AM depth-discrimination thresholds. This leads back to the alternative explanation, namely that there is negligible

perceptual adaptation to AM stimulation observed with the ramp and exposure durations used in the current study.

Perceptual coding of AM is usually assumed to be strongly influenced by central processing factors. This is because modulation-tuned channels are not found in the periphery, and the temporal responses of ANFs can robustly follow modulations to significantly higher rates than the several hundred Hertz (Joris and Yin, 1992) that human listeners can detect as a temporal (i.e., not spectrally resolved) cue (Kohlrausch *et al.*, 2000). The responses of cells in the inferior colliculus (IC) appear to be more tightly coupled to psychophysical measures than peripheral responses (Joris *et al.*, 2004), and temporal adaptation in responses of IC neurons to AM stimuli with relatively long onset and offset ramps is often negligible (Nelson and Carney, 2007).

An alternative way to conceptually account for the audio-frequency gated-continuous difference is to assume the existence of a modulation filter bank that processes the stimuli at the output of peripheral filters, generating an effective additional dimension. An increment in the SPL of a sound activates at least the low-frequency modulation channels, where the amount of activity depends on the exact stimulus characteristics and the transfer functions of the modulation filters. As recently shown by Gallun and Hafter (2006), increment detection thresholds can be quantitatively accounted for by assuming a modulation-frequency selective analysis. In contrast, in the gated-carrier level-discrimination conditions, the most effective cue is reflected in the dc component (or in the lowest available modulation filter) in such a model. The finding that a similar asymmetry between increment detection and level discrimination was not found in the AM domain may suggest that analogous circuitry, i.e., another "independent" (venelope) domain, does not exist, or has a negligible influence on perception.

2. Relation to previous work

Wojtczak and Viemeister (1999) compared intensity discrimination and low-f_m SAM detection with continuous-carrier pure tones across a wide range of standard SPLs and arrived at an empirical formulation of the relationship between the two measures: $10\log(\Delta I/I)=0.44(20\log m)+D(f_m)$, where $D(f_m)$ is a constant that depends only on modulation frequency. For a 4-Hz signal AM, Wojtczak and Viemeister determined this constant to be 1.7; for the 10-Hz signal AM used in the current study, $D(f_m)$ would probably take on a slightly lower value. With continuous 70-dB SPL carriers, a 10-Hz modulation rate, and assuming $D(f_m)$ to be 1.7, our data are reasonably consistent with the proposed empirical formula: $10\log(\Delta I/I)=-10.9$ dB (INC condition in Fig. 2 and $0.44(20\log m)+1.7=-12.5$ dB (AM2 condition in Fig. 2). Decreasing the value of $D(f_m)$ or inserting the modulation thresholds measured with the AM2 stimuli (SAM with a dc component) would make the equation's predictive ability worse. Wojtczak and Viemeister (1999) speculated that the empirical relationship might also hold for gated carriers, but did not test this hypothesis explicitly. The current data allow for such a test. In the gated intensity discrimination task (ID) here, $10\log(\Delta I/I)=-6.6$ dB, while in the gated AM detection condition (AM), $0.44(20\log m)+1.7=-10.0$ dB. The match to the proposed formula is worse in this condition, suggesting that it does not directly generalize to describe the relationship using gated stimuli.

III. EXPERIMENT II. TONE-IN-NOISE DETECTION WITH A MODULATED MASKER IN THE ENVELOPE-FREQUENCY DOMAIN

The goal of the second experiment was to determine the extent to which listeners could use slow and regular temporal fluctuations in the instantaneous depth of a (stochastic) masker modulation to aid in the detection of a (deterministic) sinusoidal signal modulation. The stimuli were designed to maximize the availability of potential release-from-masking cues in an envelope-domain transposition of a typical audio-frequency CMR paradigm.

A. Methods

Details of the listeners, apparatus, and procedure were the same as in the first set of experiments. This section addresses any remaining differences, which were mainly limited to stimulus parameters.

The carrier was again a 5.5-kHz pure tone with 50-ms raised-cosine windows applied to the onset and offset. The overall SPL of the standard and target were normalized to have the same rms level as a 70-dB SPL pure tone. Observation intervals were separated by a 500-ms silent interval. In the two standard intervals, the tonal carrier was modulated by a Gaussian noise, which had a 120-Hz bandwidth (BW) and was geometrically centered around 32 Hz, from 8 to 128 Hz. The average masker modulation depth was −13.2 dB rms ($m=0.22$; for a 120-Hz BW the noise had a "spectrum level" of −34 dB). This combination of masker depth and BW was chosen to (1) ensure significant masking of the 32-Hz signal AM (presented only in the target interval), (2) avoid overmodulation (no stimuli with modulation depths greater than one were presented to the listeners), and (3) provide a potential opportunity for across-modulation-channel processes to enhance detection performance [by using a BW greater than that of the putative modulation filters, which are typically described as having half-power Q-values between 0.5 and 2 (e.g., Lorenzi *et al.*, 2001; Wojtczak and Viemeister, 2005), or effective BWs between 16 and 64 Hz for a channel tuned to 32 Hz].

Masker waveforms in each interval were independent noise realizations, generated digitally by setting the Fourier coefficients outside the desired pass-band to zero. In the baseline conditions (analogous to unmodulated conditions in audio-frequency CMR experiments), no further manipulations were made of the masker waveform before the 32-Hz sinusoidal AM (always in sine phase) was added and the resulting envelope signal imposed on the carrier. A general equation for the final stimulus is

$$s(t) = c\{\sin(2\pi f_c t)[1 + m\sin(2\pi f_m t) + M(t)]\},$$

where c is a scalar that equalizes the overall audio-frequency power in each interval, f_c is the carrier frequency, m is the target modulation depth (zero in the standard interval), f_m is the signal modulation frequency (32 Hz), and $M(t)$ is the

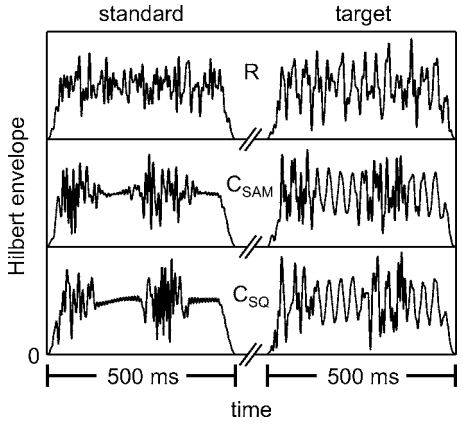

FIG. 3. Example temporal envelopes of the stimuli used to test for envelope-domain comodulation masking release. Standard-interval envelopes (left) are defined by the masker-alone waveform; target interval envelopes (right) are made up of an additive combination of masker and sinusoidal signal AM. R: Baseline (unmodulated, or random masker) condition. C_{SAM}: Sinusoidal venelope fluctuations. C_{SQ}: Square-wave venelope fluctuations. For all stimuli, f_c=5500 Hz; standard SPL=70 dB; masker BW =120 Hz, geometrically centered on the 32-Hz signal frequency; observation interval duration=500 ms; signal depth m=40%; venelope fluctuation rate=4 Hz.

masker waveform. In contrast to the baseline condition, the comparison, or "comodulated," masker waveforms were further processed before being combined with the signal SAM. Slow, coherent, and regular (sinusoidal or square wave) temporal fluctuations were imposed on the instantaneous masker modulation depth, $M(t)$. This resulted in a stimulus with a time-varying venelope. In all of the comparison conditions, the imposed venelope fluctuations were maximal in the sense that the nominal envelope depth of the masker varied between zero and the peak value. Examples of the time waveforms that were used to modulate the carrier are shown in Fig. 3; masker-alone (standard) waveforms are illustrated on the left, signal-plus-noise envelopes are shown on the right. Baseline unmodulated masker (R), sinusoidally modulated masker (C_{SAM}), and square-wave-modulated masker (C_{SQ}) conditions are shown.

Imposing slow fluctuations in the envelope can affect the resulting modulation spectra (i.e., sidebands are generated when the envelope is modulated, just as they are in the audio-frequency spectrum when a carrier is modulated). One way to avoid this complication is to filter the noise after it is modulated; the trade-off when using this strategy is that the temporal waveform is slightly changed, usually in the form of ringing caused by the band-limiting. To control for this issue, thresholds were measured when the masker envelope bandwidth was limited either before (condition C'_{SAM}) or after (condition C_{SAM}) imposing the slow venelope fluctuations in both the baseline and comparison conditions.

In the first part of the experiment, the venelope fluctuation rate was fixed at 4 Hz (two cycles were presented for each 500-ms signal), and the waveform used to modulate the (first-order) AM noise was varied, both in terms of its shape and its amplitude. In the extension of the experiment, the duration of the signals was extended to 2 s to allow for the use of even slower venelope fluctuation rates (from 0.5 to 4 Hz). An equal-energy (in terms of the envelope rms), square-wave venelope masker was used with the 2-s signals.

For comparison, thresholds were also measured for the same listeners in an audio-frequency CMR paradigm with parameters designed to (loosely) parallel those used in the envelope-frequency experiment. In both the audio- and envelope-domain experiments, detection thresholds of a mid-frequency sinusoidal signal embedded in a moderately intense and wideband (with regard to putative bandwidths of modulation or auditory filters) Gaussian masker were measured. Slow and regular fluctuations were imposed on the masker in both domains; release from masking was defined as improved thresholds in the conditions using modulated maskers over those using noises with flat temporal venelopes or envelopes.

Specific audio-frequency parameters were 2-kHz signal frequency, 800-Hz masker bandwidth (geometrically centered on the signal frequency), a masker spectrum level of 30 dB SPL (overall rms level=59 dB SPL), and a 32-Hz (first-order) sinusoidal AM imposed on the masker. Observation and interstimulus intervals were 500 ms. The tone level was adaptively varied initially in steps of 8 dB; the initial step size was halved after each of the first two track reversals occurring after consecutive correct responses until it reached 2 dB. Again, the mean of six reversals was taken as threshold for a given track. These parameters were chosen to maintain the same within-channel to across-channel energy ratios in the audio- and envelope-frequency CMR experiments. Assuming a typical 3-dB effective bandwidth of 200 Hz at 2 kHz in the audio-frequency task, the ratio of within-channel to across-channel energy was approximately 200 Hz: 800 Hz (or $\sim$1:4), which is similar to the ratio of 32 Hz: 120 Hz used in the envelope-domain task.

B. Results

1. No release from masking in the envelope-frequency domain for 500-ms *stimuli*

The magnitude of audio-frequency CMR with a wideband masker centered on the tone frequency and fully modulated by a deterministic waveform in the comodulated conditions is shown in the left panel of Fig. 4. Thresholds were about 10 dB lower in conditions with a comodulated masker (C) than in the random (flat masker) case (R). The magnitude of the effect is close to that observed in a previous study using similar stimulus parameters (Verhey *et al.*, 1999).

The new contribution of the current experiment was to translate the parameters that result in significant audio-frequency CMR into the envelope-frequency domain. Pure-tone carrier SAM detection thresholds were measured in the presence of several types of additive modulation maskers. A release from masking would take the form of lower thresholds in the conditions with slow and regular variations imposed on the instantaneous masker modulation depth when compared to performance in the conditions with a flat-envelope (Gaussian) masker modulation.

Within the right panel of Fig. 4, SAM detection thresholds are shown for the four masker conditions described

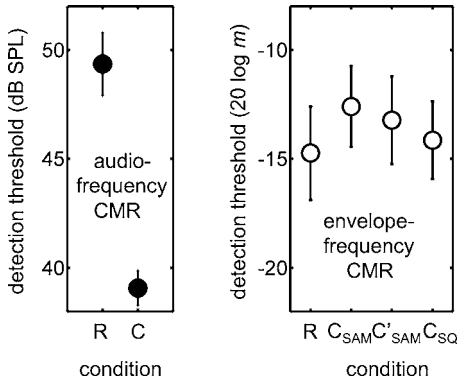

FIG. 4. Effects of imposing slow and regular fluctuations on the masker amplitude in the envelope- and audio-frequency domains. Conditions correspond to different temporal shapes imposed on the masker amplitude. Left panel: Audio-frequency thresholds. R: random flat masker envelope (unmodulated). C: 32-Hz SAM masker envelope, filtered after modulation and not equalized for overall energy increment caused by modulation. Right panel: Envelope-frequency thresholds. R: flat venelope masker (unmodulated). C_{SAM}: 4-Hz SAM venelope, noise filtered after modulation. C'_{SAM}: 4-Hz SAM venelope, noise only filtered prior to modulation. C_{SQ}: 4-Hz square-wave venelope, noise filtered after imposing the 4-Hz fluctuations. Conditions C_{SAM}, C'_{SAM}, and C_{SQ} were compensated for the small overall increase in masker energy caused by the modulation. Error bars indicate ±1 standard deviation of the individual mean thresholds

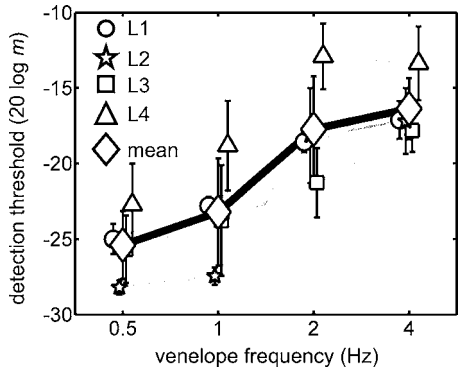

FIG. 5. Individual and mean 32-Hz SAM detection thresholds as a function of the frequency of the square-wave venelope fluctuations imposed on the first-order masker modulation. Stimulus parameters were the same as those in condition C_{SQ} of Fig. 4, except the overall duration was increased to 2 s. Error bars plotted with individual listener data represent across-track standard deviations; those with the mean data indicate across-listener standard deviations.

above. The average thresholds were all between −12 and −15 dB (20 log m), and none of the comodulated condition thresholds were significantly different from those measured in the random condition (t-test p values=0.18, 0.34, and 0.68). The results indicate that the listeners were unable to take advantage of the slow and regular venelope fluctuations imposed on the first-order masker.

2. Extending the time course of the slow masker fluctuations

In the square-wave venelope masker conditions above, the listeners were presented with two 125-ms segments of the unmasked pure SAM 32-Hz signal in the target interval (four complete cycles) between two 125-ms segments containing both the signal and masker modulation. This duration of pure-signal AM was insufficient to give rise to a release from masking. However, intuitively, one expects that there *must* be a release from masking if the periods of low masker energy are long enough. To further investigate the time course of the effect, it was necessary to increase the overall duration of each interval to 2 s to accommodate more than one cycle of the slow masker modulation. Square-wave venelope waveforms with rates of 0.5, 1, 2, and 4 Hz were imposed on the same modulation masker used with the 500-ms signals (120-Hz BW geometrically centered on the 32-Hz signal rate, with an average depth of −13.2 dB rms), and detection thresholds were determined again for a 32-Hz signal AM.

Detection thresholds for the 2-s stimuli are shown in Fig. 5. Individual thresholds are shown in addition to the mean results (diamonds) because of the relatively high intersubject variability. For all four of the listeners, performance improved with decreasing venelope fluctuation rates over the range of frequencies tested. The parameters of the stimuli used in the 4-Hz condition were identical to those used with the 500-ms stimuli (condition C_{SQ} in Fig. 4); mean performance improved by about 2.5 dB as a result of increasing the stimulus duration.

Thresholds asymptote near the expected pure-AM detection thresholds for listeners L1, L2, and L3; the remaining listener (L4) was less sensitive overall and continued to show improvement between the 1- and 0.5-Hz conditions. Overall, the listeners required at least 500-ms periods of unmasked SAM signal between masker bursts to reach performance near the 32-Hz SAM detection thresholds expected for a 70-dB SPL, 5.5-kHz pure tone carrier (between −25 and −30 dB; e.g., Kohlrausch *et al.*, 2000). The relatively high thresholds observed for the 2- and 4-Hz venelope frequencies suggest that the perceptually relevant decision variable is either integrated over a long interval, or that the internal representation of the signal AM is affected by preceding masker modulations for several hundred milliseconds. Anecdotally, listeners reported that the masker bursts were perceptually fused for venelope rates above 2 Hz and gradually became identifiable as temporally distinct and separable events at rates below 2 Hz.

C. Discussion

The findings from experiment II are consistent with those of experiment I in that venelope fluctuations did not appear to contribute to performance in envelope-frequency-domain versions of either task. In other words, there does not appear to be an additional independent coding dimension that the listeners have access to in the translated (modulation domain) experiments as there apparently is in the audio-frequency domain.

1. Relation to previous work

The results of our effort to measure CMR in the envelope-frequency domain are in qualitative agreement with the findings of several previous studies. Wojtczak and Viemeister (2005) also showed that a modulated envelope preceding a SAM signal imposed on the same carrier could affect detection performance for masker-probe delays of up

to 200 ms. Their AM forward-masking paradigm used a wideband noise carrier, a sinusoidal masker AM, and a short (50-ms) signal that was present only after the masker. The broadly similar time course of masking observed in the two studies suggests that a single mechanism could underlie both effects, and that it is independent of the statistical description of the carrier and masker. It is not obvious what this mechanism might be; in fact, if one assumes ringing or persistence as the underlying mechanism, neither effect would be expected based on the broadly tuned nature of the putative modulation filters, since the "trade-off" for implementing broad signal-processing filters is that the response recovers quickly from stimulation.

The information available to cochlear implant (CI) users is provided largely in the form of temporal envelope fluctuations imposed on the amplitude or duration of current pulses presented to stimulating electrodes. Nelson *et al.* (2003) and Nelson and Jin (2004) measured performance of CI users in a speech recognition task with a temporally modulated noise masker, and varied the "gate frequency" from 1 to 32 Hz. They found that CI users did not benefit from temporal gaps in the noise masker as long as 500 ms (performance was independent of the gate frequency). Conclusions from the CI experiments seem consistent with those from the current study: in conditions with severely impoverished spectral-frequency cues, listeners are unable to use relatively long temporal gaps in a noise masker to aid in the detection of a signal.

2. Interpreting time courses

The extended-time-course AM-detection experiment suggests a long integration time constant operating at some stage presumably central to the putative envelope-filtering process. Such a statement is consistent with "long-term" masked AM detection decision statistics that quantify responses based on an averaged representation of the processed stimulus envelope, such as envelope rms (e.g., Strickland and Viemeister, 1996; Ewert and Dau, 2000; Ewert *et al.*, 2002) or the average firing rate of a model IC cell (Nelson and Carney, 2006). However, it is worth pointing out that the current data set does not necessitate the assumption of such a time-averaged decision variable. It remains possible that a "local feature" decision variable (e.g., envelope max/min ratio or maximum local modulation depth) could be used, but that the listeners combine information from multiple looks (e.g., Sheft and Yost, 1990; Viemeister and Wakefield, 1991) of the details of the local features.

Cortical physiological studies have provided evidence for long-lasting modulation of responses to envelope fluctuations that might underlie the apparent perceptual sluggishness observed here. Using pure-tone forward masking paradigms in the primary auditory cortex (A1), several groups have shown that the response to a short probe signal could be affected by the presence of a masker that preceded the probe by as much as several hundred ms or longer (e.g., Calford and Semple, 1995; Brosch and Schreiner, 1997; Ulanovsky *et al.*, 2004). If the masker had a similar audio-frequency composition to that of the probe (as it did in the current study), the response to the probe was usually suppressed. In a recent study of the unanesthetized marmoset A1 that used stimuli more similar to those used in the current psychophysical experiments, Bartlett and Wang (2005) examined the contextual dependence of AM responses on previous stimulation. Their findings were in qualitative agreement with those of the AM forward masking studies, but the observed suppression (or facilitation) of a probe AM stimulus could last longer than 1 s in some neurons and depended on the modulation properties of the preceding stimulus. To date, the authors are not aware of physiological results at any level that address the effect of a nonsimultaneous masker modulation imposed on the same tone carrier as a deterministic signal modulation. In all of the studies mentioned above, the probe and masker were imposed at different times on separate carriers (i.e., the stimuli were gated). It would be interesting to know whether physiological time courses of adaptation to AM are different for gated and continuous carriers.

IV. MODELING

A. Methods

A physiologically motivated processing model developed to predict peripheral and central neural responses to pure SAM tones (Nelson and Carney, 2004) and psychophysical responses to masked SAM tones (Nelson and Carney, 2006) was used to simulate responses to the stimuli used in the current study. The peripheral model was a modification of previous AN models (Carney, 1993; Zhang *et al.*, 2001; Heinz *et al.*, 2001a), and the final model output can be compared to pure-tone onset response cells in the IC. Interactions between fast excitation and slow inhibition give rise to modulation tuning in the model IC cells. Since the two inputs are matched in audio-frequency CF, the model is referred to as the same-frequency inhibition and excitation (SFIE) model, as in Nelson and Carney (2004). In the current study, the relative strength of the inhibition with respect to the excitation of the model IC cells ($S_{INH,IC}$) was set equal to 1.0. This parameter was important for determining the threshold modulation depth required to elicit firing in the model cells: values of $S_{INH,IC} \leqslant 1$ resulted in lower depth thresholds than $S_{INH,IC}$ values greater than one (i.e., stronger inhibition re: excitation). The time constants of inhibition (τ_{inh}) and excitation (τ_{exc}) were chosen to yield a cell that was tuned to the signal f_m of interest (see Nelson and Carney, 2004).

B. Results and discussion

Simulation results are described and discussed with three specific psychophysical observations in mind: (1) audio-frequency level-discrimination thresholds depend on the choice of gating mode (experiment I), (2) AM detection thresholds depend on gating mode (experiment I extension) while AM depth-discrimination thresholds do not (experiment I), and (3) masked SAM detection thresholds do not readily improve when the masker is comodulated (experiment II). The first finding is examined most carefully with the model, and those results are used as justification for the assumptions made with the remaining sets of data. In general, the modeling work is meant to qualitatively test the ability of an existing, physiologically motivated envelope-

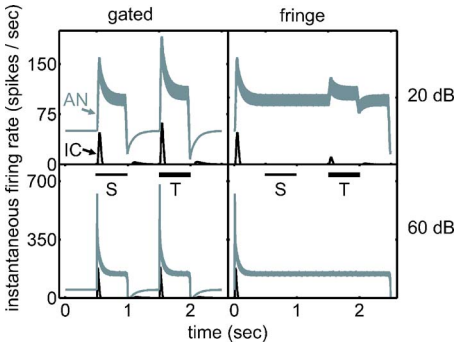

FIG. 6. Simulated responses to standard and target stimuli for the AN and IC levels of the SFIE model. Upper panels: 20 dB SPL standard level; lower panels: 60 dB SPL standard level. The target interval level was 3 dB higher than the standard. Left panels: gated stimuli; right panels: fringe presentation mode.

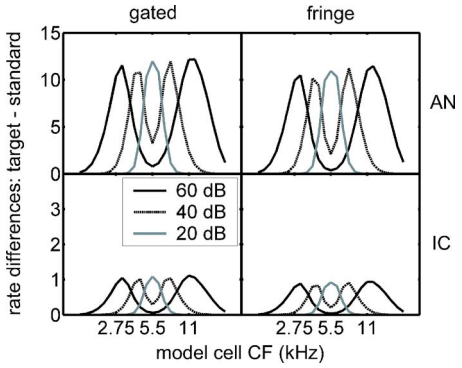

FIG. 7. Rate difference profiles for different gating modes (left and right panels) and levels of the model (upper and lower panels) in response to a 3-dB level increment. Each curve represents changes in the model rate responses for a 5500-Hz tone with a fixed standard SPL. Twenty-five model cells, log-spaced from 1375 to 22 000 Hz, were simulated for each standard level and gating mode.

processing model to account for broad and basic features of the data. This approach intentionally lies between speculating on potential mechanisms and explicitly predicting listeners' thresholds with a specific (i.e., fitted) model.

1. Audio-frequency domain level discrimination with gated and continuous carriers

By definition, modulation-tuned neurons are also envelope change detectors, and the properties that underlie AM responses in these neurons can also qualitatively explain the audio-frequency gated-continuous difference. To illustrate this point and to provide a concrete example of a component of a realistic neural circuit that predicts a heightened sensitivity to increments over changes in the level of gated tone bursts, the SFIE model was applied to the audio-frequency stimuli used in the current study.

Instantaneous firing rate (IFR) functions are shown in Fig. 6 for the responses of two stages of the model. The functions are comparable to physiological peri-stimulus time histograms (PSTHs) obtained for the AN (e.g., Harris and Dallos, 1979) and the IC (e.g., Langner and Schreiner, 1988) and were generated for eight illustrative conditions. Both AN and IC model responses are included in each of the panels of Fig. 6, which correspond to a specific combination of gating mode (gated or fringe) and standard SPL (20 or 60 dB). The timing of the standard (S) and target (T) stimulus presentation is marked by the two horizontal bars from 0.5–1 s and 1.5–2 s. A 3-dB level increment in the target interval was used. Other parameters of the stimuli were matched to those used in the psychophysical experiment. The IC model time constants ($\tau_{exc}=10$ ms; $\tau_{inh}=20$ ms) were chosen to yield a cell tuned to the effective 10-Hz modulation rate caused by the 50-ms onset and offset ramps.

First, consider the model outputs in response to a 20-dB standard-interval SPL tone (upper panels of Fig. 6). Here, the most critical differences between the modeled AN and IC responses are in the steady-state portion: ANFs respond with sustained firing to pure-tone stimulation, while the IC model only fires when the stimulus envelope elicits a change in the peripheral response. This is most clearly seen at envelope transitions with rising slopes, but offset adaptation in the peripheral model also results in a small response at the offset of the gated stimuli in the IC model. Note also that the IC model responds to both standard and target intervals when the intervals are gated (left panel), but only to the target interval in the fringe conditions (right panel). The AN model responses are always nonzero when a stimulus is present, although there is no change in the response from the baseline to the fringe-condition standard interval.

Because of rate saturation effects in the AN model, the pattern of model responses was quite different when a 60-dB standard tone SPL was used (Fig. 6, lower panels). Specifically, the target interval increment did not elicit a change in the fringe-condition responses of either the AN or IC model. As a result, a model consisting only of high-spontaneous rate (SR), on-CF ANFs at medium to high SPLs predicts unrealistically high level discrimination thresholds (Colburn *et al.*, 2003). There are two popular ways to account for psychophysical performance at high SPLs and high frequencies. One is to unevenly and heavily weight the contribution of low-SR ANFs (Winslow and Sachs, 1988; Delgutte, 1987; Viemeister, 1988). This approach is not completely satisfying, because such high-threshold, wide dynamic range ANFs make up only ~15% of the total population in cat (Liberman, 1978) and only exist at high CFs (>1500 Hz, Winter and Palmer, 1991). Another aspect of the response to high-level tones that may provide information for discrimination is in the spread of excitation across a population of neurons (Viemeister, 1972; Florentine *et al.*, 1987; Heinz *et al.*, 2001b; Colburn *et al.*, 2003). To address this issue, standard and target stimuli were presented to a group of model cells with different CFs.

Responses across the population were quantified in terms of their average rate over the entire 500 ms of the stimulus. Peripheral (AN) differences in the model's rate responses are shown in the upper panels of Fig. 7 and central (IC) rate differences are plotted in the lower panels. The parameter in each panel of Fig. 7 is the standard level. As in Fig. 6, gated and fringe conditions are illustrated in the left and right panels, respectively. For all four combinations of gating mode and model stage, the biggest differences in rate between the target and standard interval moved progressively away from the tone frequency (5500 Hz) as the standard SPL

was increased. This is consistent with previous studies using simulations of high-SR ANFs (e.g., Siebert, 1968; Heinz *et al.*, 2001b; Colburn *et al.*, 2003) and is caused entirely by saturation in the present model (because a linear basilar membrane model was used). A model version with level-dependent bandwidth and gain [i.e., a time-varying compressive nonlinearity (Heinz *et al.*, 2001a)] was also tested, and a similar pattern was obtained, suggesting that effects caused by saturation dominate those caused by compression when low-threshold, high-SR ANFs are used to estimate the population response.

Another feature of the differences in model rates was that the shapes of the profiles were similar for the responses of both stages of the model. This was a direct result of the simplified nature of the SFIE IC model neurons. Absolute values of the rate differences were significantly higher in the AN model (note the different scales for the upper and lower panels); this was not surprising given the sustained nature of responses to pure tones in the AN model and transient characteristics of the IC model responses. Finally, comparing across the gated and fringe conditions, the changes in model rates were not strongly dependent on the mode of gating for either the AN or IC model population.

The similarity of absolute rate differences for the gated and fringe conditions for both stages of the model does not provide a compelling explanation of the gated-continuous difference. It does, however, lead to a consideration of another feature of neural responses that must be known (or assumed) before predicting performance: the variability of rate estimates (e.g., Siebert, 1965). In actual ANFs, rate variability can be reasonably described as Poisson, with spike-count variance approximately equal to the mean count [at least at low rates, see Young and Barta (1986) and Winter and Palmer (1991) for a description of the reduced-variance deviation from Poisson at high rates]. The situation is less clear in more central processing stations, but, for simplicity, Poisson variance will be assumed in the responses of both stages of the model. These variance characterizations allow for a relatively simple formulation of the information provided by each frequency channel in the population rate response [following the approach of Siebert (1965); for details and derivations, also see Heinz *et al.* (2001b)]:

$$\delta'^2(icf) = \frac{[(\text{rate}_T - \text{rate}_S)/I_{\text{inc}}]^2}{\sigma^2_{\text{rate}}},$$

where $(\text{rate}_T - \text{rate}_S)$ is the rate difference term plotted in Fig. 7, I_{inc} is the size of the target-interval increment (a value of 1 results in the 3-dB amplitude increment in the examples shown), and σ^2_{rate} is the variance of the rate response. Our Poisson assumption allows for a simple estimate of σ^2_{rate}: it is assumed to be equal to the average rate (across both standard and target responses).

Information profiles, which incorporate both changes in rate and contributions of assumed neural variability, are shown in Fig. 8, in a format identical to that used to visualize the rate differences alone in Fig. 7. The limits of the ordinates are identical in all four panels. When the response variability is taken into account, the AN population model still predicts no advantage in the fringe condition relative to the gated presentation mode. In contrast, the envelope-change-detecting IC model clearly predicts a heightened sensitivity to the fringe condition, for all three tested standard levels. In addition, the overall summed population d' (related to the area under the information profile curve) is higher for the fringe-stimulus IC model rate responses than for the peripheral AN model responses. The gated-continuous difference in the IC model is strongly influenced by the lower average rate in the fringe condition, which translates into lower assumed variability and higher values of d'.

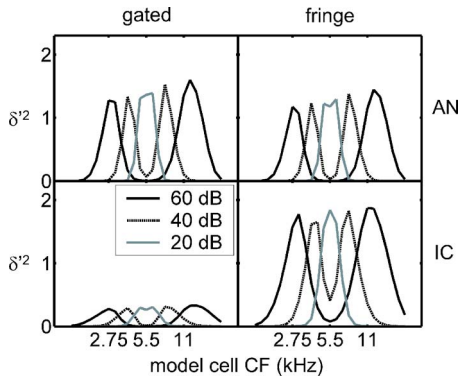

FIG. 8. Across-frequency information profiles, arranged in the same format as Fig. 7. This measure of sensitivity takes both neural variance and changes in average rates into account.

While the exact values of predicted d' and thresholds depend on details of the parameters of the model and the statistical description of the chosen internal noise assumption, overall trends and the difference between gated and fringe conditions for the SFIE model with equal amplitude inhibition and excitation do not. One of the key features of the IC model that underlies the current explanation of the audio-frequency gated-continuous difference is the fact that there is some response to both intervals in the gated condition, and only a response to the target interval in the fringe condition. The other critical assumption is an internal noise process that predicts response variability that increases with average rate. Such a change-detection mechanism could in theory be independent of peripheral adaptation, although there is an interaction between the two in the model, and some interplay probably exists in the real system as well.

In contrast to psychophysics, where at standard levels lower than about 20 dB above detection threshold, continuous and gated pedestals yield similar measures of performance (Carlyon and Moore, 1986; Viemeister and Bacon, 1988), the SFIE model predicts a fringe advantage for all supra-threshold standard SPLs. One way to potentially modify the model to account for this level dependence would be to add a second internal noise source with a fixed variance (in addition to the assumed Poisson noise) to the final model response.[1]

2. Envelope-frequency domain modulation detection and discrimination with gated and continuous carriers

AM detection thresholds (at least at low modulation rates) depend on whether the carrier is gated or quasi-continuous (Fig. 2 of the current study; see also Viemeister,

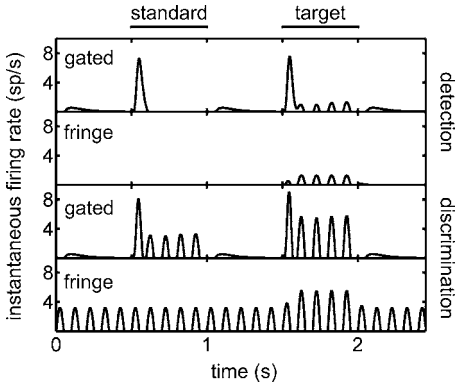

FIG. 9. SFIE model responses are qualitatively consistent with a fringe advantage in AM detection and no fringe advantage in AM depth discrimination. Model responses are shown for a 2.5-s window centered on the presentation of a standard followed by a target modulation. Simulated PSTHs are shown for an AM detection paradigm (top two panels) and an AM depth discrimination task (bottom two panels); gated and fringe conditions are included for both tasks. Stimulus parameters: f_c=5.5 kHz; SPL =70 dB; f_m=10 Hz; detection m=−20 dB; discrimination m_s=−12 dB, m_c =−7 dB. Key model parameters: τ_{exc}=10 ms; τ_{inh}=20 ms; $S_{INH,IC}$=1; AN CF=2000 Hz.

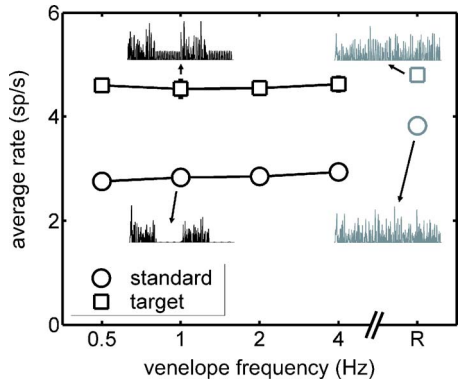

FIG. 10. Model IC cell average rates and example IFR functions in response to the stimuli used in the envelope-frequency CMR paradigm. Model parameters were the same as those used to generate the responses in Fig. 9, except the AN CF=2250 Hz, τ_{exc}=3 ms, and τ_{inh}=10 ms, which resulted in a cell rate-tuned to the 32-Hz signal AM frequency. The signal depth was fixed at −15 dB (20 log m) in the target interval, and a square-wave venelope was imposed on the masker in the comodulated conditions (as in the extension to experiment II). For comparison, rates and IFRs elicited by the random (R), or unmodulated, condition are also included in the plot. The duration of stimuli and IFRs was 2 s.

1979; Sheft and Yost, 1990; Yost and Sheft, 1997): thresholds are significantly higher when a gated carrier is used. In contrast, it was found that AM depth-discrimination thresholds (m_s=0.25) were statistically identical for gated and quasi-continuous carriers. Based on the results from the preceding section, analysis of the model in this section will be focused on IC model responses away from CF, where the biggest differences between standard and target were found.

IC model IFR functions are shown in Fig. 9 for 10-Hz SAM detection (top two panels) and 10-Hz SAM depth discrimination (bottom panels) paradigms. The observation intervals were 0.5 s; the standard interval started at 0.5 s and the target at 1.5 s. Labels in the upper left corner of each panel indicate the gating mode for each response. Simulated PSTHs for the AM-detection paradigm were similar to the IC model responses in Fig. 6 for audio-frequency level discrimination, in that the gated stimuli elicited a response in both the standard and target interval, while the fringe stimulus resulted in a model response only in the target interval. Again, if differences in both rate and variance are considered, model responses predict an enhanced sensitivity to the fringe condition compared to the gated condition (see preceding section).

In contrast, for AM depth discrimination (lower panels of Fig. 9) the IC model responded to both standard and target interval in the gated and fringe conditions. Since both rate differences and assumed rate variability are similar in these conditions, the IC model predicts little or no difference in thresholds between the gated and fringe presentation modes (as observed in the data of experiment I).

3. CMR experiment

Two questions were investigated concerning the ability of the model responses to qualitatively predict psychophysical trends observed with the stimuli used in the envelope-frequency-domain CMR experiment. First, does the model correctly predict the absence of a release from masking with the 4-Hz venelope fluctuation rate and 32-Hz signal SAM rate, as used in the base experiment? Second, is there an effect of venelope fluctuation rate over the range considered in the extension to the base CMR experiment?

To address the first question, a fixed-level signal SAM (−15 dB in 20 log m) was added to the masker in the target interval; model responses were quantified in terms of their average firing rates (across ten independent noise waveforms) in the random (R) unmodulated condition and in the 4-Hz square-wave comodulated masker condition for both standard and target intervals. If the model were consistent with the listeners' thresholds, the difference in the response to the standard and target intervals should be independent of the masker condition (random or comodulated at 4 Hz). Figure 10 shows that this was not the case; average rates in the random condition (gray symbols) were more similar in the standard (○) and target (□) intervals than the rates in the 4-Hz comodulated condition (right-most connected points). The disparity is caused by the reduced response magnitude in the standard intervals of the comodulated conditions; target-interval rates were largely independent of the venelope fluctuation patterns imposed on the masker.

The overall long-term envelope rms energy was identical in all of the noise-alone (standard) intervals shown in Fig. 10; the suppression in rate for the comodulated condition relative to that for the random condition was therefore caused by a nonlinear relationship between envelope rms and model IC cell average rate. The main factor contributing to this relationship was the change in the slope of the rate versus stimulus modulation depth function: at low m, the slope was shallower than at higher m. For example, when the modulation depth of a pure 32-Hz SAM signal was varied and presented to the cell simulated in Fig. 10, the slope of the function was ~0.1 sp/s/dB for $-30 \text{ dB} < 20 \log m < -25$ dB, and ~0.9 sp/s/dB for $-5 < 20 \log m < 0$ dB. This means that responses to small effective modulation depths (such as the "ripples" caused by postmodulation filtering of

the masker, or fluctuations in the masker away from the cell's best modulation frequency) were strongly attenuated in the model, which resulted in a reduced overall response to the modulated standard-interval stimulus.

Taken together, the rate responses of the model IC cell were not consistent with the listeners' inability to use the fluctuations in the masker to improve thresholds in the masked detection task. The schematic IFR functions included in Fig. 10 along with the rate quantifications show that the signal representation in the temporal responses of the model IC cells also suggest an increased salience of the signal in the comodulated conditions. The long effective time constants apparently necessary to explain the time course of release from masking observed in the extension of experiment II (on the order of hundreds of ms) are not included in model IC cells tuned to a 32-Hz signal frequency; as a result, the current model does not predict the psychophysical increase in thresholds with venelope fluctuation rate (connected symbols in Fig. 10.)

V. SUMMARY AND CONCLUSIONS

Two audio-frequency paradigms were translated into the envelope-frequency domain to assess the perceptual salience of venelope (second-order envelope) cues in continuous-carrier depth discrimination and SAM detection in the presence of a slowly varying noise masker. The experiments described here suggest a weak effect of temporal structure on performance in the two envelope-processing tasks. Several conclusions can be drawn from the empirical data and modeling results:

(i) Tone-carrier SAM-depth discrimination thresholds are not dependent on the gating mode of the carrier (i.e., gated or quasi-continuous), for a standard modulation depth of −12 dB (m_s=0.25) and modulation frequencies from 4 to 64 Hz. This contrasts with audio-frequency level discrimination results, which clearly indicate a heightened sensitivity to level increments when compared to gated-carrier level discrimination measurements.

(ii) SAM detection thresholds (or discrimination with a standard depth m_s=0) are approximately 5–6 dB lower when a quasi-continuous carrier is used than when the observation-interval stimuli are gated (for f_m=10 Hz).

(iii) Masked detection thresholds of a 32-Hz signal AM do not improve when the masker is slowly and coherently modulated with a 4-Hz venelope fluctuation rate. This is true for both sinusoidal and square-wave shaped comodulation. Audio-frequency tone detection thresholds, on the other hand, are strongly affected by the properties of the masker modulation.

(iv) To observe CMR in the envelope-frequency domain, the period of masker modulation must be lengthened until the masker bursts occur as perceptually distinct events (i.e., venelope fluctuation rates $\leqslant$1–2 Hz for a 32-Hz signal).

(v) A simple model developed to predict responses to SAM tones in the auditory midbrain can qualitatively account for several of the results, including the gated-continuous difference for pure-tone level discrimination and AM detection and the gated-continuous "similarity" for AM depth discrimination. The model does not, however, explain the listeners' inability to use relatively slow fluctuations in the instantaneous masker modulation depth to improve performance in the envelope-domain CMR task. Higher-order processing, possibly associated with auditory grouping and/or segregation mechanisms, may need to be considered to account for results in the CMR task.

ACKNOWLEDGMENTS

This research was supported by Grant Nos. NIH-NIDCD F31-7268 (PCN) and NIH-NIDCD R01-01641 (LHC, PCN) and by the Danish Research Council (TD, SDE). Discussions with Magdalena Wojtczak and Neal Viemeister were particularly helpful in preparing this paper.

[1]This adjustment would reduce the difference in rates at very low SPLs, when the standard level is near threshold. Because the effect of standard SPL was not a focus of the current behavioral experiments, such a modification was not included in the simulations. An addition of such a fixed-variance (stimulus-independent) internal noise would be conceptually similar to one of the sources of internal noise required to account for AM depth discrimination in Ewert and Dau (2004).

Bacon, S. P., and Viemeister, N. F. (**1994**). "Intensity discrimination and increment detection at 16 kHz," J. Acoust. Soc. Am. **95**, 2616–2621.

Bartlett, E. L., and Wang, X. (**2005**). "Long-lasting modulation by stimulus context in primate auditory cortex," J. Neurophysiol. **94**, 83–104.

Brosch, M., and Schreiner, C. E. (**1997**). "Time course of forward masking tuning curves in cat primary auditory cortex," J. Neurophysiol. **77**, 923–943.

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

Calford, M. B., and Semple, M. N. (**1995**). "Monaural inhibition in cat auditory cortex," J. Neurophysiol. **73**, 1876–1891.

Campbell, R. A., and Lasky, E. Z. (**1967**). "Masker level and sinusoidal-signal detection," J. Acoust. Soc. Am. **5**, 972–976.

Carlyon, R. P., and Moore, B. C. J. (**1986**). "Continuous versus gated pedestals and the 'severe departure' from Weber's Law," J. Acoust. Soc. Am. **79**, 461–464.

Carney, L. H. (**1993**). "A model for the responses of low-frequency auditory-nerve fibers in cat," J. Acoust. Soc. Am. **93**, 401–417.

Colburn, H. S., Carney, L. H., and Heinz, M. G. (**2003**). "Quantifying the information in auditory-nerve responses for level discrimination," J. Assoc. Res. Otolaryngol. **4**, 294–311.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892–2905.

Delgutte, B. (**1987**). "Peripheral auditory processing of speech information: implications from a physiological study of intensity discrimination," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Dordrecht, Nijhoff), pp. 333–353.

Derleth, R. P., and Dau, T. (**2000**). "On the role of envelope fluctuation processing in spectral masking," J. Acoust. Soc. Am. **108**, 285–296.

Durlach, N. I., and Braida, L. D. (**1969**). "Preliminary theory of intensity resolution," J. Acoust. Soc. Am. **46**, 372–383.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

Ewert, S. D., and Dau, T. (**2004**). "External and internal limitations in amplitude-modulation processing," J. Acoust. Soc. Am. **116**, 478–490.

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am. **112**, 2921–2931.

Florentine, M., Buus, S., and Mason, C. R. (**1987**). "Level discrimination as a function of level for tones from 0.25 to 16 kHz," J. Acoust. Soc. Am.

81, 1528–1541.
Füllgrabe, C., and Lorenzi, C. (**2005**). "Perception of the envelope-beat frequency of inharmonic complex temporal envelopes," J. Acoust. Soc. Am. **118**, 3757–3765.
Füllgrabe, C., Moore, B. C. J., Demany, L., Ewert, S. D., Sheft, S., and Lorenzi, C. (**2005**). "Modulation masking produced by second-order modulators," J. Acoust. Soc. Am. **117**, 2158–2168.
Gallun, F. J., and Hafter, E. R. (**2006**). "Amplitude modulation sensitivity as a mechanism for increment detection," J. Acoust. Soc. Am. **119**, 3919–3930.
Hafter, E. R., Bonnel, A. M., and Gallun, E. (**1997**). "A role for memory in divided attention between two independent stimuli," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London), pp. 228–237.
Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectro-temporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.
Harris, J. D. (**1963**). "Loudness discrimination," J. Speech Hear. Disord. Monogr. Suppl. **11**, 1–59.
Harris, D. M., and Dallos, P. (**1979**). "Forward masking of auditory nerve fiber responses," J. Neurophysiol. **42**, 1083–1107.
Heinz, M. G., Zhang, X., Bruce, I. C., and Carney, L. H. (**2001a**). "Auditory-nerve model for predicting performance limits of normal and impaired listeners," ARLO **2**, 91–96.
Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001b**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neural Comput. **13**, 2273–2316.
Houtgast, T. (**1989**). "Frequency selectivity in amplitude-modulation detection," J. Acoust. Soc. Am. **85**, 1676–1680.
Joris, P. X., and Yin, T. C. T. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," J. Acoust. Soc. Am. **91**, 215–232.
Joris, P. X., Schreiner, C. E., and Rees, A. (**2004**). "Neural processing of amplitude-modulated sounds," Physiol. Rev. **84**, 541–577.
Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723–734.
Langner, G., and Schreiner, C. E. (**1988**). "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms," J. Neurophysiol. **60**, 1799–1822.
Lee, J., and Bacon, S. P. (**1997**). "Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration," J. Acoust. Soc. Am. **101**, 3688–3693.
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.
Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442–455.
Lorenzi, C., Simpson, M. I. G., Millman, R. E., Griffiths, T. D., Woods, W. P., Rees, A., and Green, G. G. (**2001**). "Second-order modulation detection thresholds for pure-tone and narrow-band noise carriers," J. Acoust. Soc. Am. **110**, 2470–2478.
Lüscher, E., and Zwislocki, J. (**1949**). "Adaptation of the ear to sound stimuli," J. Acoust. Soc. Am. **21**, 135–139.
Macmillan, N. A. (**1971**). "Detection and recognition of increments and decrements in auditory intensity," Percept. Psychophys. **10**, 233–238.
Moore, B. C. J., Alcantara, J. I., and Dau, T. (**1998**). "Masking patterns for sinusoidal and narrowband noise maskers," J. Acoust. Soc. Am. **104**, 1023–1038.
Moore, B. C. J., Sek, A., and Glasberg, B. R. (**1999**). "Modulation masking produced by beating modulators," J. Acoust. Soc. Am. **106**, 938–945.
Nelson, P. B., and Jin, S. H. (**2004**). "Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **115**, 2286–2294.
Nelson, P. B., Jin, S. H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.
Nelson, P. C., and Carney, L. H. (**2004**). "A phenomenological model of peripheral and central neural responses to amplitude-modulated tones," J. Acoust. Soc. Am. **116**, 2173–2186.
Nelson, P. C., and Carney, L. H. (**2006**). "Cues for masked amplitude-modulation detection," J. Acoust. Soc. Am. **120**, 978–990.
Nelson, P. C., and Carney, L. H., (**2007**). "Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus," J. Neurophysiol. **97**, 522-539.
Schooneveldt, G. P., and Moore, B. C. J. (**1989**). "Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration," J. Acoust. Soc. Am. **85**, 273–281.
Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude-modulation detection," J. Acoust. Soc. Am. **88**, 796–805.
Shofner, W. P., Sheft, S., and Guzman, S. J. (**1996**). "Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency, two-tone complexes," J. Acoust. Soc. Am. **99**, 3592–3605.
Siebert, W. M. (**1965**). "Some implications of the stochastic behavior of primary auditory neurons," Kybernetik **2**, 206–215.
Siebert, W. M. (**1968**). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. A. Kolers (MIT, Cambridge, MA), pp. 104–133.
Smith, R. L., and Brachman, M. L. (**1982**). "Adaptation in auditory-nerve fibers: A revised model," Biol. Cybern. **44**, 107–120.
Smith, R. L., and Zwislocki, J. J. (**1975**). "Short-term adaptation and incremental responses in single auditory-nerve fibers," Biol. Cybern. **17**, 169–182.
Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). "Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments," J. Acoust. Soc. Am. **78**, 1310–1316.
Strickland, E. A., and Viemeister, N. F. (**1996**). "Cues for discrimination of envelopes," J. Acoust. Soc. Am. **99**, 3638–3646.
Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (**2004**). "Multiple time scales of adaptation in auditory cortex neurons," J. Neurosci. **24**, 10440–10453.
Verhey, J. L., Dau, T., and Kollmeier, B. (**1999**). "Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model," J. Acoust. Soc. Am. **106**, 2733–2745.
Verhey, J. L., Pressnitzer, D., and Winter, I. M. (**2003**). "The psychophysics and physiology of comodulation masking release," Exp. Brain Res. **153**, 405–417.
Viemeister, N. F. (**1972**). "Intensity discrimination of pulsed sinusoids: The effects of filtered noise," J. Acoust. Soc. Am. **51**, 1265–1269.
Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.
Viemeister, N. F. (**1988**). "Intensity coding and the dynamic range problem," Hear. Res. **34**, 267–274.
Viemeister, N. F., and Bacon, S. P. (**1988**). "Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones," J. Acoust. Soc. Am. **84**, 172–178.
Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865.
Wakefield, G. H., and Viemeister, N. F. (**1990**). "Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise," J. Acoust. Soc. Am. **88**, 1367–1373.
Wegel, R. L., and Lane, C. E. (**1924**). "The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear," Phys. Rev. **23**, 266–285.
Winslow, R. L., and Sachs, M. B. (**1988**). "Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle," Hear. Res. **35**, 165–189.
Winter, I. M., and Palmer, A. R. (**1991**). "Intensity coding in low-frequency auditory-nerve fibers of the guinea pig," J. Acoust. Soc. Am. **90**, 1958–1967.
Wojtczak, M., and Viemeister, N. F. (**1999**). "Intensity discrimination and detection of amplitude modulation," J. Acoust. Soc. Am. **106**, 1917–1924.
Wojtczak, M., and Viemeister, N. F. (**2005**). "Forward masking of amplitude modulation: Basic characteristics," J. Acoust. Soc. Am. **118**, 3198–3210.
Yost, W. A., and Sheft, S. (**1997**). "Temporal modulation transfer functions for tonal stimuli: Gated versus continuous conditions," Aud. Neurosci. **3**, 401–414.
Young, E. D., and Barta, P. E. (**1986**). "Rate responses of auditory nerve fibers to tones in noise near masked threshold," J. Acoust. Soc. Am. **79**, 426–442.
Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," J. Acoust. Soc. Am. **109**, 648–670.