Investigating temporal asymmetry using masking period patterns and models of peripheral auditory processing

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

(Received 17 March 2010; revised 7 March 2011; accepted 10 March 2011)

Two experiments were conducted in conjunction with modeling to evaluate the role of peripheral nonlinearity and neural adaptation in the perception of temporally asymmetric sounds. In both experiments, maskers were broadband noises amplitude modulated with ramped and damped exponential modulators that repeated at 40 Hz. Masking period patterns (MPPs) were constructed by measuring detection threshold of a 5-ms, 1000-Hz tone burst as function of the signal's onset delay. Experiment I showed that varying modulator half-life from 1 to 16 ms led to differences in the damped and the ramped MPPs that were largest at the short half-lives and diminished at the longer half-lives. When masker level was varied (experiment II), the largest difference between ramped and damped MPPs occurred at moderate stimulus levels. Two peripheral auditory models were evaluated, one a simple auditory filter followed by a power-law nonlinearity and another, a model of auditory nerve processing [J. Acoust. Soc. Am. **126**, 2390–2412 (2009)] that includes neural adaptation. Neither models predicted differences between the ramped and damped MPPs, providing indirect support that the central auditory system has a role in perceptual temporal asymmetry.
 [DOI: 10.1121/1.3573979]

PACS number(s): 43.66.Dc, 43.66.Mk [MAA] Pages: 3194–3205

I. INTRODUCTION

There is a widespread presence of asymmetrically modulated sounds in the natural listening environment that are relevant to the perception of pitch, rhythm, timbre, and speech. Despite the pervasiveness of such stimuli, it has been more common to study the auditory representation of symmetrically modulated sounds, such as sinusoidal- and narrowband-noise forms of modulation (e.g., Viemeister, 1979; Fastl, 1982; Strickland, 2000). Numerous features of auditory temporal processing are also asymmetric (e.g., the impulse response of the auditory filter and the temporal integration window), and natural environments (such as room reverberation) can introduce asymmetries to sounds as well. Consequently, using temporally asymmetric stimuli in psychoacoustic experiments is likely to further our understanding of the auditory mechanisms responsible in the perception of all modulated sounds. Over the past two decades, a limited number of studies have measured the perceptual asymmetry of asymmetrically modulated sounds (e.g., Patterson, 1994a,b; Akeroyd and Patterson, 1995, 1997; Carlyon, 1996; Irino and Patterson, 1996). The following experiments expand upon these studies by evaluating the perception of asymmetrically modulated sounds using a temporal masking paradigm, similar to that adopted by Ries *et al*. (2008) for single pulses of ramped and damped sounds. A model of the auditory periphery is applied in an attempt to elucidate the auditory mechanisms responsible for the perception of temporal asymmetry.

Patterson (1994a,b) first demonstrated that modulated tones with exponentially rising and falling modulation patterns could be readily discriminated from each other, even though the stimuli had identical long-term power and envelope spectra. Further, the ramped tones (rising modulators) were often perceived as having a stronger tonal quality than damped tones (falling modulators), which were dominated by a "drum-roll" quality. Akeroyd and Patterson (1995, 1997) extended this finding by using noise carriers and showed that under certain conditions, ramped and damped noises were also easily discriminable, with ramped noises having a more "noisy" quality than damped noises which sounded more like "drumming." In addition, non-repeating ramped and damped sounds (here, referred to as rising and falling sounds to distinguish them from periodic ramped sounds) have been shown to elicit distinct perceptual differences with rising sounds often heard as being louder (e.g., Neuhoff, 1998; Stecker and Hafter, 2000; Ries *et al*., 2008) and longer (e.g., Schlauch *et al*., 2001; Grassi and Darwin, 2006; DiGiovanni and Schlauch, 2007) than falling sounds.

The perception of asymmetrically modulated sounds is a process which likely involves both central and peripheral auditory effects. Peripheral psychophysical models yield a small difference in the shape of the internal envelope representations of ramped and damped noises (see Patterson, 1994a), but these models have had limited success in modeling the magnitude of the perceptual temporal asymmetry (see Akeroyd and Patterson, 1995; Carlyon, 1996; Irino and Patterson, 1996). Consequently, the role of the auditory periphery in temporal asymmetry has not been firmly established. Neurophysiological measurements also have been conducted in attempts to further understand the processes underlying perceptual asymmetry. Pressnitzer *et al*. (2000) measured responses of three different cell types in the cochlear nucleus (primary-like, onset, and chopper) to ramped and damped tones and showed that the responses of the

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

 0001-4966/2011/129(5)/3194/12/$30.00

primary-like cells exhibited smaller asymmetry than the onset and chopper cells. Asymmetric neural responses to ramped and damped sinusoids also have been measured in the inferior colliculus (Neuert *et al*., 2001) and auditory cortex (Lu *et al*., 2001). Although it is difficult to make strong comparisons due to differences in the techniques used across the various studies, the asymmetry in the stimulus representation may be enhanced by each subsequent higher auditory center. In lower nuclei (such as the cochlear nucleus), the asymmetry is represented as differences in the envelope of a post-stimulus time histogram and the firing rate, whereas the auditory cortex contains a significant portion of neurons that respond exclusively to either ramped or damped stimuli. Because of the possible enhancement by higher auditory centers, the central auditory system may have an important role in forming the perceptual difference between ramped and damped stimuli.

Nevertheless, one cannot rule out a major contribution of the auditory periphery on the perception of temporal asymmetry. To that end, we measured masking period patterns (MPPs) of ramped and damped noises. Masking was used here instead of discrimination due to confound of the multiple available cues in discrimination tasks. To assess the contribution of the auditory periphery on the measured effects, two models of peripheral processing were implemented to simulate the experimental results. Model A consisted of an auditory filter, a half-wave rectifier, and a nonlinear compressor. Despite its simplicity, this model has previously shown promising performance in encapsulating the signal detection threshold in the presence of temporally fluctuating maskers (e.g., Alcantara *et al*, 1996; Wojtczak *et al*., 2001). Model B was a modern physiological model of the auditory periphery developed by Zilany *et al*. (2009). This model demonstrates excellent performance in capturing auditory-nerve responses to dynamic stimuli and includes a realistic adaptation characteristic at the auditory nerve level. Irino and Patterson (1996) demonstrated that a model which includes adaptation was capable of producing a temporal-asymmetry effect with appropriate parameters. Applying the model to this data set can help evaluate whether auditory-nerve adaptation could be a source of any observed temporal asymmetry.

MPPs were measured to provide an estimate of the internal envelope shape of ramped and damped noises. A MPP plots the detection threshold of a tone burst measured as a function of signal delays which sample the period of a periodic stimulus (cf. Zwicker, 1976a; Fastl, 1977, 1982). For temporally modulated maskers, thresholds are typically lower when the tone burst is added to a temporal valley and higher at the temporal peaks. Consequently, these patterns are often interpreted as reflecting the internal representation of the masker envelope (Zwicker, 1976a). MPPs have been measured for numerous types of stimuli, including pure tones (Zwicker, 1976b), harmonic complexes (Zwicker, 1976a; Kohlrausch and Sander, 1995), and noise bursts (Fastl, 1977, 1982). To our knowledge, no MPPs exist for asymmetrically modulated noises. However, Ries *et al*. (2008) did measure temporal masking patterns of rising and falling noises and showed that the change in threshold from masker maximum to minimum was less for rising than for falling sounds. Yet, it is unclear whether the results of this study generalize to periodic ramped and damped sounds, as aspects of auditory adaptation and efferent effects which are mediated over longer time courses may have different roles to play for the two types of stimuli. In contrast to rising and falling sounds, modulated stimuli consist of repeated onsets over time and may "reset" the efferent or adaptation effects. Generally, our expectation was that MPPs of ramped stimuli would be less modulated than the MPPs of damped stimuli as a consequence of peripheral filtering, neural adaptation, or other auditory processes.

II. GENERAL METHODS

In these experiments, the damped stimulus was an amplitude modulated noise, with a modulator containing a periodically repeating pattern of exponential decay. A single period of the damped modulator took the form

$$\mathrm{damp}(t) = e^{-ct/h}, \tag{1}$$

where h is the half-life of the modulator in seconds, c has a value of $\ln 2$, and t is the time in seconds. The ramped modulator was the time reverse of the damped modulator; therefore, the ramped and damped noises with the same half-life had identical long-term power spectra.

A. Stimuli

MPPs were measured for ramped and damped noises as maskers at different rates of exponential decay (half-life; experiment I) and different stimulus levels (experiment II). The maskers were broadband, Gaussian noises multiplied by a repeating exponential modulator having a 25-ms period. Even though most of the work on ramped and damped sounds has been done with tonal carriers, noise carriers were used here in order to eliminate confounds related to tone-on-tone masking (e.g. off-frequency listening and combination tones).

Each period of a damped modulator was given by Eq. 1, and the ramped modulator was the time-reversed damped modulator. Sample waveforms of the ramped and damped maskers generated at a half-life of 4 ms are illustrated in Fig. 1(a). Maskers were 250 ms in duration including 30-ms cosine-squared onset and offset ramps. On each stimulus presentation, an independently chosen sample of noise was presented.

The signal to be detected was a 5-ms, 1000-Hz sinusoid with no steady state (2.5-ms raised-cosine ramped at onset and offset), generated with random initial phase for each presentation. The tone burst was calibrated such that its dB sound pressure level (SPL) was equivalent to a long-duration tone with the same peak amplitude. Signal detection thresholds were measured at signal onset delays of 125, 130, 135, 140, and 145 ms relative to the onset of the masker. The signal delays were defined by the onset of the signal, not from the center. In this way, the masker's 25-ms period was fully sampled in 5-ms steps. For damped maskers, these five

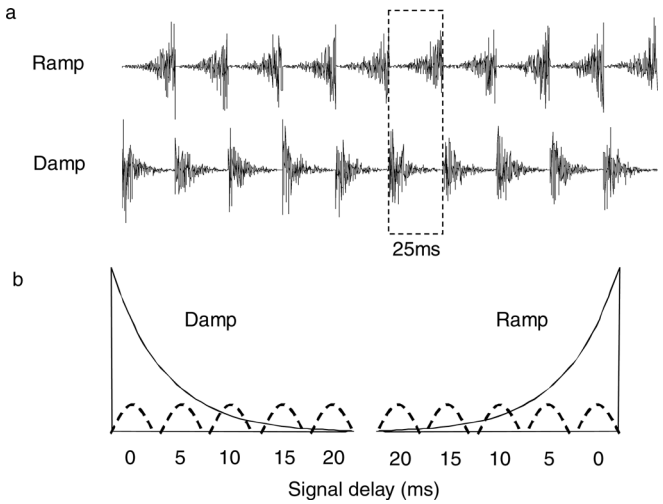

FIG. 1. Schematic examples of the stimuli. Panel (a) shows sample waveforms of ramped and damped maskers with a 4-ms half-life. The sixth period (25-ms in duration) within which the tone-pip signals are presented is indicated by the dashed box. Panel (b) shows the signal locations within the sixth masker period. Note that signal delays that will be used for data presentation are time-reversed for the ramped conditions. In this way, the same signal de lay used for ramped and damped stimuli corresponds to equal power in the maskers.

delays were labeled relative to the onset of the sixth masker period as 0, 5, 10, 15, and 20 ms.

One way to visualize the difference between ramped and damped MPPs is to reverse the signal delays for the ramped conditions thereby enabling direct comparisons of the shape of the MPPs. Thus, we can see if the MPPs mirror each other or whether there are asymmetric changes imposed by the auditory system. Therefore, the five signal delays in the ramped condition are labeled relative to the offset of the sixth masker period as 20, 15, 10, 5, and 0 ms [see Fig. 1(b)]. At the same labeled signal delay, the power of the two types of the maskers within the 5-ms time frame of signal presentation was equal.

Stimuli were generated digitally and played through a 24-bit DAC (Tucker Davis Technologies, Alachua, FL; TDT RP2.1) at a sampling rate of 48 824 Hz. The output of the DAC was fed into a single attenuator (TDT PA5) and then into a headphone buffer (TDT HB7). The resulting stimulus was presented to the listener through one earphone of a Sennheiser HD250 linear II headset (Old Lyme, CT). All data were collected with listeners seated in a double-walled sound-attenuating room.

B. Experimental procedure

Signal detection thresholds were estimated using a two-alternative forced-choice three-down, one-up adaptive procedure, targeting the 79.4% correct point on the psychometric function (Levitt, 1971). The masker was present in both intervals, with the signal presented randomly with equal probability in one of the two intervals on each trial. Participants indicated the signal interval using a button box and received correct-answer feedback after each trial. Threshold tracking began at a signal level approximately 15–20 dB above the final estimated threshold for each condition. The initial step size used in the tracking procedure was 5 dB. Following three reversals, the step size was decreased to 2.5 dB. Tracking ended after a total of nine reversals, and threshold was estimated as the average of the final six reversal points. Masked thresholds are reported in dB SPL.

C. Modeling of MPPs

Two models of auditory peripheral processing were incorporated in the current study to assist interpreting the experimental findings. Both models were implemented with the following structure: the peripheral model (model A or model B), a leaky temporal integrator, and a decision device. The models only differed in the peripheral model and the amount of internal noise used in the decision device.

Model A consisted of an auditory filter, a half-wave rectifier and a power-law nonlinearity (see Oxenham and Moore, 1994; Alcántara *et al*., 1996; Wojtczak *et al*., 2001). The auditory filter was a single gammatone filter centered at 1 kHz with a 3-dB bandwidth of 118 Hz. The power-law nonlinearity following the half-wave rectifier raised the rectified waveform to the power of 0.5 as Oxenham and Moore (1994) showed this exponent to be appropriate for modeling forward-masking data.

Model B consisted of a middle-ear filter, a nonlinear cochlear filtering stage, and a model of the synapse between the inner hair cell and the auditory nerve fiber [see Zilany *et al*. (2009) for details]. This model includes exponential adaptation with a short time constant to respond to onsets and a slowly adapting power-law component to predict recovery of the auditory nerve response at stimulus offset. The asymmetric response and its performance in capturing the auditory-nerve response to dynamic stimuli for physiological stimuli make this model an excellent choice for modeling perceptual asymmetry. As noted in Sec. I, the adaptation characteristic could be important for producing a temporal asymmetry effect (Irino and Patterson, 1996; Ries *et al*., 2008). For the current implementation of the model, normal inner- and outer-hair cell functionality was assumed. The best frequency of the modeled auditory nerve fiber was 1 kHz, and the fiber was set to have medium spontaneous firing rate. For each presentation of the stimuli, the model generates a synaptic output from the inner hair cell to the auditory nerve.[1]

A leaky integrator was added as the final stage of both models. The leaky integrator was included because the Zilany *et al*. (2009) model can follow the stimulus envelope up to about 1 kHz (Zilany *et al*., 2009, Fig. 14) whereas psychophysical estimates are generally closer to 20–60 Hz.[2] The leaky integrator was implemented as a low-pass filter having a cutoff frequency of 30 Hz. This filter resembled a temporal integrator with an exponential impulse response having an equivalent rectangular duration (ERD) of 5.3 ms.[3,4] This cutoff frequency was selected as Moore *et al*. (1988) and Plack and Moore (1990) estimated the length of the leaky integrator to be about 8 ms (i.e., a 20-Hz cutoff frequency), and Alcántara *et al*. (1996) estimated about 4 ms (i.e., a 40-Hz cutoff frequency).

The models were implemented following this procedure: For each model, two unique masker stimuli were

generated on each trial of the model. The 5-ms signal was added to one of the maskers during the threshold estimation. Then, signal + masker and masker stimuli were passed through the models. The detection threshold of the signal was estimated using a procedure similar to that of Mott *et al.* (1990) and Pressnitzer *et al.* (2001). The model outputs for signal + masker and masker were divided into 5-ms bins, and the mean and variance of the responses within each bin were calculated.[5] The bin that shared its onset with the signal was called the signal bin, and the decision variable D was estimated based on the signal bin and four bins following the signal bin covering a full masker period. D was calculated as:

$$D = \sqrt{\sum_{i=0}^{4} \frac{[\mu_i(\mathrm{s}+\mathrm{m}) - \mu_i(\mathrm{m})]^2}{\sigma_i^2(\mathrm{s}+\mathrm{m}) + \sigma_i^2(\mathrm{m}) + \sigma_0^2}}, \quad (2)$$

where $\mu_i(\mathrm{s}+\mathrm{m})$ and of $\sigma_i^2(\mathrm{s}+\mathrm{m})$ were the response means and variances to the signal + masker stimulus, respectively, with $i = 0,1,\ldots,4$ being the number of bins from the signal bin. $\mu_i(\mathrm{m})$ and $\sigma_i^2(\mathrm{m})$ were the response means and variances to the masker alone, and σ_0 was the standard deviation of a internal decision noise, which limited model performance when no masker was presented.[6] The signal was said to be detected whenever D was larger or equal to the criterion K.

To enable a more time-efficient simulation, a Yes/No procedure combined with one-down, one-up tracking was implemented for threshold estimation. Each track contained a total of seven reversals. The signal levels at last five reversals were averaged to form a threshold estimate. For each experimental condition, this process was repeated four times, and the reported results were based on the average of these repetitions.

In order to match the model predictions to the experimental data, two free parameters (σ_0 and K) were manipulated with one pair of σ_0 and K values assigned to each of the models. These values were varied under the following constraints: (1) the predicted threshold for a signal delay of 0 ms and a half-life of 2 ms at the 70-dB SPL masker level in the damped condition of experiment I was matched to the measured thresholds in the corresponding condition, and (2) the tone-pip threshold was between 30 and 35 dB SPL when no masker was presented. There are various combinations of σ_0 and K that meet these criteria, and so the values of these parameters were determined by good visual agreement between the model predictions of the shape of the MPPs.

III. EXPERIMENT I: EFFECT OF HALF-LIFE

Experiment I measured MPPs for ramped and damped noises at different modulator half-lives. It was expected that the behavioral data would confirm the modeling results of Patterson (1994a) by indicating less modulated internal envelopes for ramped than damped noises. Moreover, as the modulator half-life increases, modulation depth effectively decreases. Therefore, we also anticipated a reduction in the difference between ramped and damped MPPs with increasing half-life.

A. Experimental methods

Four observers, ranging in age from 21 to 24 years, participated. All observers had pure tone audiometric thresholds of 15 dB HL or better at audiometric frequencies between 250 Hz and 4 000 Hz (ANSI, 2004). The right ear was tested for two of the four listeners.

For this experiment, the maskers were presented at an rms level of 70 dB SPL, corresponding to a spectrum level of approximately 30 dB SPL for all half-lives. Data collection followed a randomized block design in which the half-life of the stimulus (1, 2, 4, 8, or 16 ms) was chosen at random. For each half-life, the five signal delays were tested in random order. Thresholds were calculated for each signal delay before the next signal delay was tested. In this way, a single MPP was collected for each half-life prior to testing the next randomly selected half-life. This process was repeated five times. The first threshold collected in each condition was treated as practice, and reported thresholds reflect the average of the last four thresholds obtained.

B. Experimental results

Figure 2 plots the average MPPs obtained at each half-life for both ramped and damped waveforms in the leftmost panels with data obtained from each individual listener in the right panels. As mentioned in Sec. II, for ease in comparing the shape of the ramped and damped MPPs, the MPPs of the ramped noise are shown time reversed [see Fig. 1(b)].

In general, Fig. 2 illustrates that signal detection threshold was dependent on signal delay and that the dependence on signal delay varied with half-life. MPPs for ramped and damped noises obtained at the 1- and 2-ms half-lives (upper rows) were highly modulated and exhibited a large change in threshold (over 20 dB) across the different signal delays. The amount of modulation present in the MPP (the MPP modulation depth) then decreased with increasing half-life. At half-lives of 8 and 16 ms, the change in threshold across the MPP was less than 7 dB (6.2 dB for 8 ms and 2.5 dB for 16 ms). It is notable that at 16 ms, the damped MPP was associated with slightly higher thresholds than the ramped MPP, a different direction than observed at the shorter half-lives.

In addition to a decrease in the MPP modulation depth with increasing half-life, the ramped and damped MPPs had different shapes. The ramped MPP had a shallower slope, and it also did not reach as low of a minimum as the damped MPP. In this way, the damped MPP was more modulated than the ramped MPP, with greater masking being produced by the ramped masker than the damped masker, though this effect was dependent on half-life. The difference between the ramped and damped MPPs decreased with increasing half-life resulting in only small differences present for the 8- and 16-ms half-lives.

To verify that the observations mentioned above have statistical support, a repeated measures analysis of variance (ANOVA) conducted on the data treating masker type, half-life, and signal delay as within-subjects factors was conducted to establish the main trends in the data. Note that the values for the thresholds obtained using the ramped masker

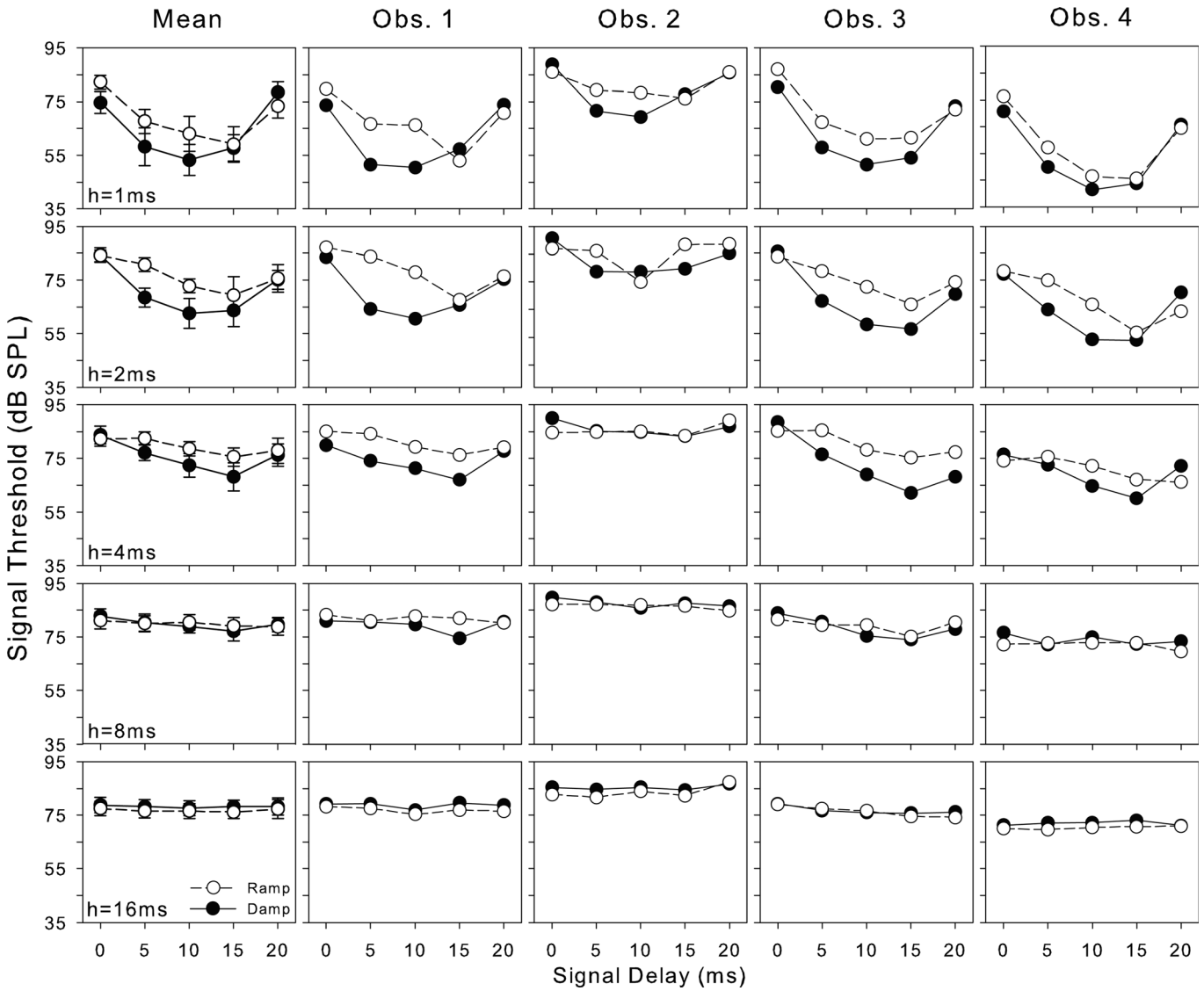

FIG. 2. Masked detection thresholds of a 1-kHz tone burst are plotted as a function of signal delay for ramped (unfilled circles) and damped (filled circles) noises. Mean MPPs are plotted in the leftmost panels with data obtained from individual listeners in the right panels. MPPs obtained at different half-lives are shown from top to bottom. Thresholds from the ramped-noise conditions are reversed in time to enable better visual comparison between the ramped and the damped masking period patterns. Error bars (left panel only) represent standard error of the mean across listeners.

were also time-reversed in this ANOVA. As reported in Table I, the ANOVA revealed significant main effects of half-life and signal delay but not masker type (ramped versus damped). All three two-way interactions and the three-way interaction were significant.

First, the involvement of masker type in all interactions was suggestive of differential patterns of masking by ramped and damped noises even though there was no main effect of masker type. To determine at which half-lives ramped and damped maskers yielded differences in masking, *post hoc* paired-sample t tests using the Bonferonni correction to the critical p for multiple comparisons ($p < 0.01$) were conducted. Significant differences between ramped and damped MPPs at half-lives of 1 [$t(19) = -3.5$; $p < 0.002$], 2 [$t(19) = -3.5$; $p < 0.002$], 4 [$t(19) = -3.1$; $p < 0.006$], and 16 ms [$t(19) = 5.4$; $p < 0.001$] were revealed indicating that the damped stimulus produced less masking than the ramped stimulus for the half-lives of 1, 2, and 4 ms. The reversal in the 16-ms data, where the damped masker produced slightly more masking than the ramped masker was also shown to be significant. The source of this reversal is unclear.

Second, MPP depth appeared to vary with half-life. To determine which half-lives were associated with a significant

TABLE I. Results of ANOVA for experiment I.

Factor	F	Significance
Half-life	$F(4,12) = 27.1$	$p < 0.001$[a]
Masker type	$F(1,3) = 5.9$	$p = 0.09$
Delay	$F(4,12) = 20.4$	$p < 0.001$[a]
Half-life × masker type	$F(4,12) = 16.8$	$p < 0.001$[a]
Half-life × delay	$F(16,48) = 17.1$	$p < 0.001$[a]
Masker type × delay	$F(4,12) = 7.2$	$p < 0.003$[a]
Half-life × masker type × delay	$F(16,48) = 3.2$	$p < 0.001$[a]

[a]Significant at $p < 0.05$.

"valley," *post hoc* paired-sample t tests were conducted on the interaction between half-life and delay. In this case, t tests were conducted within a half-life in which the threshold obtained at the delay of 0 was compared to the other delays. In this case, 20 comparisons were made, and so the critical p for multiple comparisons was $p < 0.0025$. Thresholds obtained at the 0-ms signal delay were significantly different from thresholds obtained at the delays of 5, 10, and 15 ms [$t(7) = 9.3$; $p < 0.001$, $t(7) = 7.7$; $p < 0.001$, and $t(7) = 7.7$; $p < 0.001$, respectively] for the 1-ms half-life and the 10- and 15-ms delays at the 2-ms half-life [$t(7) = 6.5$; $p < 0.001$ and $t(7) = 5.3$; $p < 0.001$, respectively]. In sum, there is evidence to suggest that the MPPs at the 1- and 2-ms half-lives have significant degrees of modulation but the MPPs at 4, 8, and 16 ms may not.

The third observation was that the MPP of the ramped stimulus was more modulated than the damped stimulus. Using the t tests and Bonferonni correction for five comparisons ($p < 0.01$), masking produced by the damped stimulus was significantly different from the masking produced by the ramped stimulus at delays of 5 and 10 ms [$t(19) = -3.5$; $p < 0.002$ and $t(19) = -3.6$; $p < 0.002$], respectively. These results suggest that the difference in the shape of the ramped and damped MPP is mainly due to the differences in masking at two delays which are near, but not at, the peak of the masker.

Although the mean does a reasonable job at capturing the trends observed across listeners, the individual data are also shown in Fig. 2 to illustrate the large differences in masked threshold measured across listeners. Large individual differences have been reported previously in masking experiments with modulated maskers (cf. Oxenham and Dau, 2001). All listeners generally showed similar trends of highly modulated MPPs at short half-lives, a deeper MPP for damped over ramped noises, and a decrease in the depth of MPP modulation with increasing half-life. However, Obs. 2 had higher thresholds compared to the other listeners for all half-lives tested. Further, this listener's data show less MPP modulation than the other listeners and a smaller difference between the ramped and damped MPPs. In particular, the time-reversed ramped and damped MPPs tend to be similar for this listener at half-lives of 4 ms and greater, whereas this overlap does not occur until half-lives of 8 ms and greater for the other listeners. Although the factors that contribute to the high thresholds exhibited by this listener are not obvious, the higher masked thresholds may have not allowed this listener to receive a release from masking afforded by the temporal structure of the stimuli. It is worth noting that we were able to measure absolute thresholds for the 1-kHz tone burst for three of the four listeners, and this listener's thresholds were on the same order as the other two listeners (about 30 dB SPL).

In summary, a large temporal asymmetry was observed in the present experiment, though the degree of asymmetry was dependent on modulator half-life. These results are in agreement with Ries *et al.* (2008) who showed that rising noises were more efficient maskers (i.e., they produce more masking) and had a shallower internal envelope shape than damped noises.

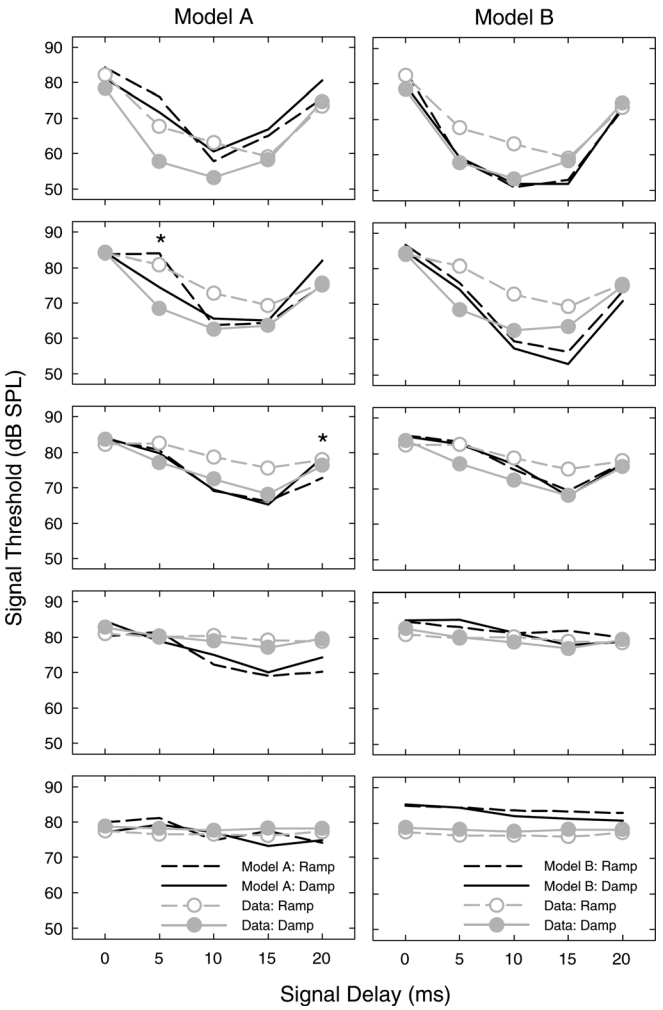

FIG. 3. Model A and model B predictions are plotted in the left and right panels, along with the mean data of experiment I. Predictions for ramped and damped stimuli are shown as dashed and solid lines, respectively.

C. Model results

Predictions from models A and B are illustrated in the left and right panels of Fig. 3, respectively, together with the average data from experiment I. In general terms, both models captured a number of the main trends in the data. First, salient peaks and valleys can be observed in the predicted MPPs (i.e., the MPPs contain modulation on the order of that observed in the data). Second, the shapes of the predicted MPPs, with the highest thresholds at 0- and 20-ms signal delays with valleys at 10- and 15-ms signal delays, were also similar to those observed in the data. Third, the depth of the MPPs tended to decrease with increasing half-life. In contrast to the experimental data, however, the predicted ramped and damped MPPs were similar to one another.

For model A (left panels of Fig. 3), the predicted threshold differences between ramped and damped maskers were on the order of 3 dB. Among all half-life and signal-delay conditions, the difference between thresholds for the ramped and damped maskers was tested for statistical significance using a t test assuming unequal variances with a Bonferonni correction ($p < 0.002$). For model A, only two out of 25 conditions demonstrated statistically significant differences

(marked with asterisks in the left panels of Fig. 3). For model B (right panels of Fig. 3), the predicted threshold differences were even smaller than that of model A, on the order of 1 .5 dB. No conditions were statistically different for ramped and damped maskers.

Despite the fact that models A and B had different levels of sophistication and came from psychophysical (model A) or physiological (model B) origins, they produced very similar results. Both of the models failed in producing temporal asymmetry even though both included peripheral nonlinearities and model B included relatively realistic auditory-nerve adaptation stages. The failure of these peripheral models to produce a large temporal asymmetry in the MPPs suggests that central processes are involved in the perception of temporal asymmetry.

IV. EXPERIMENT II: EFFECT OF MASKER LEVEL

In experiment I, large MPP differences were measured between the ramped and the damped noise maskers at short half-lives. This experiment investigated whether this asymmetric effect is level dependent. The results from this experiment should help revealing the nonlinear nature of the phenomenon, which might further indicate possible origins of temporal asymmetry. Failure of the models of peripheral processing to account for level-dependent effects would further implicate a central contribution to temporal asymmetry.

A. Methods

Three normal-hearing listeners, ranging in age from 19 to 25 years, with a mean age of 21 years, participated. Two of the listeners were participants in experiment I. All had absolute thresholds of 15 dB HL or better at audiometric frequencies between 250 and 4 000 Hz (ANSI, 2004).

MPPs were measured for ramped and damped maskers having 4-ms half-lives presented at four different stimulus levels: 40, 55, 70, and 85 dB SPL. These overall levels correspond to spectrum levels of approximately 0, 15, 30, and 45 dB SPL. As in experiment I, MPPs were measured using a randomized block design. Each trial block consisted of five threshold estimates, one for each signal delay. The sequence of signal delays was randomly selected at the start of each block. The overall masker level to be tested was chosen randomly, and then the masker presentation sequence (either ramped or damped) was selected. Once two blocks were completed at each stimulus level (one MPP obtained for the ramped and damped modulators), the two blocks were run at a new stimulus level. After an MPP was obtained for ramped and damped maskers at each level, additional blocks were tested in the same random order. The experiment was concluded once five thresholds for each masker type were obtained. Reported thresholds are the average of the final four threshold estimates in each condition. Note that because two of the listeners completed the 70-dB stimulus level in experiment I, their blocks only consisted of maskers presented at 40, 55, and 85 dB SPL.

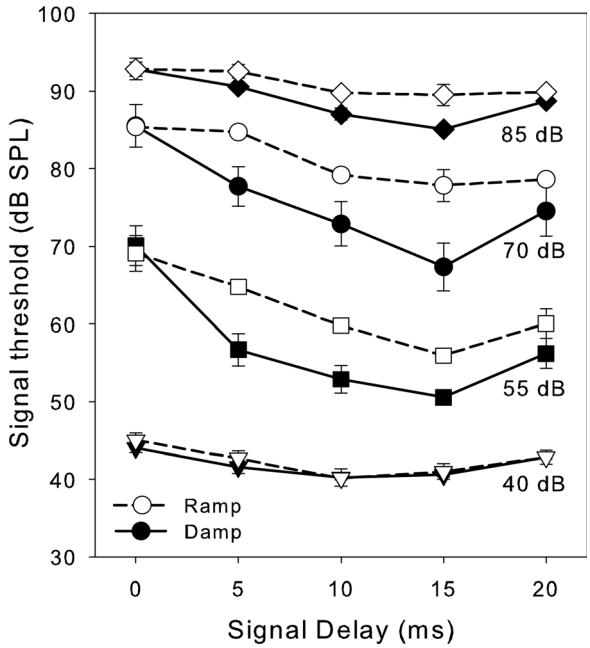

FIG. 4. Masking period patterns for ramped (unfilled symbols) and damped (filled symbols) noises measured at different masker levels are shown. Different symbols are used to denote the data obtained at the different stimulus levels. As in Fig. 2, thresholds for the ramped stimulus are time-reversed. Error bars represent standard errors of the mean.

B. Experimental results

Figure 4 plots MPPs at four different stimulus levels averaged across the three listeners. As in Fig. 2, the MPPs for the ramped stimulus were time-reversed for ease in visualization. The MPPs for ramped and damped noises were highly modulated with thresholds being dependent on signal delay for all stimulus levels tested. MPPs at the moderate levels (55 and 70 dB SPL) tended to have the greatest amount of modulation, whereas the MPPs at 40 and 85 dB SPL tended to be more flat. At all stimulus levels except at 40 dB SPL, where damped and ramped MPPs overlap, masked thresholds at the signal delays of 5–20 ms were higher for the ramped noise than for the damped noise. All masked thresholds were at least 10 dB above the threshold for the 1 000-Hz tone burst in quiet (about 30 dB SPL), but we note that the flat MPPs at 40 dB SPL could have occurred because portions of the stimuli may have been inaudible to the listeners due to the low stimulus levels.

A repeated measures ANOVA was conducted on the data treating stimulus level, masker type, and signal onset as within-subjects factors. As for experiment I, the values for the thresholds obtained using the ramped masker are also time-reversed in this ANOVA. Table II reports that the ANOVA revealed significant main effects of all factors: stimulus level, masker type, and signal delay. The only interaction not significant was that between masker type and signal delay, indicating that overall, the shape of the ramped MPP may have been similar to the shape of the damped MPP. The three-way interaction, however, implies that stimulus level has an impact on this relationship. Essentially, the ANOVA is consistent with a change in MPP depth associated with changes in stimulus level and also a change in the

TABLE II. Results of ANOVA for experiment II.

Factor	F	Significance
Level	$F(3,6) = 625.7$	$p < 0.001$[a]
Masker type	$F(1,2) = 19.8$	$p < 0.05$[a]
Delay	$F(4,8) = 83.8$	$p < 0.001$[a]
Level × masker type	$F(3,6) = 9.2$	$P < 0.02$[a]
Level × delay	$F(12,24) = 8.3$	$p < 0.001$[a]
Masker type × delay	$F(4,8) = 2.9$	$p = 0.09$
Level × masker type × delay	$F(12,24) = 4.0$	$p < 0.002$[a]

[a]Significant at $p < 0.05$.

ramped/damped difference associated with increases in stimulus level.

C. Growth of masking

Another informative way to evaluate the nature of the level-dependent effect is by plotting growth of masking (GOM) functions, which describe the change in signal threshold with increasing masker levels. GOM functions are plotted in Fig. 5 for the two masker types (left and right panels) and the five signal delays used for each stimulus. Note that again, thresholds for the ramped stimulus are time-reversed to better visualize differences between the two stimuli. The dotted line is included as a linear reference.

Although many of the GOM functions appear linear, others appear to have a curvature, with a steeper slope at the lower stimulus levels and a shallower slope at the higher stimulus levels. To determine which of these functions have a significant curved shape, we conducted a second order curve fit and tested whether the quadratic term was significantly different from zero. In this regression, negative terms indicate that the functions are steeper at lower levels than at higher levels (e.g., damp at 0 ms), whereas positive terms would indicate functions growing more quickly at the higher stimulus levels. When this term is zero, the functions are linear.

Table III, which shows the values of the quadratic terms, indicates that the majority of the functions have square terms that are negative (all of the ramped functions and the 0- and 5-ms damped functions). However, the square terms that are significantly different from zero occur at the 0-ms delay for both ramped and damped functions (diamonds in Fig. 5) and not at the other delays. Thus, for both ramped and damped stimuli, the nonlinear GOM functions occur only at the temporal peaks of the stimuli. The GOM functions obtained at the other signal delays are fitted well with a linear function. The nonlinear GOM function has similar characteristics to that observed in overshoot experiments and is commonly observed at the onsets of stimuli (Bacon and Savel, 2004).

TABLE III. Values of the square term in the quadratic regression for the GOM functions shown in Fig. 5.

	Signal delay (ms)				
	0	5	10	15	20
Ramp	−0.018[a]	−0.016	−0.010	−0.004	−0.007
Damp	−0.021[a]	−0.003	0.002	0.009	0.001

[a]Significant at $p < 0.05$.

D. Modeling results

Predictions from models A and B are shown in the left and right panels of Fig. 6, respectively, together with the average data of experiment II. The predicted MPPs from both models illustrated modulated patterns with distinct peaks and valley. At moderate stimulus levels, the shape of the MPPs for the model predictions was similar to the shape measured in the data. At 40 dB SPL, the measured thresholds varied little with signal delay and were between 40 and 45 dB SPL, whereas the predicted thresholds were higher and varied more within the masker period. At 85 dB SPL, the range of thresholds in the predicted MPPs was also larger than the data, which showed fairly flat MPP shapes. As in

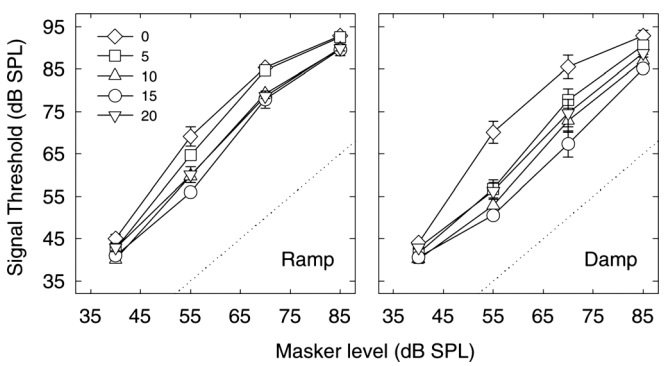

FIG. 5. Growth of masking function for ramped and damped maskers are plotted in the left and right panels, respectively. Five different signal delays are indicated with different symbols. The dotted line without symbols is a linear reference. Error bars represent standard errors of the mean.

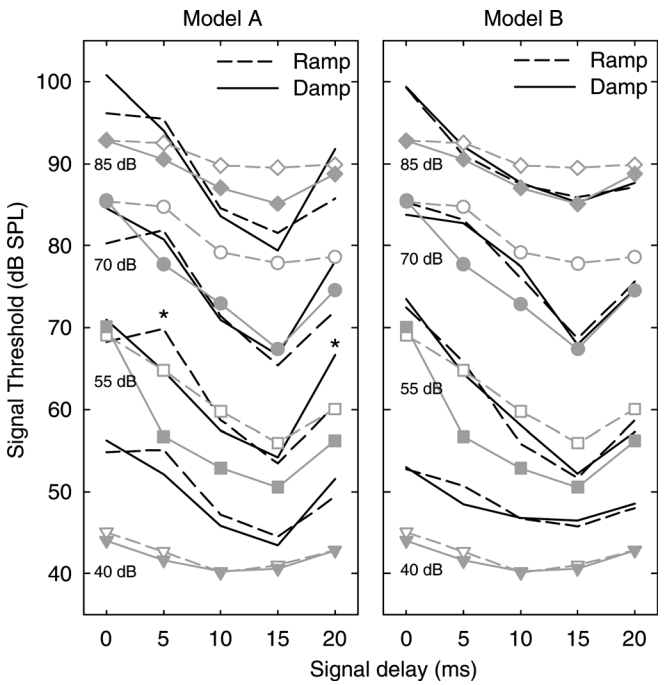

FIG. 6. Model A and model B predictions are plotted along with the mean data of experiment II in the left and right panels, respectively. Predictions for the ramped and damped stimuli are shown as dashed and solid lines, respectively.

experiment I, no salient temporal asymmetry effect was predicted by the two models for all conditions.

Model A (left panel of Fig. 6) predicted a relatively small level effect suggesting a linear growth of masking. Due to this linearity, the model mis-predicted the MPPs at masker levels of 40 and 85 dB SPL. The deviations from the predicted MPPs to the average MPPs from the experiment were as large as 13 dB (at the 40-dB SPL masker level). The predicted threshold differences between the ramped and damped MPPs were small (within 3 dB) across conditions. As for experiment I, the model did not predict temporal asymmetry for any of the stimulus levels. The difference between thresholds for the ramped and damped maskers was tested for statistical significance using a t test assuming unequal variances with a Bonferonni correction ($p < 0.0025$). For model A, only two out of 25 conditions demonstrated statistically significant differences (marked with asterisks in the left panels of Fig. 6).

Compared to model A, Model B predicted the level effect on the MPP shape slightly better (right panel of Fig. 6). The predicted MPP shape was relatively flat at 40 dB SPL, steepest at 55 and 70 dB SPL, with some flattening at 85 dB SPL. This result indicates that that model B encapsulated the growth of masking relatively better than model A. However, at 40 dB SPL, model B consistently predicted higher thresholds than those measured in the data by about 7 dB, and model B did not produce statistically significant temporal asymmetry in any of the conditions. Although model B simulated the growth of masking more closely than model A, it still failed to generate threshold differences between ramped and damped MPPs.

V. DISCUSSION

A. Failure of peripheral processing models

The results indicate that the MPPs for damped and time-reversed MPPs for ramped maskers show substantial differences in their shape, with the MPP obtained using damped stimuli being more modulated than the time-reversed ramped MPP. This finding suggests that the ramped noise has a shallower internal envelope and less internal modulation than the damped noise. Two models of auditory peripheral processing were not able to capture the temporal asymmetry in the MPPs, nor could they encapsulate the half-life and level dependence of the phenomenon.

Within model A the processing stages that could potentially introduce temporal asymmetry are the auditory filter and the temporal leaky integrator, both of which have temporally asymmetric impulse responses. Previous studies using a similar model were also unsuccessful in predicting temporal asymmetry (e.g., Akeroyd and Patterson, 1995; Carlyon, 1996; Irino and Patterson, 1996). These studies demonstrated that the combination of the auditory filter and temporal integrator predicts temporal asymmetries that are much smaller than observed (Irino and Patterson, 1996). Indeed, we replicated these findings in modeling MPP data and showed that model A predicted almost identical MPPs for ramped and damped noises.

Compared to model A, model B possesses an advanced processing stage of auditory-nerve adaptation, which enables the model to encapsulate auditory-nerve responses to amplitude-modulated stimuli reasonably well (Zilany *et al.*, 2009; Zilany and Carney, 2010). However, model B also did not lead to differences between the ramped and damped MPPs. To establish whether other adaptation mechanisms might have accounted for the temporal asymmetry, we evaluated adaptation loops (e.g., Dau *et al.*, 1996, 1997) and a hair-cell model (e.g., Meddis, 1986; Meddis *et al.*, 1990), but neither of these models created threshold differences between the predicted ramped and damped MPPs near the MPP valleys. In a similar vein, Irino and Patterson (1996) reported that extreme parameter settings were required for the adaptation stage in a peripheral processing model to encapsulate the large perceptual asymmetry. It is worth noting that these models may not adequately capture all aspects of peripheral processing, and small differences in the output of the auditory periphery are likely enhanced by higher stages in the auditory system. Consequently, we do not reject the idea that the peripheral nonlinearity is a source of temporal asymmetry. However, converging evidence based on auditory modeling and physiological experiments indicate that the central auditory system has a significant role to play in producing temporal asymmetry.

One possible source of temporal asymmetry is related to the idea that broadband masker energy is known to influence masking results for amplitude-modulated maskers, such as those used here. Coherently modulated masker energy in on- and off-frequency channels can cause a release from masking, typically referred to as comodulation masking release or CMR (Hall *et al.*, 1984). Across-frequency suppression arising from the central auditory pathway might contribute to the CMR effect (e.g., Verhey *et al.*, 2003). Because the current ramped and damped stimuli are amplitude modulated, there is a possibility that across-frequency processing underlying CMR plays an important role in temporal asymmetry. Akeroyd and Patterson (1995) argued that damped noises have a greater amount of envelope coherence across frequency than the ramped stimuli. As such, it seems plausible that the lower signal detection thresholds observed in the damped MPPs could be due to a greater CMR for damped than for ramped stimuli. The current experiment did not include appropriate conditions to asses the role of CMR and would require future tests, but greater masking release for the damped stimuli seems possible.

B. Effect of masker level

In experiment II GOM functions were measured for ramped and damped noises. For both ramped and damped maskers, GOM functions were nonlinear at the envelope peaks (signal delay = 0) but not significantly different from linear at the other signal delays tested. The nonlinear GOM functions had a more rapid growth of masking at the lower stimulus levels and a shallower growth of masking at the higher stimulus levels. This shape is similar to GOM functions measured in overshoot (or "temporal effect") experiments (Bacon, 1990; Bacon and Savel, 2004).

The phenomenon of overshoot occurs when the detection threshold of a tone pip is higher at the onset of a masker relative to its threshold in the temporal center (Zwicker, 1965). Using an on-frequency broadband masker, Bacon and Savel (2004) observed a linear GOM in the center of the masker and a nonlinear GOM at the onset similar to that observed here. Bacon and Savel (2004) explained the nonlinear GOM in terms of the compressive cochlear nonlinearity, which has a compressive input–output (I–O) function at moderate levels and is relatively linear at very low and very high levels (e.g., Ruggero *et al*., 1997). At stimulus onset, the signal level at threshold is often much higher than the masker level (large signal-to-masker ratio). When the masker is presented at low levels, the signal falls within the compressive region of the I–O function while the masker is processed by the linear (low-level) portion of the I–O function. When the masker level is moderately high, the masker is processed by the compressive portion of the I–O function while the signal enters into the upper linear (high-level) portion of the I–O function. Thus, the slope of the GOM function can be steep at low masker levels and linear or shallow at the higher masker levels.

A similar rationale has been adopted to explain the growth of masking under forward masking conditions. Plack and Oxenham (1998) found the GOM for forward masking to be nonlinear with a shallow slope at low masker levels that became steeper with increasing level. Plack and Oxenham (1998) suggested that the compressive nonlinearity was the source for the nonlinear GOM. In this case, the signal level was typically lower than the masker level at threshold (small signal-to-masker ratio) producing an increase in GOM slope with increasing masker level.

Bacon *et al*. (1997) and Bacon and Lee (1997) showed that when a broadband masker was highly modulated, signal detection in the temporal peaks was dominated by simultaneous masking whereas signal detection in the temporal valleys of the masker was dominated by forward masking from the preceding masker peaks. Forward masking might play a role in the current data because highly modulated maskers were used. At the temporal peaks, thresholds would be likely to be governed by simultaneous masker energy while at the temporal valleys thresholds could also be determined by the amount of forward masking. For simultaneous masking, a large signal-to-masker ratio would be needed for detection, thereby leading to nonlinear GOM functions similar to those obtained from overshoot experiments. The GOM functions for the ramped and damped maskers at the 0-ms signal delay were consistent with this interpretation.

For the case of forward masking in the valleys, the masker peak would be processed at the same location on the I–O function as in the simultaneous masking case, but the location of the signal would change depending on the amount of forward masking produced by a stimulus. If one stimulus produces less forward masking than another, the signal-to-masker ratio needed for detection would be lower for the stimulus that produces less forward masking. This would shift the location of the signal to a lower portion on the I–O function. Consequently, the masker and signal would now be processed by more similar amounts of compression, leading to a more linear GOM. Following this argument, a stimulus that produces less forward masking would yield more linear GOM functions in masker valleys. Recall that the statistical analysis of the GOM functions showed GOM functions obtained at delays away from the masker peaks were not significantly different from linear. These linear GOM functions could have occurred if the temporal peaks in the ramped and damped stimuli produce similar forward masking effects on the subsequent masker valleys. Although DiGiovanni and Schlauch (2007) found that a single ramped period produces more forward masking than a damped period, data collected from a follow-up experiment using a single pulse of the 70-dB maskers used here (4-ms half-life and 25 ms in duration) indicated very little differences in forward masking between the rising and falling stimuli. Consequently, the similarity in the GOM functions for ramped and damped maskers suggests that simultaneous masking dominated at the temporal masker peaks and that there were little differences in within-period forward masking for the ramped and damped maskers.

C. Relating temporal asymmetry in masking, discrimination, and detection tasks

One motivation of the current experiments was that Patterson (1994a,b) and Akeroyd and Patterson (1995) showed that ramped noises often can be easily discriminated from damped noises with one possibility being that internal representation of a ramped stimulus has a shallower (flatter) envelope than the damped stimulus. The MPPs measured here support this hypothesis, and they also follow the same pattern as the masking patterns of single pulses of the ramped and damped noises measured by Ries *et al*. (2008). However, there are some notable discrepancies between the MPPs and the data from Patterson and Akeroyd's experiments.

First, if the MPPs measured here were used to predict Patterson and Akeroyd's data, one would predict that the stimuli with the 1-ms half-life would be easier to discriminate from one another than stimuli at the longer half-lives because the differences measured between ramped and damped MPPs are larger then the differences at the other half-lives. However, Akeroyd and Patterson (1995) showed that ramped and damped noises with 1-ms half-life could be discriminated from one another 85%–90% of the time, whereas the stimuli with 2-, 4-, and 8-ms half-lives were 100% discriminable. A decrease in discrimination ability was measured for stimuli with a 16-ms half-life, which had 85% discriminability. Thus, there is no evidence that 1-ms ramped and damped noises are more easily discriminable than ramped and damped noises with longer half-lives. The MPPs measured here at the 16-ms half-life are effectively flat, and would suggest relatively poor discriminability, yet Akeroyd and Patterson (1995) showed that the 16 and 1-ms stimuli had similar levels of discriminability. To the extent that MPPs provide a measurement of the internal modulation present in the stimuli, it seems unlikely that ramped and damped noises are discriminable based on their internal envelope shapes alone.

One difference between discrimination experiments and the masking experiments used here is that in discrimination, listeners are able to use a multitude of cues to determine whether two the sounds differ. Although we have illustrated that ramped and damped stimuli have different internal modulation depths, there are likely other differences in the representations of these stimuli. Akeroyd and Patterson (1995) used the Auditory Image Model [described in Patterson *et al.* (1995)] to show that the damped noises have greater envelope coherence across frequency than for ramped noises. Thus, the modulation patterns across frequency are more consistent for damped noises than ramped noises. These across-frequency differences could be used by listeners to discriminate the two types of stimuli.

Second, the masking data presented also seem to be inconsistent with modulation detection data. For example, the data of both experiments would lead to a prediction that damped modulation should often be easier to detect than ramped modulation because of the larger MPP depth for the damped noises. Akeroyd and Patterson (1997) measured modulation detection thresholds for broadband noise stimuli modulated with asymmetric sinusoidal modulators of varying modulation rates. One modulator consisted of only the rising portion of a sinusoidal cycle and an instantaneous decrease (U-SAM). The other modulator contained the downward portion of a sinusoidal cycle with an instantaneous increase (D-SAM). No difference in the ability to detect U-SAM or D-SAM was measured for any of the modulation rates.

In summary, it appears that the most likely explanation for these discrepant results is that listeners use different cues for detection of modulation, discrimination of modulated stimuli, and detecting a tone added to amplitude-modulated stimuli.

VI. SUMMARY AND CONCLUSIONS

MPPs demonstrated that ramped maskers have less modulated internal waveforms than damped noises. This effect was also found to be level dependent with the greatest differences occurring at moderate stimulus levels. The shape of the MPPs and the level dependent effects are likely due to both peripheral and central auditory processes. Two models of auditory peripheral processing did not adequately account for the observed differences between ramped and damped maskers. Consequently, the role of the auditory periphery in temporal asymmetry is potentially limited. Centrally mediated mechanisms, such as overshoot and CMR, could have contributed to the asymmetric masking effects of the ramped and damped noises. The MPPs measured do not fully explain the experimental data on discriminating or detecting ramped and damped sounds, which suggests the importance of different cues in discrimination and detection tasks that are not captured in the current masking experiments.

ACKNOWLEDGMENTS

This work was supported by the Indiana University's Faculty Research Support Program. We thank Matthew Rosenthal, Lindsay Weberling, and Joshua Elzinga for their contributions to data collection and subject recruitment.

[1]The model of Zilany *et al.* (2009) is capable of generating neural firing patterns of the modeled auditory nerve fiber in terms of peri-stimulus time histograms (PSTHs). Using PSTHs as the output increased computational effort significantly compared to using the inner-hair cell synaptic output. Although the expectation would be that the PSTHs would give different predictions than the synapse output due to the neural refractory period in the nerve fiber, we found that the two types of outputs did not produce observable differences in the model predictions. Consequently, the synaptic output was used as the model output.

[2]We did run the simulations both with and without the leaky integrator and were not able to predict the shape of the MPP functions without using the leaky integrator.

[3]The low-pass filter in the frequency domain is equivalent to a integrator in the time domain. The ERD of the filter's impulse response or the ERD of the temporal window, is related to the filter's cutoff frequency f_{cutoff} by $\text{ERD} = 1/(2\pi f_{\text{cutoff}})$.

[4]Pilot tests showed that the model predictions were not very sensitive to the cutoff frequency of the temporal integrator between values of 20–40 Hz. Moreover, when the cutoff frequency was increased to 65 Hz to represent auditory temporal resolution assessed via amplitude-modulation detection tasks (Viemeister, 1979), model A showed poor performance in encapsulating the measured MPP shapes, while the predictions from model B did not change significantly.

[5]Note that the delay introduced by the Zilany *et al.* (2009) model was measured and compensated for.

[6]The presence of the internal noise was represented by the term σ_0 in Eq. (2).

Akeroyd, M. A., and Patterson, R. D. (**1995**). "Discrimination of wideband noises modulated by a temporally asymmetric function," J. Acoust. Soc. Am. **98**, 2466–2474.

Akeroyd, M. A., and Patterson, R. D. (**1997**). "A comparison of detection and discrimination of temporal asymmetry in amplitude modulation," J. Acoust. Soc. Am. **101**, 430–439.

Alcántara, J. I., Holube, I., and Moore, B. C. J. (**1996**). "Effects of phase and level on vowel identification: data and predictions based on a nonlinear basilar-membrane model," J. Acoust. Soc. Am. **100**, 2382–2392.

ANSI (**2004**). ANSI S3.6–2004, American National Standard Specification for Audiometers (Revision of ANSI S3.6–1996) (American National Standards Institute, New York).

Bacon, S. P. (**1990**). "Effect of masker level on overshoot," J. Acoust. Soc. Am. **88**, 698–702.

Bacon, S. P., and Lee, J. (**1997**). "The modulated–unmodulated difference: Effects of signal frequency and masker modulation depth," J. Acoust. Soc. Am. **101**, 3617–3624.

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

Bacon, S. P., and Savel, S. (**2004**). "Temporal effects in simultaneous masking with on- and off-frequency noise maskers: Effects of signal frequency and masker level," J. Acoust. Soc. Am. **115**, 1674–1683.

Carlyon, R. P. (**1996**). "Spread of excitation produced by maskers with damped and ramped envelopes," J. Acoust. Soc. Am. **99**, 3647–3655.

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.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the effective signal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am. **99**, 3615–3622.

DiGiovanni, J. J., and Schlauch, R. S. (**2007**). "Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds," Ecol. Psychol. **19**, 239–264.

Fastl, H. (**1977**). "Subjective duration and temporal masking patterns of broadband noise impulses," J. Acoust. Soc. Am. **61**, 162–168.

Fastl, H. (**1982**). "Fluctuation strength and temporal masking patterns of amplitude-modulated broadband noise," Hear. Res. **8**, 59–69.

Grassi, M., and Darwin, C. J. (**2006**). "The subjective duration of ramped and damped sounds," Percept. Psychophys. **68**, 1382–1392.

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.

Irino, T., and Patterson, R. D. (**1996**). "Temporal asymmetry in the auditory system," J. Acoust. Soc. Am. **99**, 2316–2331.

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.
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.
Lu, T., Liang, L., and Wang, X. (**2001**). "Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates," J. Neurophysiol. **85**, 2364–2380.
Meddis, R. (**1986**). "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711.
Meddis, R., Hewitt, M. J., and Shackleton, T. M. (**1990**). "Implementation details of a computation model of the inner hair-cell auditory-nerve synapse," J. Acoust. Soc. Am. **87**, 1813–1816.
Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). "The shape of the ear's temporal window," J. Acoust. Soc. Am. **83**, 1102–1116.
Mott, J. B., McDonald, L. P., and Sinex, D. G. (**1990**). "Neural correlates of psychophysical release from masking," J. Acoust. Soc. Am. **88**, 2682–2691.
Neuert, V., Pressnitzer, D., Patterson, R. D., and Winter, I. M. (**2001**). "The responses of single units in the inferior colliculus of the guinea pig to damped and ramped sinusoids," Hear. Res. **159**, 36–52.
Neuhoff, J. G. (**1998**). "Perceptual bias for rising tones," Nature (London) **395**, 123–124.
Oxenham, A. J., and Dau, T. (**2001**). "Reconciling frequency selectivity and phase effects in masking," J. Acoust. Soc. Am. **110**, 1525–1538.
Oxenham, A. J., and Moore, B. C. J. (**1994**). "Modeling the additivity of nonsimultaneous masking," Hear. Res. **80**, 105–118.
Patterson, R. D. (**1994a**). "The sound of a sinusoid: Spectral models," J. Acoust. Soc. Am. **96**, 1409–1418.
Patterson, R. D. (**1994b**). "The sound of a sinusoid: Time-interval models," J. Acoust. Soc. Am. **96**, 1419–1428.
Patterson, R. D., Allerhand, M. H., and Giguere, C. (**1995**). "Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform," J. Acoust. Soc. Am. **98**, 1890–1894.
Plack, C. J., and Moore, B. C. J. (**1990**). "Temporal window shape as a function of frequency and level," J. Acoust. Soc. Am. **87**, 2178–2187.
Plack, C. J., and Oxenham, A. J. (**1998**). "Basilar-membrane nonlinearity and the growth of forward masking," J. Acoust. Soc. Am. **103**, 1598–1608.
Pressnitzer, D., Meddis, R., Delahaye, R., and Winter, I. M. (**2001**). "Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus," J. Neurosci. **21**, 6377–6386.
Pressnitzer, D., Winter, I. M., and Patterson, R. D. (**2000**). "The responses of single units in the ventral cochlear nucleus of the guinea pig to damped and ramped sinusoids," Hear. Res. **149**, 155–166.
Ries, D. T., Schlauch, R. S., and DiGiovanni, J. J. (**2008**). "The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds," J. Acoust. Soc. Am. **124**, 3772–3783.
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.
Schlauch, R. S., Ries, D. T., and DiGiovanni, J. J. (**2001**). "Duration discrimination and subjective duration for ramped and damped sounds," J. Acoust. Soc. Am. **109**, 2880–2887.
Stecker, G. C., and Hafter, E. R. (**2000**). "An effect of temporal asymmetry on loudness," J. Acoust. Soc. Am. **107**, 3358–3368.
Strickland, E. A. (**2000**). "The effects of frequency region and level on the temporal modulation transfer function," J. Acoust. Soc. Am. **107**, 942–952.
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. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.
Wojtczak, M., Schroder, A. C., Kong, Y. Y., and Nelson, D. A. (**2001**). "The effect of basilar-membrane nonlinearity on the shapes of masking period patterns in normal and impaired hearing," J. Acoust. Soc. Am. **109**, 1571–1586.
Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," J. Acoust. Soc. Am. **126**, 2390–2412.
Zilany, M. S. A., and Carney, L. H. (**2010**). "Power-law dynamics in an auditory-nerve model can account for neural adaptation to sound-level statistics," J. Neurosci. **30**, 10380–10390.
Zwicker, E. (**1965**). "Temporal effects in simultaneous masking and loudness," J. Acoust. Soc. Am. **38**, 132–141.
Zwicker, E. (**1976a**). "Masking period patterns of harmonic complex tones," J. Acoust. Soc. Am. **60**, 429–439.
Zwicker, E. (**1976b**). "Psychoacoustic equivalent of period histograms," J. Acoust. Soc. Am. **59**, 166–175.