

# Investigating the auditory enhancement phenomenon using behavioral temporal masking patterns

Yi Shen<sup>a)</sup> and Virginia M. Richards

*Department of Cognitive Sciences, University of California, Irvine, California 92697-5100*

(Received 8 July 2011; revised 22 August 2012; accepted 28 August 2012)

A narrowband signal is subjected to less masking from a simultaneously presented notched masker if it is preceded by a precursor that occupies the same spectral region as the masker, a phenomenon referred to as enhancement. The present study investigated (i) the amount of enhancement for the detection of a narrowband noise added to a notched masker, and (ii) masking patterns associated with the detection of tone pips added to the narrowband signal. The resulting psychophysical data were compared to predictions generated using a model similar to the neural adaptation-of-inhibition model proposed by Nelson and Young [(2010b). *J. Neurosci.* **30**, 6577–6587]. The amount of enhancement was measured as a function of the temporal separation between the precursor and masker in Experiment I, and as a function of precursor level in Experiment II. The model captured the temporal dynamics of psychophysical enhancement reasonably well for both the long-duration noise signals and the masking patterns. However, in contrast to psychophysical data which indicated reliable enhancement only when the precursor and masker shared the same levels, the model predicated enhancement at all precursor levels. © 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4754527]

PACS number(s): 43.66.Dc, 43.66.Lj [CJP]

Pages: 3363–3374

## I. INTRODUCTION

Simultaneous masker energy is not the only determining factor for the detection threshold of a signal. A copy of the masker (a precursor) presented prior to the concurrent signal and masker (e.g., the left panel of Fig. 1) can release the signal from simultaneous masking by as much as 15 dB (e.g., Viemeister, 1980; Carlyon, 1989; McFadden, 1989; Strickland, 2004). Psychophysically, such auditory enhancement has been demonstrated using various types of stimuli, including tone complexes (e.g., Viemeister, 1980; Richards and Neff, 2004; Richards *et al.*, 2004), noises (e.g., Wright *et al.*, 1993), and speech stimuli (e.g., Summerfield *et al.*, 1987; Summerfield and Assmann, 1989). The amount of enhancement is largest when the temporal separation between the precursor and the masker is short (Viemeister, 1980; McFadden, 1989), and when the precursor/masker level is high (Viemeister, 1980; Hicks and Bacon, 1992). The current study replicates some of these previous findings by measuring enhancement as a function of the duration of the temporal gap between the precursor and masker and the precursor level. Additionally, temporal masking patterns are measured to probe the temporal dynamic of the enhancement phenomenon. The psychophysical data are compared to predictions given by a phenomenological model of auditory processing, in order to test an adaptation-of-inhibition hypothesis for enhancement.

Various hypotheses have been proposed previously to explain the origins of the enhancement effect. One explanation proposes that enhancement is a consequence of peripheral adaptation (Viemeister, 1980), i.e., the reduction in the

sensitivity of the peripheral auditory system following sound exposure. By this argument, if the precursor and masker have the same spectra, the precursor would selectively reduce the internal representation of the masker but not the signal whose frequency resides outside the precursor's bandwidth. As a result, the signal-to-noise ratio at the output of the auditory periphery would be increased by the presence of the precursor. However, as pointed out by Summerfield *et al.* (1987), auditory nerve adaptation alone cannot fully account for the amount of enhancement because the time constants associated with the recovery from peripheral adaptation are typically shorter than those for the enhancement effect.

An alternative explanation for the enhancement phenomenon proposed by Viemeister and Bacon (1982) suggested a more complex role of adaptation. In one experiment, Viemeister and Bacon (1982) studied the enhancement of a target tonal component in a harmonic complex following a precursor that consisted of the frequency components of the masker with the target component and the components just above and below the target component removed. The level of the target component was held constant and its efficiency in forward masking a brief tone pip was measured. The authors found that introducing the precursor increased the forward masked thresholds, suggesting that the excitation generated by the target was increased in the auditory system. Therefore, the enhancement effect was probably not due to the attenuation of the masker by peripheral adaptation alone; it was accompanied by (or due to) an amplification in the target region. Viemeister and Bacon (1982) proposed an adaptation-of-suppression explanation of the enhancement phenomenon, which stated that a notched masker could cause reduced sensitivity in the region of its spectral notch through a lateral suppression mechanism. However, due to adaptation, the amount of suppression is

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: shen.yi@uci.edu

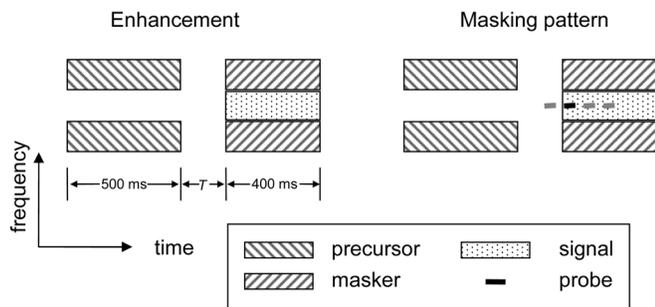


FIG. 1. Schematics of the stimuli. In the enhancement measurements (left) the detection thresholds for an added narrowband noise signal were determined. In the masking pattern measurements (right), the detection thresholds for 6-ms tonal probes were obtained. For the latter measurements, the signal level was 5 dB lower than the masker level (see text for details).

reduced following the precursor, leading to an effectively more intense signal in the notch.

Inspired by the adaptation-of-suppression hypothesis, several studies have been conducted to investigate whether the “suppression” described by Viemeister and Bacon (1982) is peripheral with mixed results. Strickland (2004) extended the hypothesis and suggested that the signal receives suppression from the surrounding spectral region occupied by the masker. When a precursor is presented immediately before the masker, the cochlear gain in the masker frequency region is decreased, possibly by the activation of the olivocochlear efferent feedback system. This causes a decrease in suppression received at the signal frequency. Using a phenomenological model of auditory peripheral processing, Strickland (2004, 2008) demonstrated that the extended adaptation-of-suppression hypothesis could potentially be used to explain the enhancement effect. On the other hand, Wright *et al.* (1993) measured both enhancement and psychophysical two-tone suppression and found that listeners with stronger suppression exhibited reduced enhancement, which was not predicted by the adaptation-of-suppression hypothesis.

Palmer *et al.* (1995) investigated the enhancement phenomenon physiologically at the level of the auditory nerve in guinea pigs. Their stimuli paralleled those used by Viemeister and Bacon (1982). For auditory nerve fibers tuned to the signal frequency, changes in spike rate were not found to depend on the presence or absence of a precursor. While no precursor-based gain in spike rate was observed for fibers tuned to the signal frequency, off-signal-frequency fibers exhibited more adaptation of responses owing to the presence of the precursor. In a framework of the adaptation-of-suppression hypothesis, this physiological finding suggested that a lateral inhibition mechanism at a location more central relative to the auditory nerve may be responsible for psychophysical enhancement. Note that the findings of Palmer *et al.* (1995) did not eliminate the possibility of a peripheral contribution to the enhancement phenomenon because the anesthesia used in this study might have affected the olivocochlear efferent feedback system, which could potentially contribute to the phenomenon (Nelson and Young, 2010a). As an example of evidence of a non-peripheral component to enhancement, Wang *et al.* (2012) observed an enhancement effect in cochlear implant users.

Nelson and Young (2010a,b) measured neural correlates of psychophysical enhancement at the level of the inferior colliculus (IC). They found that many neurons in the IC exhibited enhanced firing during the presentation of a signal when a precursor was presented. They suggested a phenomenological model to explain the observed data. This consisted of a model of auditory peripheral processing (Heinz *et al.*, 2001), coupled with a wideband inhibition mechanism. The responses of units tuned to the signal frequency and inhibition from units tuned to a broad range of off-signal frequencies were integrated. By this mechanism the masker inhibits activity in response to the signal. Moreover, when the precursor is present, the response to the masker is adapted, and thus the inhibition to the signal is reduced, providing an effective gain in response to the signal. This model provides a neural account that is consistent with the adaptation-of-suppression hypothesis proposed by Viemeister and Bacon (1982). Recognizing the role of inhibition in the Nelson and Young (2010b) model, here the term “adaptation of inhibition,” rather than “adaptation of suppression” is used to differentiate this class of models from those models based on peripheral suppression (e.g., Strickland, 2004; Jennings *et al.*, 2011).

The current study compared psychophysical data and predictions of a model based on the work of Nelson and Young (2010a,b). In particular, the goal was to determine whether a variant of their model could provide an understanding of (i) the effects of the temporal separation between the precursor and masker and (ii) the effects of precursor level on auditory enhancement. In an effort to psychophysically evaluate changes in enhancement over relatively short time spans, masking patterns (e.g., Zwicker, 1965; Fastl, 1977) were measured. Additionally, the model proposed by Nelson and Young (2010b) was implemented to provide a comparison between predictions of an adaptation-of-inhibition model and psychophysical data. If the model was able to accurately predict both the enhancement and masking pattern data, it would provide support for an adaptation-of-inhibition hypothesis of psychophysical enhancement.

## II. GENERAL METHODS

### A. Stimuli

#### 1. Enhancement measurements

In both Experiments (Exps.) I and II, two psychophysical measurements were completed: enhancement and masking pattern. In the enhancement measurements, detection thresholds of a narrowband-noise (approximately 0.6-octave wide) signal were measured in the presence of a masker preceded by a precursor. Before each interval of each trial, the center frequency of the signal,  $f_c$ , was drawn at random from a uniform distribution on a log-frequency scale spanning a one-octave range centered at 2 kHz. By randomizing the spectral location of the signal, the influence of the previous presentation was minimized, i.e., to ensure that the measured enhancement depended on the precursor, not the stimulus from the previous presentation. The bandwidth of the signal spanned from  $f_c/1.2$  ( $0.833f_c$ ) to  $1.2f_c$ . To generate the signal, a spectrum was first constructed by combining a

Raleigh-distributed magnitude response and a uniformly distributed phase response ( $0-2\pi$ ) and setting all spectral content outside of the signal bandwidth to zero. The spectrum was then inverse-Fourier transformed into the time domain. The same procedure was used to generate the masker and precursor. The masker and precursor were notched noises, each of which consisted of an upper frequency band between  $1.2f_c$  and  $2.4f_c$  and a lower frequency band between  $f_c/2.4$  ( $0.417f_c$ ) and  $f_c/1.2$  ( $0.833f_c$ ). The widths of the spectral notches in the precursor and masker were chosen so that an enhancement effect would be expected according to previously published enhancement studies (e.g., Carlyon, 1989). In Exp. I the masker and precursor levels were always presented at the same spectrum levels, while in Exp. II the level of the masker was fixed and the level of the precursor was systematically varied.

As illustrated in the left column of Fig. 1, the signal and masker had total durations of 400 ms and were gated on and off together with 2-ms raised-cosine ramps. Note that this signal/masker duration was longer than those typically used in previous studies (e.g., 80 ms in Viemeister, 1980; 5 ms in Carlyon, 1989; 62 ms in Wright *et al.*, 1993). The long signal and masker durations were used here to allow the transitions to the masking pattern measurements (see Sec. II A 2). The precursor had a duration of 500 ms and was also gated with 2-ms onset and offset ramps. The temporal gap between the precursor and masker,  $T$ , was the duration from the end of the offset ramp of the precursor to the beginning of the onset ramp of the masker. Thus, when  $T=0$  ms, the drop of the stimulus amplitude between the precursor and masker had an equivalent duration of 2 ms, which was below the gap detection threshold at 2 kHz measured using band-pass noises with bandwidths comparable to those of the stimuli used here (Shailer and Moore, 1983). The amount of enhancement in the current study was defined as the decrement in signal detection threshold for a short gap duration ( $<500$  ms) compared to a gap duration of 500 ms. This definition was used because the focus of the current study was the short-term ( $<500$  ms) component of the enhancement phenomenon. In Exp. I, the amount of enhancement was assessed at various precursor-masker gap durations. In Exp. II, the dependence of the enhancement effect on precursor level was investigated.

## 2. Masking pattern measurements

In addition to enhancement measurements, masking patterns of the signal band were also obtained. In acquiring these masking patterns, detection thresholds of a brief tonal probe were measured in the presence of the stimuli from the corresponding enhancement measurement which consisted of a signal, a masker, and a precursor. The tonal probe was 6 ms in duration including 2-ms raised-cosine onset/offset ramps. Note that the signal and precursor in the enhancement measurements may be described as components of the “total masker” when measuring the masking patterns. However, we use the same terminology to be consistent across the two types of measurements (see Fig. 1). The masker was presented at a spectrum level of 30 dB SPL (sound pressure level), and the signal level was set to be 5 dB below the masker level. This

signal level was approximately the average threshold of the signal across all precursor manipulations in the enhancement measurements determined in pilot measurements.

The center frequency of the signal  $f_c$ , which was also the probe frequency, was randomized following the same procedure as in the enhancement measurements on a trial-by-trial basis. To obtain masking patterns, the probe was presented at different probe delays. A probe delay was the delay of the temporal onset of the probe relative to that of the signal. A positive probe delay means the signal onset led the probe onset, whereas a negative probe delay means the probe was gated on before the signal. The probe delays ranged from  $-50$  to 200 ms. As in the enhancement measurements, the signal and masker durations were 400 ms. This long duration was implemented to ensure that at the longest probe delay (200 ms) the probe was placed at the steady-state region of the signal’s masking pattern (e.g., Formby *et al.*, 2000). Example probe positions are schematized in the right panel of Fig. 1, with a single bold line indicating one possible temporal placement of the probe (at a probe delay of 0 ms) and gray lines indicating examples of other possible positions. The resulting probe thresholds formed a masking pattern covering a time period from shortly before the signal onset to the temporal center of the signal.

## B. Procedure

Thresholds for both enhancement and masking pattern measurements were estimated using a two-interval, two-alternative forced-choice procedure combined with a 2-down, 1-up tracking algorithm, which estimated thresholds at the 70.7% correct point on the psychometric function (Levitt, 1971). The inter-stimulus interval between the two presentation intervals on each trial was 500 ms. For the enhancement measurements, the adaptive tracking algorithm was based on the spectrum level of the signal. At the beginning of each track, the initial signal spectrum level was 40 dB SPL, and the level was decreased after two consecutive correct responses and increased after a single incorrect response. For the masking pattern measurements, the initial probe level was 85 dB SPL. The step size for the level increment or decrement was 8 dB for the first two reversals. It was reduced to 5 dB after the first two reversals, and ultimately reduced to 2 dB after the first four reversals. Each track terminated after a total of ten reversals. The threshold was estimated as the average of the signal levels over the last six reversals.

All stimuli were generated digitally at a sampling frequency of 44100 Hz on a personal computer, which also controlled the experimental procedure and data collection through custom written software in Matlab (The MathWorks, Inc., Natick, MA). Stimuli were presented monaurally to the listeners through the PC’s 24-bit soundcard (Envy24 PCI audio controller, VIA Technologies, Inc., Fremont, CA), a programmable attenuator (PA, Tucker-Davis Technologies, Inc., Alachua, FL), and a Sennheiser HD410 SL headphone (Old Lyme, CT). The left ear of each listener was tested. Each stimulus presentation was followed by visual feedback as to the correct response. The experiment was conducted in a double-walled, sound-attenuating booth.

### C. Subjects

Six listeners with normal hearing, including the first author (S4), participated. Listeners S1–S4 participated in Exp. I. Due to limited availability of S1 and S2, these two listeners did not participate in Exp. II. For this latter experiment, two more listeners S5 and S6 were recruited. All listeners were between the ages of 18 and 30 and had audiometric thresholds at or better than 15 dB HL (hearing level) between 250 and 8000 Hz in both ears. Listeners were paid for their participation except for the author. The experiments were conducted in 2-h sessions. No listener participated in more than one session per day. Before data collection began, listeners practiced for at least 4 h, which included at least one repetition of all experimental conditions in the two experiments presented here.

### D. Model implementation

A phenomenological model, consisting of a peripheral processing stage, a wideband inhibition mechanism, and a decision device, was implemented to compare the psychophysical results with model predictions. The model closely followed the one proposed by Nelson and Young (2010b), which was inspired by the adaptation-of-inhibition hypothesis of the enhancement phenomenon.

The peripheral stage of the model included 16 modeled auditory nerve fibers. These fibers had characteristic frequencies from one octave below to one-half of an octave above 2 kHz with one tenth octave spacing. The auditory nerve responses to acoustic inputs were simulated using the model of Zilany *et al.* (2009). This model differed from its predecessors (e.g., Zhang and Carney, 2005; Zilany and Bruce, 2006) in that it included a power-law adaptation algorithm. This algorithm realistically captured both short- and long-term adaptation properties at the level of the auditory nerve. The current implementation of the model set all fibers to have medium spontaneous rates (5 spikes per second) and healthy inner and outer hair cell functions. The synapse output from the model (Zilany *et al.*, 2009), expected firing rate as a function of time, was used as the output.

Following the peripheral stage, the outputs from the 16 model auditory nerve fibers were fed into a wideband inhibition mechanism. This implementation closely followed the one described by Nelson and Young (2010b). This processing stage combined responses from a narrowly tuned excitatory unit and a broadly tuned inhibitory unit. The excitatory unit received input only from the fiber tuned to 2 kHz. The inhibitory unit received inputs from all fibers using a triangular weighting function peaked at 2 kHz (with a maximal value of  $S_{\text{inh}}$  at 2 kHz and minimal values of 0 at 1 kHz and 2.8 kHz, where  $S_{\text{inh}}$  will be referred to as the inhibition strength). The responses from the excitatory and inhibitory units were smoothed using  $\alpha$  functions (e.g., Nelson and Carney, 2004) with a time constant of 4 ms. The inhibitory response was then passed through a thresholding device and subtracted from the excitatory response, producing a final model output. The thresholding device produced an output response of zero when the input was below a threshold  $\phi$ , otherwise the output was the input minus  $\phi$ . The implementation of the

thresholding device was added to the model of Nelson and Young (2010b) in an effort to match the predictions and psychophysical data with an assumption that the inhibitory unit had a higher threshold for spike generation than the excitatory unit.

For all model simulations, the spectral center of the precursor, masker, and signal was fixed at 2 kHz. In order to simulate behavioral thresholds, the model outputs were passed to a decision device to calculate a decision variable. Stimuli in both signal and no-signal intervals from 50 independent trials were presented to the model. For each interval, the response at the output of the model was segmented into 20-ms bins. For the  $i$ th bin, a parametric  $d'$ -like decision variable  $D_i$  was calculated as

$$D_i = \frac{\mu_{\text{sig},i} - \mu_{\text{nsig},i}}{\sqrt{(\sigma_{\text{sig},i}^2 + \sigma_{\text{nsig},i}^2)/2 + \sigma_0^2}}, \quad (1)$$

where  $\mu_{\text{sig},i}$  and  $\mu_{\text{nsig},i}$  are the grand average responses within the  $i$ th bin and across all stimulus repetitions;  $\sigma_{\text{sig},i}$  and  $\sigma_{\text{nsig},i}$  are the standard deviations of the time-averaged response within the  $i$ th bin across all repetitions. The subscripts “sig” and “nsig” refer to the model outputs for the signal and no-signal intervals, respectively. The value of the standard deviation of the assumed neural internal noise,  $\sigma_0$ , limits the magnitude of  $D_i$ .

For the enhancement measurements, the decision variables in various bins were combined using a strategy that assumed the listener relied on just one of the bins at the beginning of the signal for detection. In particular, the expected final decision variable  $D$  was the maximum of the  $D_i$  values in the three bins spanning from 0 to 60 ms relative to the onset of the signal. For the masking pattern measurements, the final decision variable  $D$  was the maximum of the  $D_i$  values in the three consecutive bins following the onset of the probe.

The values of  $D$  were calculated for a range of signal levels (20–40 dB SPL for the enhancement measurements, and 55–85 dB SPL for the masking pattern measurements) for each experimental condition. The resulting values of  $D$  as a function of signal level were fitted using a linear least-squares fit. The signal level that led to a  $D$  value of  $K$  was defined as the threshold, where  $K$  was the criterion for the threshold determination.

The model had four free parameters: the inhibition strength  $S_{\text{inh}}$ , the threshold for the inhibitory unit  $\phi$ , the internal noise  $\sigma_0$ , and the threshold criterion  $K$ . These parameters were adjusted to achieve threshold predictions that were visually close to the measured data. For all simulation data reported below,  $S_{\text{inh}}$  was 0.18,  $\phi$  was 80,  $\sigma_0$  was 20, and  $K$  was 0.6.

## III. EXPERIMENT I: EFFECT OF PRECURSOR-MASKER GAP

### A. Method

Both enhancement and masking patterns were measured at several precursor-masker gaps ( $T=0, 10, 50, 100, 200,$  and 500 ms for the enhancement measurements and  $T=0,$

200, and 500 ms for the masking pattern measurements). The amount of enhancement was defined as the amount of threshold decrement from the reference condition ( $T = 500$  ms). In both enhancement and masking pattern measurements, the masker and precursor spectrum levels were fixed at 30 dB SPL. To obtain the masking patterns, probe delays of  $-50$ ,  $-10$ ,  $0$ ,  $10$ ,  $50$ ,  $100$ , and  $200$  ms were tested at each of the three  $T$  values ( $0$ ,  $200$ , and  $500$  ms).

Listeners S1–S4 participated in this experiment. Masking pattern data were collected before the enhancement measurements. A total of 84 thresholds were obtained for each listener in the masking pattern measurements (3 gap durations  $\times$  4 repetitions  $\times$  7 probe delays), which were collected in the following order: Listeners S2 and S3 ran the 0-ms gap condition first, followed by the 500-ms condition, while listeners S1 and S4 ran the 500-ms condition followed by the 0-ms condition. All listeners ran the 200-ms condition after the other two gap conditions because this condition was added later in the experiment. For each of the gap durations, the order in which the seven probe delays were tested was random. Then, this process was repeated three more times, generating four replicates of the masking patterns. The reported data were based on the average of the four replicates.

For the enhancement measurements, a total of 24 thresholds were obtained for each listener (4 repetitions  $\times$  6 gap durations). The six gap durations were initially tested in random order, and then the process was repeated three more times, each with an independent random testing sequence through the gap durations.<sup>1</sup> The mean thresholds across the four repetitions were reported.

## B. Enhancement

Figure 2 plots the results from the enhancement measurements as unfilled symbols. Detection thresholds for the narrowband noise signal are shown for the individual listeners in the left panel of Fig. 2 as a function of the precursor-masker gap duration. Because the 500-ms gap duration was treated as the reference condition, the amount of enhancement (right panel of Fig. 2) indicates the threshold decrement

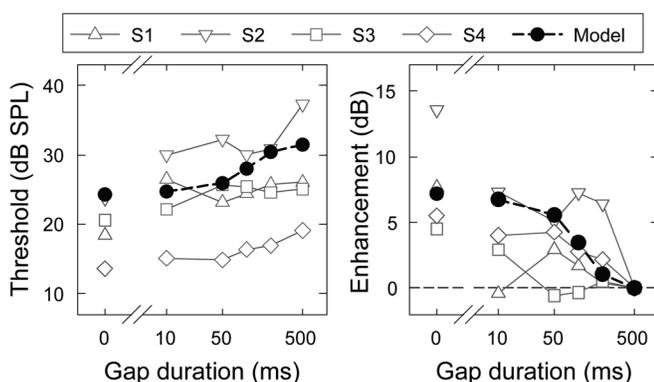


FIG. 2. The left panel plots the signal thresholds as a function of gap duration for the enhancement measurements of Exp. I. The right panel plots the amount of enhancement. Results for different listeners are plotted using different symbols. The experimental data are plotted using light unfilled symbols, while the data predicted by a model of auditory processing are shown as the dark filled symbols.

relative to the 500-ms gap condition. Although individual thresholds varied across a wide range (approximately 20 dB from the poorest to the best performer), most listeners showed an enhancement effect at shorter precursor-masker gaps. At the shortest gap duration ( $T = 0$  ms), the amounts of enhancement ranged from 4.5 to 13.6 dB. A repeated-measures analysis of variance (ANOVA) using precursor-masker gap as the within-subject factor revealed a significant main effect of gap duration on the detection threshold of the signal [ $F(5,15) = 7.64$ ,  $p < 0.001$ ].

Using a shorter signal duration (80 ms), Viemeister (1980) found an average of 11.2 dB threshold increment from the “pulsed” (long precursor-masker gaps) to the “continuous” (no precursor-masker gap) conditions for narrowband noise signals centered at 1 kHz. The amounts of enhancement measured in the present experiment were similar to those measured by Viemeister (1980), except that larger individual differences in both thresholds and amounts of enhancement were observed here. This might reflect, in part, the presentation-by-presentation randomization of the signal center frequency  $f_c$  used in our experiment.

## C. Masking pattern

Figure 3 shows the individual masking patterns (probe thresholds as functions of probe delay) for three precursor-masker gap durations (in separate panels). The signal level during the masking pattern measurement was fixed at a spectrum level of 25 dB SPL, which was close to the average signal threshold from the enhancement measurement at the 500-ms gap duration (26.9 dB).

The left panel of Fig. 3 plots the masking patterns when there was no temporal gap between the offset of the precursor and the onset of the masker. The probe thresholds in the masking pattern did not seem to systematically depend on the probe delay. For two of the listeners (S1 and S2), the threshold increased with probe delay, while this pattern of threshold was not observed for listeners S3 and S4. For all listeners, the probe threshold at the  $-10$ -ms delay was closer to that at 0 ms than at  $-50$  ms, despite the fact that the probe and the signal did not physically overlap at negative probe delays. The high thresholds at the  $-10$ -ms delay might indicate that the listeners confused the probe and signal onset and so were not efficient in attending to the probe (e.g., Hill *et al.*, 2004).

When a silent gap of 200- or 500-ms duration was present between the precursor and masker (middle and right panels in Fig. 3), the probe threshold tended to rise quickly at the signal onset, forming a peak in the masking pattern. As the probe delay increased further, the probe threshold fell to a steady-state threshold. This result was consistent with previous experiments where temporal masking patterns were measured for broadband maskers using tone pips (e.g., Zwicker, 1965) and is commonly referred to as the overshoot effect of masking (or the overshoot). The current results suggested that as the precursor-masker gap increased, an overshoot characteristic began to emerge in the masking pattern.

Because a fixed signal level (25 dB SPL) was used for these measurements, it was possible that the presence of the

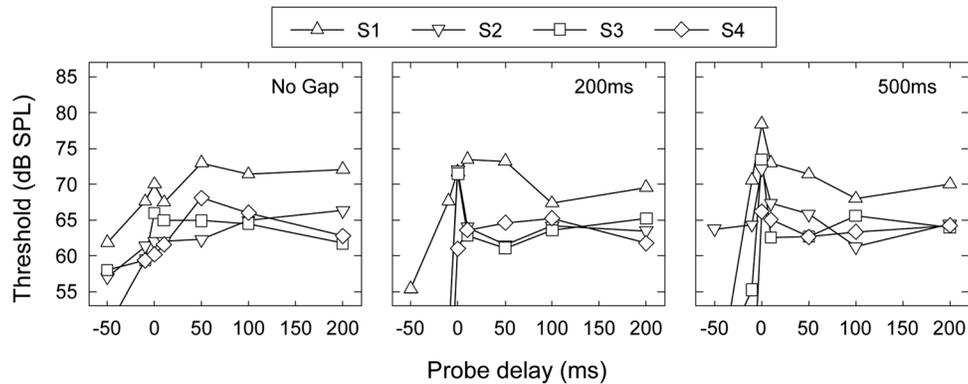


FIG. 3. Masking patterns, probe thresholds as functions of probe delay, are plotted for the individual listeners in Exp. I. The three panels, from left to right, correspond to the precursor-masker gap durations of 0, 200, and 500 ms, respectively. Note that probe thresholds at  $-50$  and  $-10$  ms probe delays were sometimes very low (they reflected the absolute threshold of the probe). To enable better clarity of the data, the range in each panel does not always include these low thresholds.

overshoot pattern reflected the fact that the signal was not audible to the listeners at larger gap durations (see the left panel of Fig. 2). However, audibility was unlikely to fully explain the differences in masking patterns. For example, listeners *S2* (downward triangles) and *S4* (diamonds) had the largest difference in signal thresholds in the enhancement measurements with all thresholds from *S2* above 25 dB SPL and all thresholds from *S4* below 25 dB SPL. However, both listeners exhibited the emergence of the overshoot characteristic with increasing gap duration in the masking pattern measurement. The probe thresholds from these two listeners were typically within 5 dB of one another.

In order to investigate whether changes in the gap duration affected the masking patterns of the signal and masker, a repeated-measures ANOVA was conducted on the probe thresholds with precursor-masker gap duration and probe delay as the within-subject factors. Because our major interest was to study the responses to the signal rather than the precursor, thresholds at negative probe delays were excluded from the analysis. The main effect of gap duration was significant [ $F(2,6) = 8.95, p = 0.016$ ], as was the interaction between gap duration and probe delay [ $F(8,24) = 4.119, p = 0.003$ ], suggesting that increasing the gap duration both decreased the probe threshold and altered the shape of the masking pattern. The main effect of probe delay [ $F(4,12) = 2.15, p = 0.137$ ] was not significant.

To further study the effect of gap duration on the shape of the masking pattern, each masking pattern for each individual listener (for non-negative probe delays) was fitted with a quadratic polynomial.<sup>2</sup> The quadratic coefficients from these regressions were indicative of the masking-pattern shape. A positive quadratic coefficient suggested that the probe threshold decreased with probe delay as a convex function, exhibiting an overshoot characteristic, while a negative quadratic coefficient suggested that the probe thresholds form a concave function, and the overshoot characteristic was absent. The resulting coefficients are listed in Table I. For all but one (*S3*) listener, the estimated quadratic coefficient increased with increasing gap durations. Taking all 12 estimates (3 gap durations  $\times$  4 listeners) into account, a significant positive correlation was found between the coefficient and the gap duration [ $r = 0.70, p < 0.05$ ]. That is,

the overshoot characteristic of the masking pattern emerged as the gap duration increased.

#### D. Model simulation

Thresholds predicted by the model in the enhancement measurements are plotted using filled circles in the left panel of Fig. 2. The predicted thresholds were within the range of the psychophysical thresholds. More importantly, as the precursor-masker gap increased from 0 to 500 ms, the predicted threshold increased by approximately 7 dB. That is, the model produced an enhancement effect. The amount of enhancement predicted by the model as a function of gap duration is plotted in the right panel of Fig. 2. The predictions are similar to the experimental results.

Figure 4 plots the masking patterns predicted by the model using filled circles. For the 0-ms condition (left panel), as the probe delay increased, the predicted thresholds tended to increase gradually, agreeing with the measured masking patterns from listeners *S1* and *S2*. For the 200- and 500-ms conditions (middle and right panels), the predicted thresholds were slightly higher than the measured thresholds. The predicted thresholds fell as the probe delay increased demonstrating strong overshoot characteristics. Therefore, the model predictions captured the trend in the experimental data in that an overshoot pattern occurred at precursor-masker durations of 200 and 500 ms, but not at 0 ms. Moreover, across the gap durations tested, differences in

TABLE I. Quadratic coefficients estimated by fitting a quadratic function to the individual masking patterns obtained in Exp. I. A positive coefficient is an indicator of a convex shape or an overshoot configuration for the masking pattern; a negative coefficient is an indicator of a concave shape or absence of overshoot. The last column on the right shows the average coefficients across listeners with standard errors indicated in the parentheses.

$T$ (ms)	Quadratic coefficient (dB SPL/s <sup>2</sup> )				Mean
	<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	
0	-194	-24	-63	-609	-223 (134)
200	198	436	529	-366	199 (201)
500	541	593	342	242	430 (83)

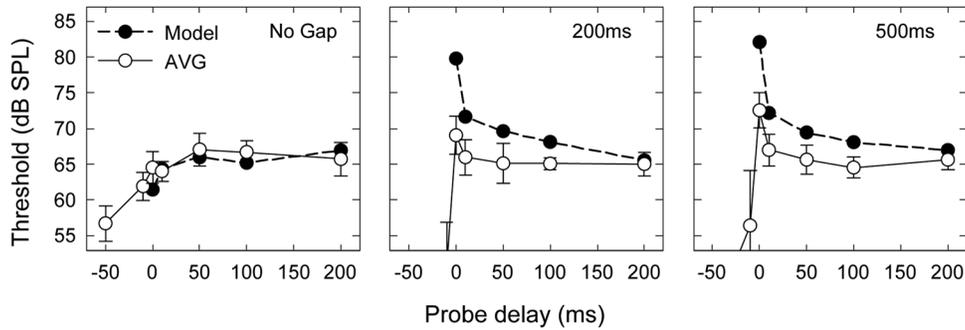


FIG. 4. Average masking patterns across listeners in Exp. I are plotted as unfilled circles. Error bars indicate  $\pm$  one standard error of the mean. The data predicted by a model of auditory processing are shown using filled symbols. Results are arranged in three panels in the same manner as Fig. 3.

thresholds were largest at the signal onset for both measured and predicted results.

The general agreement between the shapes of the measured and predicted masking patterns supports the adaptation-of-inhibition hypothesis for psychophysical enhancement. In the model predictions, the observed overshoot at the 200- and 500-ms gap durations reflected lateral inhibition at the signal onset. At short probe delays, a high probe level would be needed to overcome the inhibition, leading to higher probe thresholds compared to those measured for later probe delays. In contrast, for the 0-ms gap, the inhibition had been adapted at the signal onset, and so the probe threshold depended little on probe delay. Moreover, the model predictions also agreed with the experimental results in that when the amount of enhancement reduced the overshoot characteristic in the masking pattern seemed to emerge.

#### IV. EXPERIMENT II: EFFECT OF PRECURSOR LEVEL

##### A. Method

Experiment II measured enhancement and masking patterns as a function of the precursor level. The stimuli and procedures in this experiment were identical to those of Exp. I, except as follows. The spectrum levels of the precursor were 0, 10, 20, 30, and 40 dB SPL. For the enhancement measurements, the precursor-masker gap was either 0 or 500 ms. For the masking pattern measurements, the precursor-masker gap  $T$  was 0 ms. Probe delays of -10, 0, 10, 50, 100, and 200 ms were tested.

Listeners S3–S6 participated in this experiment. For all listeners, the masking patterns were measured first, which included a total of 120 threshold estimates (4 repetitions  $\times$  5 precursor levels  $\times$  6 probe delays). These data were collected as follows: the five precursor levels were tested in random sequences, within each of which thresholds were estimated at the six probe delays in random order. After a threshold was obtained in every condition, the process was repeated three more times. The mean thresholds across the four repetitions were reported.

The enhancement measurements included a total of 40 threshold estimates (4 repetitions  $\times$  5 precursor levels  $\times$  2 gap durations). Thresholds were estimated at the five precursor levels in random order. Within each precursor-level condition, the order in which the two gap durations were tested was also random. This process was repeated to produce four repetitions of threshold estimates in each condition. The mean thresholds across the four repetitions were reported.

##### B. Enhancement

Figure 5 plots the results from the enhancement measurements of Exp. II using light unfilled symbols. The top panels show the signal thresholds from individual listeners as a function of the precursor spectrum level in the 0-ms (left) and 500-ms (right) gap conditions. The lower panel shows the amount of enhancement, which is defined as the threshold for the 500-ms gap condition minus the threshold for the 0-ms gap condition. Little or no enhancement was measured except at the 30-dB SPL precursor level, which was also the masker level (arrow in the lower panel of Fig. 5). At this precursor level, the amounts of enhancement were comparable to those obtained from Exp. I and those reported by Viemeister (1980). The observed enhancement effect seemed to reflect the fact that the signal threshold decreased in the no-gap condition as the precursor level increased from -30 to 0 dB relative to the masker level, but then increased with the relative precursor level as it increased further to 10 dB (top left panel of Fig. 5). A repeated-measures ANOVA was performed on the threshold

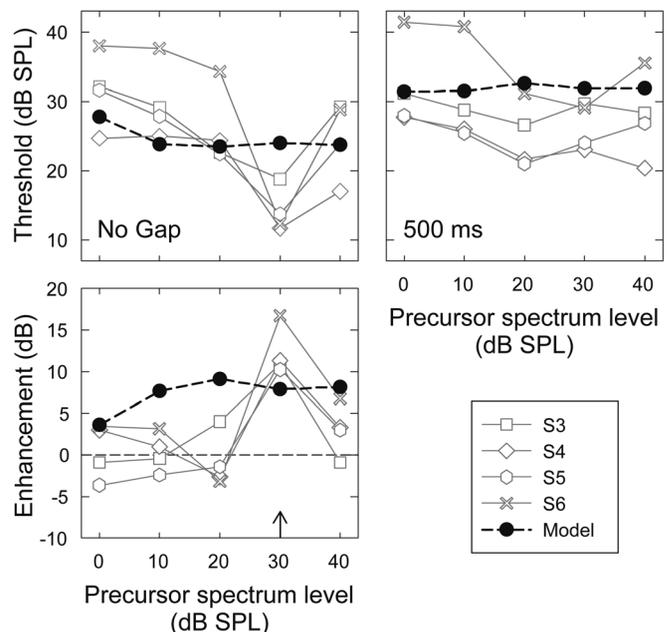


FIG. 5. The signal thresholds (upper panels) and the amount of enhancement (lower panel) measured in Exp. II as a function of precursor spectrum level. The results for individual subjects are plotted using light unfilled symbols, while the model predictions are plotted as dark filled symbols. The arrow on the abscissa indicates the spectrum level of the masker.

data treating precursor level and precursor-masker gap as the within-subject factors. The analysis revealed significant main effects of precursor level [ $F(4,12) = 10.78, p < 0.001$ ] and gap duration [ $F(1,3) = 12.08, p < 0.05$ ], and a significant interaction between these two factors [ $F(4,12) = 14.84, p < 0.001$ ]. The significant effect of gap duration indicates an enhancement effect, whereas the significant interaction suggests a dependence of the amount of enhancement on the precursor level.

The pattern of thresholds in the current experiment agrees with those measured by Bacon and Smith (1991) and Hicks and Bacon (1992). In one of their conditions, Hicks and Bacon (1992) measured the detection threshold for a pure-tone signal (at 1 or 4 kHz) in the presence of a masker and a precursor, similar to those used in the no-gap condition of the current experiment. They fixed the masker spectrum level at 40 dB SPL and varied the level of the precursor (from -40 to 10 dB relative to the masker level) and found that for the 4-kHz signal frequency, the signal threshold decreased with the precursor level up to a relative level of 0 dB, after which the threshold increased drastically. A comparable pattern was also observed for the 1-kHz signal, with a much reduced effect of precursor level on the signal threshold.

Carlyon (1989) also measured the effect of precursor level on the enhancement effect. In contrast to the data presented here and those of Bacon and Smith (1991) and Hicks and Bacon (1992), Carlyon found that the amount of enhancement was about 10 dB and did not significantly depend on precursor level when the precursor level was varied from -30 to 0 dB relative to the masker level. There were several differences in stimulus design between the current study and that of Carlyon (1989). First, the current study used longer precursor and masker durations (500 and 400 ms, respectively) compared to those used by Carlyon (200 and 5 ms, respectively). Second, in the current study the amount of enhancement was defined as the threshold difference between conditions for precursor-masker gaps of

0 and 500 ms, whereas the enhancement was assessed in terms of the threshold difference between conditions with and without the presence of the precursor in Carlyon's study. Moreover, in his study, when the precursor was present, the precursor-masker gap was 10 ms. Third, the signal in the study by Carlyon was a brief tone burst at 1 kHz, while the current experiment used narrowband noise signals centered at around 2 kHz. Last, the spectral center of the stimuli (including the precursor, masker, and signal), which was fixed in frequency in Carlyon's study, was randomly roved on a presentation-by-presentation basis in the current experiment.

Any of these differences may have contributed to the inconsistent results between the current and Carlyon's studies. For example, Hicks and Bacon (1992) showed that the dependence of enhancement on precursor level was less salient at lower signal frequencies (e.g., at 1 kHz). Therefore, the difference in signal frequency across the two studies might be the reason that the dependence of the enhancement effect on precursor level was found in the current experiment but not in the Carlyon's study.

### C. Masking pattern

Figure 6 shows the masking patterns for the five precursor spectrum levels (in separate panels). At moderate-to-high precursor levels (20–40 dB SPL), the subjects' masking patterns resemble those seen in the no-gap condition of Exp. I (e.g., see the left panel of Fig. 3): the probe threshold was relatively independent of probe delay. At the lowest precursor level (0 dB SPL), the masking patterns were overshoot-like: the threshold rose to form a peak after the signal onset, then fell to a steady-state level. However, the size of the overshoot in this condition, which is the threshold difference between the 0- and 200-ms probe delays, was smaller than the overshoot observed in Exp. I. A repeated-measure ANOVA treating precursor level and probe delay as the two within-subject factors did not reveal significant effects of precursor level

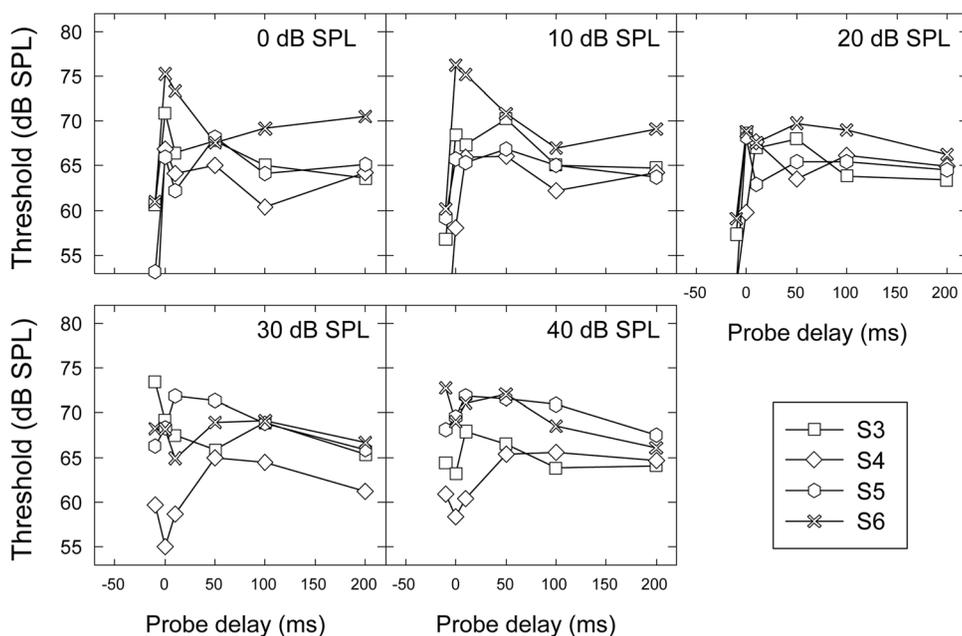


FIG. 6. Masking patterns are plotted for individual listeners in Exp. II. Data from different precursor level conditions are in separate panels.

TABLE II. Quadratic coefficients estimated by fitting a quadratic function to each of the individual masking patterns obtained in Exp. II. The last column on the right shows the average coefficients across listeners with standard errors in the parentheses.

Precursor level (dB SPL)	Quadratic coefficient (dB SPL/s <sup>2</sup> )				
	S3	S4	S5	S6	Mean
0	113	370	-84	506	226 (132)
10	-30	-178	-103	548	59 (166)
20	99	-203	30	-199	-68 (78)
30	-54	-744	-193	-243	-308 (151)
40	-37	-486	-222	-180	-231 (94)

[ $F(4,12) = 0.16, p = 0.953$ ] or probe delay [ $F(4,12) = 1.46, p = 0.274$ ], and the interaction between the two factors [ $F(16,48) = 1.64, p = 0.094$ ] did not reach significance (for  $p < 0.05$ ). Given that an interaction between probe delay and precursor-masker gap was observed in Exp. I, it seems likely that had a sufficiently low-level precursor been tested, a significant interaction would have emerged.

As in Exp. I, we explored the possibility that the shape of the masking pattern depends on the precursor level by curve-fitting a quadratic polynomial to each of the masking patterns. The resulting coefficients are listed in Table II. A negative correlation was found between the coefficient and precursor level [ $r = -0.62, p = 0.004$ ], indicating that as the precursor level increased the quadratic coefficient decreased from positive to negative values and the overshoot characteristic of the masking pattern disappeared.

#### D. Model simulation

The thresholds predicted by the model for the enhancement measurements are shown as filled circles in the upper panels of Fig. 5. The predicted amounts of enhancement are shown in the lower panel. For both the 0- and 500-ms conditions (upper left and right panels, respectively), the predicted

thresholds were approximately 30 dB SPL. In contrast to the psychophysical data, there was no systematic change in threshold with changes in precursor level except that in the no gap condition the signal threshold was slightly higher at 0 dB SPL compared to the other precursor levels.

As a result, the predicted amount of enhancement (filled circles in the lower panel of Fig. 5) increased slightly as the precursor level increased from 0 to 20 dB SPL, then reached an asymptote at about 7 dB. This prediction deviated from the experimental observation that enhancement was the largest at the 30-dB SPL precursor level and reduced at lower or higher precursor levels. At the 30-dB precursor level, the model under-predicted the amount of enhancement by about 4 dB, while at all other precursor levels, the model over-predicted the amount of enhancement. The discrepancy between the measured and predicted enhancement indicates that the model failed to fully capture the enhancement results.

Figure 7 plots the predicted masking patterns by the model as filled circles. At the lowest precursor level, the predicted threshold fell slightly as the probe delay increased from 0 to 10 ms, exhibiting a small overshoot characteristic. The predicted overshoot characteristic diminished at precursor levels higher than 0 dB SPL, where the simulated probe threshold increased slightly with increasing probe delay. This dependency of the masking-pattern shape on precursor level was seen in the measured results from listeners S3, S4, and S6 (see Fig. 6). Therefore, the model captured the masking pattern results but not the enhancement results.

### V. GENERAL DISCUSSION

#### A. Predicting the effects of precursor-masker gap and precursor level on enhancement

The experiments presented here investigated the plausibility of the adaptation-of-inhibition hypothesis of auditory enhancement. The hypothesis states that enhancement is a result of reduction of inhibition caused by the presence of a

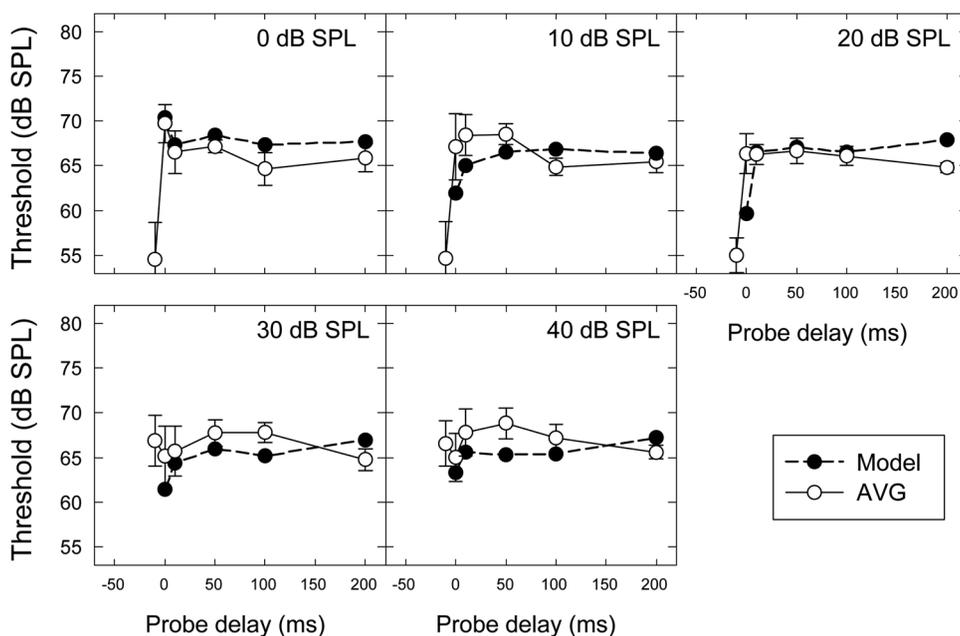


FIG. 7. Masking patterns averaged across listeners in Exp. II are plotted as unfilled circles. The filled circles indicate the predicted masking patterns for a model of auditory processing.

precursor. The hypothesis has been computationally implemented into a phenomenological model of auditory processing by Nelson and Young (2010b). Our implementation of their model produces reasonably good predictions of the signal thresholds and the amounts of enhancement measured psychophysically in Exp. I (Figs. 2 and 4).

To demonstrate how the model predicts threshold differences when the precursor-masker gap varies, Fig. 8 plots examples of model outputs as a function of time for a 500-ms gap (left) and no gap (right). First consider the model response when there is a long temporal gap. After the precursor offset, the unit returns to its spontaneous rate (marked with an arrow). Then, at the masker onset, there is a brief period of reduced activity, followed by a buildup of response. As pointed out by Nelson and Young (2010b), this build-up characteristic at the output of the model reflects the fact that the wideband inhibition mechanism is included in the model. Because the inhibitory unit in the wideband inhibition stage of the model is broadly tuned, broadband stimuli, such as the precursor or the masker, cause more rapid adaptation of the inhibitory unit compared to the excitatory unit. This difference in the rates of adaptation leads to a build-up response pattern. This class of responses has been observed in those IC neurons that also demonstrate strong physiological enhancement (Nelson and Young, 2010a,b).

When the precursor and masker are temporally contiguous (the upper right panel of Fig. 8), in contrast to the long-gap condition, a strong response to the signal following the precursor-masker boundary is observed. This strong response reflects the fact that the inhibitory unit in the wideband inhibition stage is adapted, whereas the excitatory unit is not (i.e., the adaptation-of-inhibition hypothesis of Viemeister and Bacon, 1982). Due to adaptation of inhibition, the response to the signal is effectively amplified. Therefore, as illustrated in the lower panels of Fig. 8, the detectability of the signal would be enhanced in the 0-ms gap condition compared to the 500-ms gap condition. Notably, this model predicts enhanced response primarily at the signal onset.

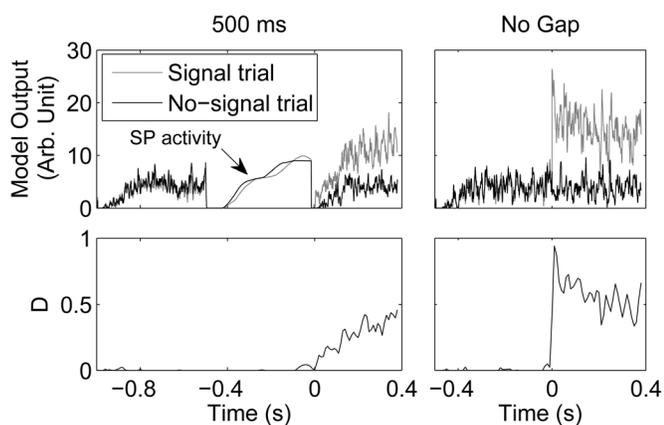


FIG. 8. Examples of model responses to the stimuli (upper panels) and the decision variable  $D$  (lower panels) in Exp. I for a signal level of 25 dB SPL. Results are shown for precursor-masker gaps of 0 and 500 ms in the left and right columns, respectively. In the upper panels, the dark curve plots the model's response to the masker alone, while the gray curve plots the model's response to the signal plus masker. Zero on the abscissa marks the onset of the masker and signal.

For Exp. II, the model is less successful in capturing the results of the enhancement measurements. One feature of the results in the current experiment is that enhancement is observed to only occur when the precursor and masker levels are equal. This result is not evident in the model predictions.

This discrepancy between the model and experimental results is, at least in part, related to the fact that the model predicted little dependence of the enhancement phenomenon on the precursor level, especially for precursor levels above 20 dB SPL. The amount of enhancement was predicted to be about 7 dB at all precursor levels except for 0 dB SPL. The lack of a precursor-level effect in the modeling results might be explained in terms of the balance between the excitatory and inhibitory responses near the signal onset. The left panel of Fig. 9 plots the average onset responses at the outputs of the excitatory and inhibitory units in the model to the stimuli of Exp. II (results shown for the signal interval only with the signal level fixed at 25 dB SPL). When a 500-ms precursor-masker gap was present ("Gap" on the abscissa), the inhibition was stronger than the excitation at the signal onset. This was similar to the results at low precursor levels (-20 and -10 dB SPL). As the precursor level increased, both the excitatory and inhibitory responses decreased, indicating a suppressive effect of the precursor. However, the inhibitory response exhibited a steeper drop with increasing precursor level than the excitatory response for precursor levels below 20 dB SPL. As a result, the inhibitory response became lower than the excitatory response and the difference remained relatively stable for precursor levels above 20 dB SPL (see the right panel of Fig. 9).

Because the final model output combines the excitatory and inhibitory responses, results shown in Fig. 9 suggest that the combined response is insensitive to changes in precursor level for precursor levels above 20 dB SPL. This might have led to the lack of dependence of the predicted enhancement phenomenon on precursor level for Exp. II. A number of modifications to the current model might lead to improved

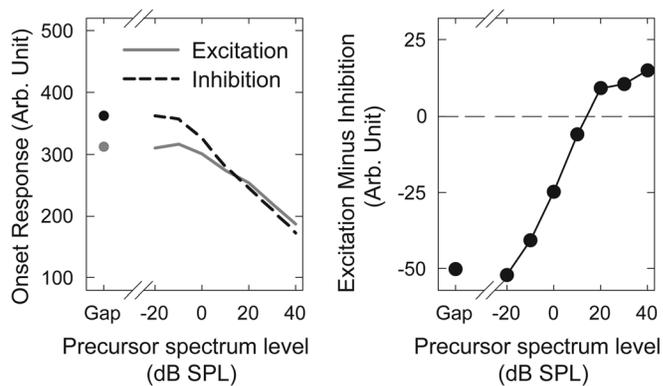


FIG. 9. Model responses near the signal onset. Left: the average responses within a 10-ms window following the signal onset and across 20 stimulus repetitions. The responses at the outputs of the excitatory (gray and solid) and inhibitory (dark and dashed) units in the model are shown for various precursor levels and for the case where a 500-ms precursor-masker gap is present (marked as "Gap"). For the generation of these results, the stimuli were identical to those used in Exp. II. The masker and signal spectrum level were fixed at 30 and 25 dB SPL, and the precursor level was 30 dB SPL in the "Gap" condition. Right: The difference between the excitatory and inhibitory responses shown in the left panel.

predictions. For example, the outer hair cell gain might be manipulated to account for the dependence of peripheral suppression on precursor level, auditory nerve fibers with low-spontaneous rates might be implemented in the model instead of mid-spontaneous fibers, or the threshold for the inhibitory unit  $\phi$  might be adjusted. Without systematically studying the effects of these parameter manipulations, it is not clear whether the model will capture the non-monotonic dependence of the enhancement phenomenon on precursor level found in Exp. II. Therefore, the discrepancy between the model and experimental results of Exp. II does not provide conclusive evidence against the adaptation-of-inhibition hypothesis.

## B. Temporal dynamics of the enhancement phenomenon

If the adaptation-of-inhibition hypothesis is valid, it not only predicts changes in the signal threshold, but also different amounts of enhancement across the duration of the signal. Predictions regarding the temporal dynamics of the enhancement effect are illustrated in Fig. 8. When the precursor immediately precedes the masker (lower right panel), the detectability of the signal,  $D$  [Eq. (1)], is high at the signal onset and falls over time. On the other hand, when a 500-ms gap is present (lower left panel), the responses for both the signal and no-signal trials of the enhancement measurement (light and dark curves in the upper left panel, respectively) fall below spontaneous activities (marked by the arrow) at the onset of the masker. This reflects strong inhibition from the spectral region surrounding the signal. Consequently, the detectability  $D$  is low near the signal onset and recovers over time.

The masking patterns presented here used a brief tone pip presented at the center frequency of the narrowband signal and at various delays relative to the signal onset to probe the temporal dynamics of the enhancement. Based on the model predictions shown in Fig. 8, high probe thresholds are expected near the signal onset when the precursor-masker gap is long because high probe levels would be needed to bring the response to the probe above spontaneous activity. In other words, the model predicts an overshoot characteristic in the masking pattern for long, but not short, precursor-masker gaps. Therefore, a test of the model would be to investigate whether or not the overshoot characteristic is observed in the corresponding conditions, and whether the overshoot pattern emerges as the amount of enhancement reduces.

The experiments presented here suggest that moving the precursor temporally away from the masker or decreasing the precursor level had a similar effect on the masking patterns at the signal frequency: the emergence of overshoot characteristics. When the overshoot pattern was robust, an enhancement effect in the corresponding condition was not observed in our data. On the other hand, in conditions where an enhancement was readily observed, the probe threshold was relatively independent of probe delay. Therefore, the results from the current experiments generally agree with the adaptation-of-inhibition hypothesis of enhancement (e.g., Viemeister and Bacon, 1982).

## C. Relations to studies of the overshoot phenomenon

In the current experiments, the precursor's effects on the amount of overshoot observed in the masking pattern measurements are in good agreement with previous studies of the overshoot phenomenon, which have consistently shown that a precursor's presence can reduce the overshoot effect (e.g., McFadden, 1989; Overson *et al.*, 1996; Bacon and Liu, 2000; Strickland, 2008). For example, Overson *et al.* (1996) measured the detection threshold for a 10-ms, 4-kHz tone probe presented with a simultaneous masker and a precursor. Both the masker and precursor were broadband noises with durations of 400 ms. The probe was placed either at the masker onset or 195 ms after the masker onset. The overshoot effect, quantified as the difference in threshold at the two probe delays, was investigated as a function of the precursor-masker gap duration. They found that the amount of overshoot decreased ( $\sim 10$  dB) as the precursor was moved closer to the masker (from 400- to 1-ms gap duration). A similar trend was observed in Exp. I (see Fig. 3). Using notched noise maskers and precursors similar to those used in the present experiments, Strickland (2008) measured the detection of tone probes at the masker onset as a function of precursor level (with no temporal gap between the precursor and masker). She found that the overshoot effect was diminished as the precursor level increased. A similar result was observed in Exp. II, where the probe threshold at the signal onset decreased with increasing precursor levels.

The current implementation of the model of Nelson and Young (2010b) produced an overshoot effect suggesting a potential common origin of the overshoot and enhancement phenomena. It is worth pointing out that although it is possible that the enhancement and overshoot phenomena are related, the experimental data from psychophysical studies of the two phenomena do not always agree. For example, frequency regions above the signal frequency are known to be of more importance to overshoot compared to low-frequency regions (e.g., McFadden, 1989). In contrast, frequency regions lower than the signal frequency seem to contribute to the enhancement effect more than frequency regions above the signal frequency (e.g., Carlyon, 1989). Broadly tuned inhibitory feedback mechanisms have been proposed previously to explain the overshoot effect (e.g., McFadden, 1989). Many recent studies suggest that the olivocochlear efferent system might be a plausible candidate for such a mechanism (e.g., von Klitzing and Kohlrausch, 1994; Strickland, 2001, 2004; Walsh *et al.*, 2010; Jennings *et al.*, 2011). Others point to the possibilities of a more central location in the auditory pathway for the feedback mechanism (e.g., Keefe *et al.*, 2009). The adaptation-of-inhibition hypothesis is functionally similar to these proposed explanations for the overshoot effect. However, whether the same neural circuitry underlies both the enhancement and overshoot phenomena requires further physiological and psychophysical investigations.

## VI. SUMMARY

Using narrowband noise signals and notched noise maskers and precursors, the current study measured the

precursor's effects on the amount of signal enhancement and the temporal masking patterns for the signal and the masker. An overshoot characteristic of the masking pattern emerged as the precursor-masker gap duration increased. In conditions where enhancement was measured, the overshoot characteristic was weakened, or not present, in the masking pattern. The adaptation-of-inhibition hypothesis of the enhancement phenomenon derived from Viemeister and Bacon (1982), and computationally implemented using a model proposed by Nelson and Young (2010b), accounted for much of the data from the enhancement and masking pattern experiments described here. However, the adaptation-of-inhibition hypothesis alone did not explain the non-monotonic dependence of the enhancement effect on the precursor level observed in Exp. II.

## ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R21 DC010058. The authors would like to thank two anonymous reviewers for their insightful comments and Eva Maria Carreira for her assistance in data collection.

<sup>1</sup>Note that a potential problem of the test sequence used in the current study was that the gap durations were tested in random order within a block for the enhancement measurements, while the gap durations were fixed within a block and the probe frequencies were randomized. This might mean that within a block, listeners may have been affected by stimulus uncertainty along different dimensions for the two measurements.

<sup>2</sup>Several types of functions were used to fit the masking pattern (between 0- and 200-ms probe delays), including linear, quadratic, and exponential functions. Among these functions, the quadratic polynomial provided the best overall fit to the data, hence it was used in further analyses as a method of quantifying the shape of masking patterns.

- Bacon, S. P., and Liu, L. (2000). "Effects of ipsilateral and contralateral precursors on overshoot," *J. Acoust. Soc. Am.* **108**, 1811–1818.
- Bacon, S. P., and Smith, M. A. (1991). "Spectral, intensive, and temporal factors influencing overshoot," *Q. J. Exp. Psychol. A* **43**, 373–399.
- Carlyon, R. P. (1989). "Changes in the masked thresholds of brief tones produced by prior bursts of noise," *Hear. Res.* **41**, 223–235.
- Fastl, H. (1977). "Subjective duration and temporal masking patterns of broadband noise impulses," *J. Acoust. Soc. Am.* **61**, 162–168.
- Formby, C., Sherlock, L. P., and Ferguson, S. H. (2000). "Enhancements of the edges of temporal masking functions by complex patterns of overshoot and undershoot," *J. Acoust. Soc. Am.* **107**, 2169–2187.
- Heinz, M. G., Zhang, X., Bruce, I., and Carney, L. H. (2001). "Auditory nerve model for predicting performance limits of normal and impaired listeners," *ARLO* **2**, 91–96.
- Hicks, M. L., and Bacon, S. P. (1992). "Factors influencing temporal effects with notched-noise maskers," *Hear. Res.* **64**, 123–132.
- Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., and Moore, D. R. (2004). "Auditory processing efficiency and temporal resolution in children and adults," *J. Speech Lang. Hear. Res.* **47**, 1022–1029.
- Jennings, S. G., Heinz, M. G., and Strickland, E. A. (2011). "Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot," *J. Assoc. Res. Otolaryngol.* **12**, 345–360.
- Keefe, D. H., Schairer, K. S., Ellison, J. C., Fitzpatrick, D. F., and Jesteadt, W. (2009). "Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot," *J. Acoust. Soc. Am.* **125**, 1595–1604.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**(2B), 467–477.
- McFadden, D. (1989). "Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot," *J. Acoust. Soc. Am.* **85**, 254–261.
- 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 Young, E. D. (2010a). "Enhancements in the marmoset inferior colliculus: Neural correlates of perceptual 'pop-out,'" in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), Chap. 15, pp. 155–165.
- Nelson, P. C., and Young, E. D. (2010b). "Neural correlates of context-dependent perceptual enhancement in the inferior colliculus," *J. Neurosci.* **30**, 6577–6587.
- Overson, G. J., Bacon, S. P., and Webb, T. M. (1996). "The effect of level and relative frequency region on the recovery of overshoot," *J. Acoust. Soc. Am.* **99**, 1059–1065.
- Palmer, A. R., Summerfield, Q., and Fantini, D. A. (1995). "Responses of auditory-nerve fibers to stimuli producing psychophysical enhancement," *J. Acoust. Soc. Am.* **97**, 1786–1799.
- Richards, V. M., Huang, R., and Kidd, G. (2004). "Masker-first advantage for cues in informational masking," *J. Acoust. Soc. Am.* **116**, 2278–2288.
- Richards, V. M., and Neff, D. L. (2004). "Cuing effects for informational masking," *J. Acoust. Soc. Am.* **115**, 289–300.
- Shailer, M. J., and Moore, B. C. J. (1983). "Gap detection as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 467–473.
- Strickland, E. A. (2001). "The relationship between frequency selectivity and overshoot," *J. Acoust. Soc. Am.* **109**, 2062–2073.
- Strickland, E. A. (2004). "The temporal effect with notched-noise maskers: Analysis in terms of input-output functions," *J. Acoust. Soc. Am.* **115**, 2234–2245.
- Strickland, E. A. (2008). "The relationship between precursor level and the temporal effect," *J. Acoust. Soc. Am.* **123**, 946–954.
- Summerfield, Q., and Assmann, P. F. (1989). "Auditory enhancement and the perception of concurrent vowels," *Percept. Psychophys.* **45**, 529–536.
- Summerfield, Q., Sidwell, A., and Nelson, T. (1987). "Auditory enhancement of changes in spectral amplitude," *J. Acoust. Soc. Am.* **81**, 700–708.
- Viemeister, N. F. (1980). "Adaptation of masking," in *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft University Press, Delft, The Netherlands), pp. 190–198.
- Viemeister, N. F., and Bacon, S. P. (1982). "Forward masking by enhanced components in harmonic complexes," *J. Acoust. Soc. Am.* **71**, 1502–1507.
- von Klitzing, R., and Kohlrausch, A. (1994). "Effect of masker level on overshoot in running- and frozen-noise maskers," *J. Acoust. Soc. Am.* **95**, 2192–2201.
- Walsh, K. P., Pasanen, E. G., and McFadden, D. (2010). "Overshoot measured physiologically and psychophysically in the same human ears," *Hear. Res.* **268**, 22–37.
- Wang, N., Kreft, H., and Oxenham, A. J. (2012). "Vowel enhancement effects in cochlear-implant users," *J. Acoust. Soc. Am.* **131**, EL421–EL426.
- Wright, B. A., McFadden, D., and Champlin, C. A. (1993). "Adaptation of suppression as an explanation of enhancement effects," *J. Acoust. Soc. Am.* **94**, 72–82.
- Zhang, X., and Carney, L. H. (2005). "Analysis of models for the synapse between the inner hair cell and the auditory nerve," *J. Acoust. Soc. Am.* **118**, 1540–1553.
- Zilany, M. S. A., and Bruce, I. C. (2006). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J. Acoust. Soc. Am.* **120**, 1446–1466.
- 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.
- Zwicker, E. (1965). "Temporal effects in simultaneous masking by white-noise bursts," *J. Acoust. Soc. Am.* **37**, 653–663.