New perspectives on the measurement and time course of auditory enhancement

Lei Feng[1] and **Andrew J. Oxenham**[1,2]

[1]Department of Otolaryngology, University of Minnesota, Minneapolis, Minnesota, 55455

[2]Department of Psychology, University of Minnesota, Minneapolis, Minnesota, 55455

Abstract

A target sound can become more audible and may 'pop out' from a simultaneously presented masker if the masker is presented first by itself, as a precursor. This phenomenon, known as auditory enhancement, may reflect the general perceptual principle of contrast enhancement, which facilitates adaptation to ongoing acoustic conditions and the detection of new events. Little is known about the mechanisms underlying enhancement, and potential confounding factors have made the size of the effect and its time course a point of contention. Here we measured enhancement as a function of precursor duration and delay between precursor offset and target onset, using two single-interval pitch comparison tasks, which involve either same-different or up-down judgments, to avoid the potential confounds of earlier studies. Although these two tasks elicit different levels of performance and may reflect different underlying mechanisms, they produced similar amounts of enhancement. The effect decreased with decreasing precursor duration, but remained present for precursors as short as 62.5 milliseconds, and decreased with increasing gap between the precursor and target, but remained measurable 1 second after the precursor. Additional conditions, examining the effect of precursor/masker similarity and the possible role of grouping and cueing, suggest multiple sources of auditory enhancement.

Keywords

Auditory perception; Contrast enhancement; Perceptual invariance

---

Our perception of a sound can be strongly affected by the context in which it is presented. For example, if a complex tone is presented twice and an additional component is added on the second presentation, the added component will often be more perceptually salient than the other components, and more salient than it would have been without the first presentation of the complex tone. This effect, known as auditory enhancement, has been demonstrated in a variety of ways across different studies. Enhancement can decrease the threshold of detectability of the added "target" component (Viemeister, 1980), and it can result in the target component producing more forward masking following its offset (Byrne, Stellmack, & Viemeister, 2011; Viemeister & Bacon, 1982). Supra-threshold effects have also been noted; for instance, enhancement can increase the effective level of a target tone monaurally, such that it affects the lateralized percept produced by combining the target with a contralateral tone at the same frequency and phase (Byrne et al., 2011). In addition, the enhanced target tone can be salient enough to be perceived as a separate entity with a

distinct pitch (Byrne, Stellmack, & Viemeister, 2013; Demany, Carcagno, & Semal, 2013; Erviti, Semal, & Demany, 2011; Hartmann & Goupell, 2006). Similar effects have been observed with artificial speech stimuli, where spectral gaps in a broadband harmonic precursor at frequencies corresponding to the first three formants of a vowel can lead to the enhancement of those frequencies, and the perception of the vowel, in a subsequent harmonic complex with a flat spectral envelope (Summerfield, Haggard, Foster, & Gray, 1984; Summerfield, Sidwell, & Nelson, 1987; Thibodeau, 1991; Wang, Kreft, & Oxenham, 2012).

Enhancement effects have been typically studied with harmonic or inharmonic complex tones, but similar effects can be observed using noise bands as the precursor and/or the masker (Summerfield et al., 1987; Viemeister, Byrne, & Stellmack, 2013). All these context effects may reflect the auditory system's ability to adapt and to normalize or "whiten" the representation of sound to improve coding efficiency (Barlow, 1961; Dean, Harper, & McAlpine, 2005) and to sensitize the system to new events or important changes in the stimulus (Stilp, Alexander, Kiefte, & Kluender, 2010). In addition, these effects may contribute to perceptual invariance, or our ability to recognize the same stimuli, and understand speech, over a wide range of acoustic conditions, including different talkers, different rooms, and background noise. Despite these large effects, and their clear importance for auditory and speech perception, surprisingly little is known about their underlying mechanisms. Some psychophysical studies have proposed a combination of neural adaptation and lateral inhibition or suppression as an explanation of the enhancement effects (Byrne et al., 2011; Carcagno, Semal, & Demany, 2012; Viemeister, 1980; Viemeister & Bacon, 1982). Other mechanisms might also play a role. For instance, the precursor and masker could be perceptually grouped together because of their shared spectral characteristics, so that the target tone stands out when it is introduced (Carlyon, 1989), as an instantiation of the "old plus new" heuristic described by Bregman (1990, p. 222). Alternatively, or in addition, the precursor could provide other cues that help listeners to segregate the target from the subsequent masker (Richards, Huang, & Kidd, 2004). Neurophysiological studies have found adaptation of maskers at the level of auditory nerve (Palmer, Summerfield, & Fantini, 1995) and facilitated responses to a target tone in the cochlear nucleus of guinea pigs and the inferior colliculus of marmoset monkeys (Nelson & Young, 2010; Scutt & Palmer, 1998) in response to stimuli that elicit auditory enhancement in humans. Based on these studies it seems likely that enhancement is generated at different levels and accumulates along the auditory pathway. Hearing-impaired listeners and cochlear-implant users have shown reduced or absent enhancement effects, which might be related to their poorer speech understanding in noisy environment (Carlyon, Long, Deeks, & McKay, 2007; Thibodeau, 1991, 1996; Wang et al., 2012). Better characterization and understanding of auditory enhancement might eventually lead to new signal processing algorithms in hearing aids and cochlear implants that could restore normal enhancement effects for people using these devices.

Auditory enhancement is often quantified as the difference between detection thresholds measured in unenhanced (no precursor) and enhanced (with precursor) conditions. In such cases, thresholds are often obtained using a two-interval forced-choice task, where the masker and precursor are presented in both intervals and the target is presented in only one.

There are at least two ways in which threshold differences measured in this way could deviate from the underlying amount of enhancement. First, the masker without the signal present could act as a precursor for the masker with the signal present, and could thus lead to *underestimates* of enhancement. Second, long-term enhancement effects may play a role (Carlyon, 1989; Viemeister, 1980), so that enhancement effects build up over multiple trials, leading to potential *overestimates* of the enhancement produced in a single trial.

A recent method used by Demany, Carcagno and colleagues (Carcagno et al., 2012; Carcagno, Semal, & Demany, 2013) overcomes these issues by employing a one-interval task with frequency components that are selected randomly from trial to trial. Their method involves pitch comparisons between a target presented within the masker and a comparison tone presented afterwards. However, in their paradigm the enhancement effect was measured in terms of changes in sensitivity (d'), rather than the effective level of the stimulus, making direct comparisons with earlier studies of auditory enhancement difficult, and leaving open the question of the amount of enhancement in terms of effective level change. Byrne et al. (2013) also used a one-interval task and roved frequencies in a paradigm that involved participants judging whether the target tone "stood out" from the remainder of the complex. By repeating the task at multiple target levels, relative to the surrounding masker tones, they were able to calculate the effective amount of enhancement produced by a precursor. In general, the effects were found to be larger than expected based on earlier studies: rather than enhancement of 5–10 dB, they found that the enhancement was equivalent to an increase in target level of about 20 dB. It is possible that this outcome represents the "true" amount of enhancement when potential confounds, such as using the same target frequency throughout a trial, are eliminated. However, the authors argued instead that "informational masking" (Durlach et al., 2003; Oxenham, Fligor, Mason, & Kidd, 2003), produced by the randomization of the target and masker frequencies, may have played a role. In addition, their task was subjective and was not a direct measure of performance, making comparisons with earlier detection studies difficult.

The aim of our study was two-fold. The first aim was to develop a one-interval task with roved frequency that avoids the potential confounds outlined above and that provides a performance-based measure of the effective amount of enhancement (in dB). The second aim was to use this task to explore the time course of enhancement, in terms of both the duration of the precursor and the gap between the precursor and the target. We used a one-interval pitch comparison task, similar to the one described by Demany, Carcagno and colleagues (Carcagno et al., 2012; Demany & Ramos, 2005) to measure enhancement as a function of precursor duration and delay between precursor offset and target onset. As shown in Figure 1, each trial consisted of a complex tone, comprising masker components and a single target component, with the frequencies roved from trial to trial. The complex tone was followed by a single pure-tone probe, and the participants' task was to compare the frequency of the probe with that of the target tone. In the enhanced condition (Figure 1), a precursor, comprising masker components only, was presented before the complex tone and probe. In the no-precursor (control) condition, there were only two sounds, the complex tone and the probe. To estimate the effective amount of enhancement, performance in this task was measured by adaptively varying the level of the target component and probe, relative to

the (fixed-level) masker components. In such a condition with large uncertainty of masker and target frequencies, the precursor can act as a cue which decreases masker uncertainty to reduce informational masking (Richards et al., 2004; Richards & Neff, 2004). To attempt to control for the effects of informational masking, we used two different tasks. The first was termed the Present/Absent task: participants were asked whether the probe tone was present in the complex (identical to the target) or absent (centered in frequency between the target and one of the adjacent masker components) (Figure 1a). The second was termed the Up/ Down task: participants were asked to judge the direction of pitch shift between the target and the probe tone, which was always either one semitone higher or lower than the target (Figure 1b). It has been proposed that automatic "frequency shift detectors" in the auditory system (Demany & Ramos, 2005; Okada & Kashino, 2003) provide us with the ability to detect small frequency shifts and identify the direction of frequency change in a way that depends less on attention than the more traditional Present/Absent task, in which the target tone has to be "heard out" from the rest of the complex (Demany, Clément, & Semal, 2001). Given that informational masking can be viewed as a failure of selective attention, the Up/ Down task should be less affected by informational masking because the detection of frequency shift is thought to be "automatic" and thus to require less attention. As a consequence, we would expect less enhancement in the Up/Down task compared to the Present/Absent task if enhancement under these conditions reflects a reduction of informational masking.

Experiment 1

Methods

Participants—Fourteen normal-hearing (NH) participants (nine females), including author LF, were tested. The participants were between 18 and 31 years old and all had absolute pure-tone thresholds below 20 dB SPL in both ears at octave frequencies from 250 to 8000 Hz. All participants provided written informed consent and, with the exception of author LF, were compensated for their time.

Stimuli—In the main experiment, each trial contained an inharmonic complex tone with 7 equal-amplitude components at a spectral spacing of 5/11 octaves between adjacent components, followed by a pure-tone probe. The target tone was randomly chosen to be the 3rd, 4th or 5th component within the complex tone with equal *a priori* probability. A schematic representation of the stimuli is shown in Figure 1. In the Present/Absent task, the frequency of the probe tone was either the same as the frequency of the target tone (present), or was geometrically centered between the frequencies of the target tone and of one of its adjacent neighbors (absent) (Figure 1a). In the Up/Down task, the probe was either one semitone (1/12 octave) higher than the target tone (up) or one semitone lower (down) (Figure 1b). In both cases the two alternatives were presented with equal *a priori* probability. From trial to trial, the frequencies of the entire inharmonic complex were randomly roved within a one-octave frequency range (with uniform distribution on a logarithmic scale). This roving, along with the randomized position of the target tone within the complex, led to the frequency of the target tone being anywhere between 800 Hz and 3 kHz on any given trial. The inharmonic complex and probe tone were each 50 ms long,

including 10-ms raised-cosine rise and fall ramps, separated by a 100-ms silent gap. The level of each masker component was 45 dB SPL. In the enhanced conditions, a precursor was presented before the inharmonic complex. The precursor frequencies matched those of the masker in each trial (i.e., no component at the target frequency). The duration of the precursor was 62.5, 250, or 1000 ms, including 10-ms raised-cosine onset and offset ramps. The delay between the precursor offset and inharmonic complex onset was 10, 100, or 1000 ms.

Procedure—Participants were seated in a double-walled sound-attenuating booth. The stimuli were generated digitally using the AFC software package (Ewert, 2013) under Matlab (Mathworks, Natick, MA) at a 48-kHz sampling rate, delivered through an L22 soundcard (LynxStudio, Costa Mesa, CA) with a 24-bit resolution, and presented monaurally to the right ear through HD650 headphones (Sennheiser, Old Lyme, CT). The level of the target tone was initially set to 65 dB SPL (i.e., 20 dB higher than the individual masker components) and was varied adaptively following a two-down one-up rule that tracks the 70.7% correct point on the psychometric function (Levitt, 1971). Feedback was provided after each trial. The level of the probe tone was always the same as that of the target tone. Initially the level of the target was varied in steps of 5 dB. After two reversals in the direction of the adaptive tracking procedure, the step size was reduced to 2 dB. The run was terminated after eight reversals and the threshold was computed as the average target level at the last six reversal points of the tracking procedure. The three precursor durations and three precursor-target gaps resulted in nine conditions, plus a control condition with no precursor, for a total of ten conditions. Each condition was tested six times for each participant, and the conditions were presented in a different random order for each repetition and each participant. Threshold was defined as the mean of the last four repetitions for each condition and participant. To control for order effects between the two tasks, half the participants completed the Up/Down task before continuing to the Present/Absent task, and the other half completed the two conditions in the opposite order. All participants provided written informed consent, and the protocols were approved by the University of Minnesota Institutional Review Board.

Screening and training—Before the main experiment, every participant was required to pass two pitch-discrimination training and screening sessions. In the first session, the participants were presented with two consecutive pure tones, each of 50 ms duration, separated from each other by a silent gap of 100 ms. The two tones were either the same or differed in pitch the same way as the target and probe tones in the Present/Absent task. Participants were asked whether the two tones had the same or different pitch. In the second session, the second tone was either one semitone higher or lower than the first tone, as in the Up/Down task. Participants were asked to judge the direction of the pitch change. The tones in both training sessions were roved in frequency from trial to trial in the same way as in the main experiment. All participants had to obtain at least 80% correct in both sessions to pass. A total of eighteen participants were tested and four were not able to pass the second screening session. Therefore only fourteen participants were included in the main experiment. The average performance for the 14 participants was 97.5% correct in the Same/ Different training task and 94.3% correct in the Up/Down training task.

Results

The thresholds of individual participants in the no-precursor control condition are shown in Figure 2 in terms of target-to-masker ratio (TMR), or the level of the target relative to the level per component in the remainder of the inharmonic complex. The thresholds in the two tasks for each participant are shown as an ellipse, marking the standard deviation around the mean of the last four of the six repetitions. Thresholds from the seven participants who completed the Present/Absent task first are shown as solid black ellipses and thresholds from the remaining seven participants who completed the Up/Down task first are shown as dashed grey ellipses. The data show considerable variability in thresholds, both within and between participants. Whereas some participants were consistently able to perform the task with the target at or below the level of the other components in the complex (a TMR of 0 dB or lower), others were above this level by 10 dB or more. A mixed-model analysis of variance (ANOVA), with task type as a within-subjects factor and task order as a between-subjects factor, revealed a significant effect of task type, $F(1,12) = 14$, $p = 0.003$, $\eta_p^2 = 0.54$, but no effect of task order, $F(1,12) = 0.293$, $p = 0.60$, $\eta_p^2 = 0.16$, and no interaction between the two, $F(1,12) = 0.99$, $p = 0.34$, $\eta_p^2 = 0.076$. (A Greenhouse-Geisser correction for lack of sphericity was included where appropriate in all ANOVAs reported in this study.) The fact that the Present/Absent thresholds were significantly higher than the Up/Down thresholds is also apparent in Figure 2 in that the centers of most of the ellipses lie above the major diagonal.

Figure 3 shows the amount of enhancement, calculated as the difference in thresholds between each of the nine conditions with a precursor and the no-precursor condition. Enhancement in the Present/Absent task is shown in black and enhancement in the Up/Down task is shown in grey. Enhancement was significantly greater than zero for all conditions tested, one sample t-test, p < 0.04 in all nine cases for both tasks. The maximal average enhancement observed was over 24 dB for the 1000-ms precursor and 10-ms gap in both tasks. Enhancement increased with increasing precursor duration and decreasing precursor-target gap (Figure 3). The time course of enhancement was similar in both tasks, although the overall enhancement appeared greater in the Up/Down tasks. A repeated-measures three-way ANOVA on the amount of enhancement with task type (Present/Absent or Up/Down), gap duration (10, 100, or 1000 ms), and precursor duration (62.5, 250, or 1000 ms) as factors showed a significant effect of task type, $F(1,13) = 4.96$, $p = 0.044$, $\eta_p^2 = 0.276$, a significant effect of gap duration, $F(1.7,21.7) = 49.6$, $p < 0.001$, $\eta_p^2 = 0.792$, and a significant effect of precursor duration, $F(1.4,18.1) = 50.8$, $p < 0.001$, $\eta_p^2 = 0.796$. The interaction between gap and precursor duration was also significant, $F(2.35, 30.6) = 19.3$, $p < 0.001$, $\eta_p^2 = 0.597$. However, there was no interaction between task and gap, $F(1.24,16.1) = 1.56$, $p = 0.235$, $\eta_p^2 = 0.107$, or task and precursor duration, $F(1.57,20.4) = 1.24$, $p = 0.30$, $\eta_p^2 = 0.087$, and no significant 3-way interaction, $F(2.9,37.5) = 0.709$, $p = 0.548$, $\eta_p^2 = 0.052$. Thus, although there was a small difference in the overall amount of enhancement between the two task types, the pattern of results (i.e., the dependence of enhancement on precursor duration and gap duration) was not affected by the task.

Discussion

In Experiment 1, we found that the thresholds in the no-precursor condition in the Up/Down task were lower than those in the Present/Absent task, consistent with a previous study that proposed automatic "frequency shift detectors" in the auditory system that can signal a change in pitch, even if the target pitch itself is not perceived (Demany & Ramos, 2005). Whereas the previous study characterized this difference in terms of a change in sensitivity (*d'*), our results show that the difference in task performance is equivalent to a change in target level of 7.4 dB on average. Although the two tasks are different and different probe tones are used, for an ideal listener without automatic frequency-shift detectors, the model described in the study by Demany and Ramos (2005) would predict a similar *d'* in the two conditions: *d'* is equal to 1.95 in the Up/Down task and 2.09 in the Present/Absent task when the internal noise has a standard deviation of 1 semitone and *d'* is equal to 0.66 and 0.53, respectively, when the internal noise has a standard deviation of 2 semitones. A certain change in the target level at threshold can only be directly compared to changes in *d'* at a fixed level if the slopes of the underlying psychometric functions are the same. We used the data from the individual trials in the adaptive-tracking procedure to estimate the psychometric functions for each listener in both tasks by fitting a Weibull function using maximum likelihood method. The equation of the Weibull function for the 2AFC task is:

$$G(x) = 1 - 0.5e^{-\left(\frac{kx}{t}\right)^b}, \text{ where } k = \left[-\ln\left(\frac{1 - 0.707}{0.5}\right)\right]^{\frac{1}{b}}$$

In this equation, *x* is the target level and *G*(*x*) is the percent correct. *t* is the threshold at 70.7% correct, *b* determines the shape of the curve. *t* and *b* are free parameters for fitting. A few examples of the psychometric functions are shown in the supplemental Figure S1. The slope was defined as the first derivative of the fitted Weibull function at the threshold. The slopes from individual listeners were estimated for the no-precursor condition and the with-precursor condition that produced the maximal enhancement (1000-ms precursor and 10-ms gap for both tasks; see Table 1). In general, the slopes were quite shallow in all conditions. In the Present/Absent task, the averaged slope in the no-precursor condition was 1.41, meaning a 1-dB change in target level resulted in a 1.41 percentage point change in correct responses, and was 1.29 in the with-precursor condition. In the Up/Down task, the averaged slopes were 1.56 and 1.52 respectively for the no-precursor and with-precursor conditions. A repeated-measures two-way ANOVA with task type and precursor presence as factors showed no significant effect of task type on the psychometric function slope, $F(1,13) = 0.83$, $p = 0.38$, $\eta_p^2 = 0.06$. There was also no significant effect of the precursor presence, $F(1,13) = 0.12$, $p = 0.73$, $\eta_p^2 = 0.009$, and no significant interaction between task type and precursor presence, $F(1,13) = 0.03$, $p = 0.87$, $\eta_p^2 = 0.002$. The fact that the averaged slopes in the two tasks are not measurably different supports the model predictions of Demany and Ramos (2005), and provides justification for directly comparing the differences between conditions in terms of dB change in threshold.

We found that precursors as short as 62.5 ms could generate significant enhancement that lasts for at least 1 s. Such slow decay of enhancement emphasizes the importance of using a one-interval task in estimating the actual enhancement to avoid cumulative effects across

trials, and to avoid the masker from the non-signal interval acting as an enhancer for the signal interval. The maximal enhancement observed in Experiment 1 was over 24 dB on average, which is much larger than the 5–10 dB enhancement typically reported in many two-interval tasks if the sound intensity of maskers, precursor duration and gap duration are comparable (Viemeister, 1980). The size of enhancement in our study is comparable with that found by Byrne et al. (2013) where participants reported the "pop-out" of the target tone, although the authors proposed that the larger enhancement was due to the reduction of informational masking. Indeed, in the no-precursor condition in our study, the individual thresholds varied from −10 dB to 29 dB TMR. Such large differences in individual performance between participants are often interpreted as being indicative of informational masking (Lutfi, Kistler, Oh, Wightman, & Callahan, 2003; Neff & Dethlefs, 1995).

If the observed enhancement was at least partially caused by a reduction in informational masking, less enhancement would be expected in the Up/Down task than in the Present/Absent task, because the detection of the frequency shift is considered more "automatic," less dependent on attention, and hence presumably less susceptible to informational masking (Demany et al., 2001). In addition, participants with higher thresholds in the no-precursor condition should show more enhancement because greater informational masking would create more opportunity for enhancement to be effective. In fact, slightly *more* enhancement was observed in the Up/Down condition than in the Present/Absent condition, contrary to predictions based on informational masking. In addition, there was no significant correlation between the amount of enhancement and the no-precursor thresholds for individual participants in most of the nine conditions, with the exception of a *negative* correlation in the 250-ms precursor and 10-ms gap condition, $r(14) = −0.64$, $p = 0.015$, and in the 62.5-ms precursor and 10-ms gap condition, $r(14) = −0.58$, $p = 0.03$ (no correction for multiple comparisons). Again, if anything, the relationship is opposite to that predicted by an explanation based on informational masking. These results suggest that the enhancement measured here is unlikely to be caused primarily by a release from informational masking.

The enhancement observed in another one-interval paradigm, this time using a binaural centering task, was about 4–5 dB (Byrne et al., 2011) – much smaller than the enhancement observed in our study. There are several possible explanations for this. One is that the Byrne et al. study used a target well above the levels of the masker components, making it quite salient even in the absence of a precursor. One possible explanation of the enhancement effect involves the perceptual segregation of the target from the masker by the precursor, with the precursor and masker forming a single stream (Carlyon, 1989). If the target is already salient regardless of the presence of the precursor, as in the Byrne et al. study, then the enhancement produced by the precursor might be expected to be less, as the target may be heard as a separate object even without the precursor. Another proposed mechanism for enhancement is the adaptation of inhibition produced by the masker on the target, such that the response to the target tone is less inhibited by the masker when a precursor is presented before the masker (Byrne et al., 2011; Viemeister & Bacon, 1982). It has been shown in both physiological and psychophysical studies that the suppression of a probe tone by a simultaneous masker is dependent on the level relationship between the probe and masker. Little or no suppression is expected when the probe level is much higher than the masker

level (Duifhuis, 1980; Sachs & Kiang, 1968; Shannon, 1976). Similarly, any inhibitory effects would be expected to decrease with increasing difference between the target and masker levels in the current paradigms. Thus, the target in the Byrne et al. (2011) study may not have been sufficiently inhibited by the surrounding masker tones to exhibit a large release from inhibition in the presence of the precursor. To provide further insights into the possible underlying mechanisms and their time courses, a series of additional conditions was tested in Experiment 2.

Experiment 2

Two broad categories of explanation have been used to understand enhancement effects. One category invokes perceptual grouping (Carlyon, 1989) or cueing (Richards et al., 2004) principles: either the precursor and masker are perceptually grouped based on their spectral similarity and temporal proximity so that the target forms a separate perceptual object, following the "old-plus-new" heuristic outlined by Bregman (1990, p. 222; earlier mentioned in Helmholtz, 1859, p. 59), or the precursor acts as a cue to help identify the masker, and thus distinguish it from the target (Kidd, Richards, Streeter, Mason, & Huang, 2011). The other category invokes neural adaptation and inhibition, whereby the target's neural representation is enhanced by adaptation of the masker due to the precursor stimulus (Nelson & Young, 2010; Viemeister, 1980). These two categories are in some ways explanations at different conceptual levels, and are not mutually exclusive (for instance, adaptation could be one mechanism by which the old-plus-new organizational principle is implemented neurally). Nevertheless, it is possible to derive different predictions from these two categories of explanation. For instance, a grouping-based hypothesis would predict that using a different stimulus for the precursor should reduce enhancement, as the similarity between the precursor and masker is reduced. In contrast, an adaptation-based hypothesis would not necessarily predict a reduction in enhancement based on dissimilarity between precursor and masker, so long as the spectral extent of the two stimuli are similar. Some studies have shown that the perceptual similarity between the precursor and the masker is not crucial for producing enhancement. For instance, a notched-noise, or an inharmonic complex tone was equivalent to a harmonic complex precursor if they had the same spectral envelope (Summerfield et al., 1987; Viemeister et al., 2013). A multi-tone precursor was still efficient even its components were gated asynchronously and the amount of enhancement was not different from that produced by the synchronized multi-tone precursor (Carcagno et al., 2013). On the other hand, one study did find that a notched-noise precursor produced smaller enhancement compared with a multi-tone precursor which was an exact copy of the masker (Kidd et al., 2011). The different conclusions on the effect of notched noise might be due to the different experiment designs. The maskers varied from trial to trial in the study by Kidd et al. (2011), whereas a fixed masker was used across trials in the other two studies (Summerfield et al., 1987; Viemeister et al., 2013). The effect of perceptual similarity might be more pronounced in a task with high masker uncertainty, which provides more opportunity for the perceptual grouping or cueing to take effect. The first condition in Experiment 2 tests the effect of perceptual similarity on enhancement with our paradigm by using spectrally notched noise as the precursor, with the same overall bandwidth as the complex-tone masker, but with a very different perceptual quality.

Even though the perceptual quality of a notched noise and an inharmonic complex tone are very different, listeners could still potentially benefit from the precursor because it provides a cue regarding the general frequency range of the expected target in each trial, along with its expected onset time. In this case, the spectral notch in the precursor may not be as important as the outer spectral edges of the precursor, which define its spectral extent. To estimate the potential benefit of such spectro-temporal cueing, we measured thresholds in a second condition in Experiment 2 with a bandpass-noise precursor, with the same overall spectral extent as the notched noise, but without the spectral notch. The bandpass-noise precursor should produce no spectral enhancement of the signal, and so any improvement of thresholds relative to the no-precursor condition may be attributable to cueing the general frequency range and the onset time of the expected target in each trial. Since different mechanisms might operate on different time scales, the time course of enhancement was also measured for different precursor types.

It is also possible that the spectral location of the notch in the notched-noise precursor could assist in cueing the frequency of the target tone. If so, listeners may be able to base their judgments in the frequency discrimination tasks on the relationship between the notched noise and the comparison tone, without any reference to the target tone at all. To test this hypothesis, a third condition was tested in which participants compared the pitch of the notched-noise precursor and the comparison tone directly.

Methods

Participants—Eight participants were tested, four of whom had participated in Experiment 1. The four new participants were all trained and screened with pure-tone pitch comparison tasks, as described in Experiment 1, and were also given training in the Present/Absent task. The Present/Absent task training was the same as that in Experiment 1 except that there were only two conditions: no precursor and a 1000-ms precursor with a 10-ms gap, which had produced the largest amount of enhancement in Experiment 1. The training used a constant stimulus task, with the target tone in all stimuli presented at 65 dB SPL. There were 40 trials in a block and each condition was repeated twice in a random order. The four new participants achieved an average 76.9% correct in the no-precursor condition and 93.8% correct in the with-precursor (enhanced) condition.

Stimuli and procedure—The procedures for measuring thresholds were the same as in Experiment 1. In the first two conditions, the masker, target, and comparison tones were also the same as those used in Experiment 1. The only difference was the precursor stimulus, which was a bandpass noise that had the same overall bandwidth and edge frequencies as the masker in each trial. In the notched-noise condition, the noise had a spectral notch that had the same width as the gap between the two masker components on either side of the target in Experiment 1; in the bandpass condition, the noise was the same, except that it had no spectral notch. Threshold equalizing noise (TEN)(Moore, Huss, Vickers, Glasberg, & Alcantara, 2000) was used to generate both noises, and the level of the noise was set to 45 dB SPL per equivalent rectangular auditory filter bandwidth (ERB)(Glasberg & Moore, 1990) at 1 kHz within the pass-band. The level was chosen to be the same as the level per component for the multi-tone precursor in Experiment 1 so that the two different precursors

should cause equivalent amount of adaptation of the masker components. Similar to Experiment 1, the time course of enhancement, as a function of precursor duration and precursor-target gap, was tested with the notched-noise and bandpass-noise precursors separately. Because Experiment 1 had not revealed any qualitative differences between the results using the Up/Down and Present/Absent procedures, only the Present/Absent procedure was used to measure thresholds in these two conditions. The same three precursor durations and three gap durations were tested for both precursor types.

The third condition was similar to the notched-noise experiment, except that the masker-plus-target complex was replaced by a silent gap, so that only the precursor and the comparison tone were presented. This condition was only tested with the 1000-ms precursor and 10-ms gap. The comparison tone was set to 45 dB SPL. Because no target was present, there was nothing on which to adapt the presentation level, so performance was tested using a constant-stimulus task with 200 trials per participant, and participants were asked to judge whether the comparison tone was higher or lower than the precursor sound. In all conditions, feedback was provided after each trial. In the third condition the feedback was based on the center frequency of the noise's spectral notch, which is where the target would be have been presented.

Results

Figure 4 shows the amount of enhancement in each of the nine conditions with notched-noise precursors (grey lines) or with bandpass-noise precursors (black lines). The enhancement for the notched-noise precursors showed a similar time course as the multi-tone masker precursors in Experiment 1 (Figure 4): enhancement increased with increasing precursor duration and decreasing precursor-target gap. The magnitude of enhancement seemed somewhat smaller when the precursor was a notched noise, compared to the complex tone of Experiment 1, with an average maximal enhancement of about 17 dB in the condition with the longest precursor and shortest gap. Four listeners participated in both Experiment 1 and 2. Figure 5a shows the amount of enhancement averaged across these four participants in each of the nine conditions for both complex-tone precursors and notched-noise precursors. When the precursor was short (62.5 ms), the difference in enhancement between the two precursor types was small regardless of the gap duration. When the precursor was longer (250 ms and 1000 ms), the difference became more pronounced at shorter gaps (10 ms and 100 ms): the enhancement produced by complex-tone precursors was larger than that produced by notched-noise precursors. This difference was about 9 dB maximally in the condition with a 250 ms long precursor and a 10 ms gap (Figure 5b). A three-way ANOVA was performed on the amount of enhancement shown in these four listeners, with precursor type (notched noise or complex tone), precursor duration, and gap duration as within-subjects factors. The ANOVA revealed a significant effect of precursor duration, $F(1.4,4.2) = 22.6$, $p = 0.007$, $\eta_p^2 = 0.88$, and of gap duration, $F(1.7,5.1) = 21.6$, $p = 0.003$, $\eta_p^2 = 0.88$. However, the effect of precursor type was not significant, $F(1,3) = 1.9$, $p = 0.26$, $\eta_p^2 = 0.385$. The interactions between precursor type and gap duration, $F(1.1,3.4) = 7.3$, $p = 0.06$, $\eta_p^2 = 0.71$, between precursor type and precursor duration, $F(1.9,5.7) = 4.5$, $p = 0.07$, $\eta_p^2 = 0.6$, and between precursor duration and gap duration, $F(1.7,5.2) = 5.5$, $p = 0.055$, $\eta_p^2 = 0.6$, all failed to reach significance, as did the three-way interaction, $F(2.2,6.6)$

= 2.3, $p = 0.17$, $\eta_p^2 = 0.44$. Although neither the main effect of precursor type nor any interaction with it was significant, this may reflect a lack of statistical power, based as it is on only four subjects, rather than an absence of any effect.

When the precursor was replaced by a bandpass noise, the difference between thresholds with and without a precursor was generally smaller, but was still positive on average in all but one case (see Fig. 4). Averaged across all conditions, the threshold difference was about 3 dB. In the third condition, where the masker and target were absent and participants had to base their judgments on a comparison of the notched noise and the comparison probe tone, the average performance was 54% correct, which was slightly higher than the chance level of 50%, $t(8) = 4.7$, $p = 0.002$, $d = 1.54$. Although chance level was exceeded, the fact that performance was very low, and well below the level required for threshold in our adaptive procedures, suggests that participants were not likely to have used the spectral notch in the precursor to match the comparison tone. Because feedback was provided after each trial, it is not likely that the participants used other attributes of the precursor sounds, such as the edges of the upper or lower frequency bands.

Discussion

The results from Experiment 2 showed that a notched-noise precursor produced large enhancement effects with similar temporal characteristics to those found with the complex-tone precursor from Experiment 1. However, the overall size of the enhancement effect with the notched-noise precursor seemed smaller than that produced by the complex-tone precursor when the precursor was long and the gap was short. The difference in enhancement between the tonal and noise precursor, which reached as much as 9 dB, may represent the portion of the enhancement effect that can be explained through perceptual grouping mechanisms. This portion of enhancement tended to be smaller and close to zero when the precursor was short (62.5 ms), or when the gap between the precursor and masker was long (1000 ms). Such a time course is consistent with previous findings that sequential grouping depends on temporal proximity (Bregman, 1990). It also remains possible that some of the remaining enhancement is due to grouping, based on the similarity of the spectral extent of the noise precursor and tonal masker. Segregation based on non-spectral cues tends to be weaker than that based on spectral separation (Vliegen, Moore, & Oxenham, 1999); thus it could be argued that the noise precursor and tonal masker might be still be grouped together. However, the large perceptual difference between the noise and tone renders this possibility less likely.

We also found that thresholds were lower in the presence of the bandpass noise than with no precursor. Because of the lack of a spectral gap in the band-pass noise precursor, this difference cannot be considered enhancement in the traditional sense; instead, it may reflect an effect of cueing the onset timing and gross spectral location of the target and masker combination. The spectral location of the target varied from 800 Hz to 3 kHz from trial to trial, and the precursor could potentially cue listeners to better focus on a narrower frequency range. This effect was about 3 dB on average and was observed for most conditions, even when the gap was 1000 ms long. This effect of spectro-temporal cueing could also contribute to the enhancement using notched-noise and complex-tone precursors.

In order to remove this portion from the total enhancement effect, we redefined enhancement as the difference between thresholds with the bandpass-noise precursor and thresholds with the notched-noise precursor in each of the nine conditions, similar to the approach used by Carcagno et al. (2012). Figure 6a shows the amount of enhancement, defined in this way, averaged across all eight participants. This newly defined enhancement had a similar time course as the enhancement in Experiment 1: enhancement increased as the precursor became longer and the gap became shorter, reaching a maximum of over 13 dB on average. This result suggests at least an effective 13 dB increase in target level was due to the frequency specific adaptation from the precursor in Experiment 1 when the precursor was 1000 ms long and the gap was 10 ms. However, the reduced enhancement effect led to the finding that the enhancement was no longer significant under all conditions: the enhancement remained significant when the gap was 10 ms, one sample t-test, $p < 0.001$, $d > 2$, and when the gap was 100 ms, $p < 0.012$, $d > 1.4$, but it was no longer significant for any precursor duration when the gap was 1000 ms, $p > 0.06$, $d < 0.8$ (no corrections for multiple comparisons). This suggests that when the gap is as long as 1000 ms, the enhancement effect in Experiment 1 was mainly produced by other non-adaptation mechanisms, such as the cueing general frequency range and onset time of the target tone. A three-way ANOVA was performed on the amount of enhancement from all eight listeners who participated in Experiment 2, with precursor type (notched noise or bandpass noise), precursor duration and gap duration as within-subjects factors. The ANOVA revealed significant main effects of precursor type, $F(1,7) = 8.2$, $p = 0.024$, $\eta_p^2 = 0.54$, precursor duration, $F(1.3,9.3) = 25.2$, $p < 0.001$, $\eta_p^2 = 0.78$, and of gap duration, $F(1.7,12.1) = 17.4$, $p < 0.001$, $\eta_p^2 = 0.71$. The interaction between precursor type and gap duration was also significant, $F(1.8,13) = 30.2$, $p < 0.001$, $\eta_p^2 = 0.81$, as were the interactions between precursor type and precursor duration, $F(2,14) = 9.6$, $p = 0.002$, $\eta_p^2 = 0.58$ and between precursor duration and gap duration, $F(1.8,12.8) = 12.3$, $p = 0.001$, $\eta_p^2 = 0.64$. There was no significant three-way interaction, $F(2.6,18.2) = 1.0$, $p = 0.39$, $\eta_p^2 = 0.13$.

In light of this different definition of enhancement, we recruited five of the original participants from Experiment 1 to take part in a new control experiment, in which the same Present/Absent task was used as in Experiment 1, except that the precursor included both the masker and target tone. We calculated enhancement with the new definition: the difference between the thresholds with and without target tone in the precursor. The average enhancement across these 5 participants in nine conditions is shown in Figure 6b. Overall, the pattern of results was very similar to the pattern of enhancement calculated by the original definition in Experiment 1, although the maximum amount of enhancement (with the longest precursor and shortest gap) was now around 24 dB, rather than the 29 dB found in the same five participants in Experiment 1. After the portion of the improvement in thresholds produced by the precursor with no spectral notch (compared to no precursor) was subtracted, the newly defined enhancement was no longer significant in all conditions. Although it remained significant for the 1000-ms precursor in all conditions, it was not significant when the precursor was shorter (62.5 or 250 ms) and the gap was 1000 ms, one sample t-test, $p > 0.15$, $d < 0.8$, or when the precursor was 62.5 ms and the gap was 100 ms, $t(4) = 2.36$, $p = 0.078$, $d = 1.1$. This time course is similar to the findings from a previous study, where the authors reported a 90% decrease in the enhancement as the gap duration

increased from 10 ms to 600 ms for a 500 ms precursor using the new definition (Carcagno et al., 2012).

Our results suggest that the time course of enhancement in Experiment 1 results from different sources that might operate on different time scales. Frequency-specific adaptation might be the main source for enhancement when the gap between precursor and target is relatively short. Perceptual grouping mechanisms might play a part in producing the enhancement effect when the precursor is long and the gap is short (< 1000 ms in our experiment). The estimated enhancement produced by perceptual grouping can be as large as 9 dB depending on both the precursor duration and the gap duration. Some previous studies did not find a difference in enhancement with different precursor types (Summerfield et al., 1987; Viemeister et al., 2013). We noticed that the magnitude of the enhancement with the previous paradigms was much smaller than the enhancement reported in our study. The possible underestimation of the true enhancement effect might make it difficult to separate the effect caused by adaptation from that produced by perceptual grouping. In our study and the one other study that also revealed a smaller enhancement effect from the notched-noise precursor (Kidd et al., 2011), the maskers and target changed in frequency from trial to trial. Such masker and target uncertainty could also explain the observed effect of perceptual similarity between the precursor and the masker in both studies. A study by Carcagno and colleagues also tested the effect of perceptual similarity by gating the components of the precursor asynchronously (Carcagno et al., 2013). They did not find any effect of perceptual similarity even with the frequency rove from trial to trial. However, in their study, the target tone was always presented first in each trial. It is likely that the target uncertainty is a critical factor to observe the effect of perceptual similarity between the precursor and the masker. For the three precursor durations tested, cueing the target onset time and general frequency range could also play a part in generating enhancement and it might be the main source for the long-lasting enhancement effects (1000 ms gap) observed in Experiment 1.

General discussion

In Experiment 1, we demonstrated that a new one-interval task with adaptive threshold tracking can be used as a robust performance measure of auditory enhancement. A precursor as short as 62.5 ms could significantly enhance the target tone and the enhancement could last at least 1000 ms. Enhancement increased with increasing precursor duration and decreasing gap between precursor and masker, consistent with previous studies (Carcagno et al., 2012; Viemeister, 1980). In the current study, we observed a maximum of over 24 dB enhancement on average, which is much greater than the magnitude of enhancement from many earlier studies using two-interval tasks. One reason for this difference may be that the one-interval task with frequency roving avoids potential confounds involving interactions between intervals and across trials. We used two different tasks that are thought to engage two different auditory processes: in the Present/Absent task, participants must hear out the target tone from the masker in order to compare it with the following test tone, whereas in the Up/Down task, participants may detect a change in pitch, even though the pitch of the target tone itself is not perceived (Demany & Ramos, 2005). The magnitude and time course of enhancement were similar in both tasks, suggesting that the enhancement observed is a

relatively general and robust effect. In particular, the finding that the amount of enhancement was no less in the case of the Up/Down task, which is thought to rely more on "automatic" frequency-shift detection, than in the Present/Absent task, which appears to involve "hearing out" the target within the masker, suggests that informational masking cannot account for the large enhancement effects observed. Further support for this conclusion was provided by the lack of correlation between thresholds without a precursor and the amount of enhancement: if thresholds without the enhancer were primarily governed by informational masking then participants with the highest thresholds without the precursor (and presumably the highest degree of informational masking) should have shown the largest enhancement effects. No such correlation was found.

Experiment 2 again examined the time-course of enhancement, while focusing on two additional factors that may influence the amount of measured enhancement: perceptual grouping and spectro-temporal cueing. The strength of grouping by perceptual similarity was reduced by using a perceptually dissimilar (notched-noise) precursor, rather than a copy of the inharmonic-tone masker. The role of spectro-temporal cueing was tested by using a bandpass-noise precursor with no spectral gap, which may have produced some cueing but should not have produced traditional enhancement, as it had no spectral gap surrounding the target. Both manipulations had some effect, and the maximum amount of enhancement (defined as the difference between thresholds with and without a spectral notch in the precursor) of around 13 dB was smaller than that found in Experiment 1 (see Fig. 6). Although the general pattern of results and the dependence on precursor duration and precursor-target gap were similar, the amount of enhancement was no longer significant for the longest-duration gap (1000 ms) between the precursor and the target.

Potential neural correlates of enhancement have been explored at different levels of the auditory pathways. The adaptation of maskers by precursors has been found at the level of the auditory nerve, although those results cannot explain the increased effective gain of the target tone shown in psychophysical studies (Palmer et al., 1995). An increase in response to the target tone has been found in the cochlear nucleus, but there it was limited to the onset responses (Scutt & Palmer, 1998). Neurons in the inferior colliculus showed a more sustained facilitation in responses to the target tone when the precursor was presented (Nelson & Young, 2010). Carcagno and colleagues measured the auditory steady-state response (ASSR) to the target tone (Carcagno, Plack, Portron, Semal, & Demany, 2014). A high amplitude modulation (AM) rate was used so that the ASSRs recorded originated mainly from the brainstem. They did not find any increase in the ASSRs to the enhanced target tone. Since the ASSR magnitude correlates well with loudness (Menard, Gallego, Berger-Vachon, Collet, & Thai-Van, 2008), the perceptual enhancement cannot be explained by the increased neural responses at midbrain level. This finding did not entirely contradict the single neuron study by Nelson and Young (2010) though. The ASSR reflects the averaged population responses which might not reveal the facilitated responses of a small subgroup of neurons in the inferior colliculus. It is also possible that the single-unit responses in inferior colliculus were not sufficient to account for the perceptual effect. It will be interesting to look at neural correlates for enhancement at the cortical level. Although those physiological studies did not examine the temporal course of enhancement, neurons at different levels of the auditory pathway have been shown to exhibit different temporal

integration characteristics. Some studies have shown long lasting stimulus-specific adaptation, up to seconds, at higher levels of the auditory pathways, such as the inferior colliculus, thalamus, and primary auditory cortex (Antunes, Nelken, Covey, & Malmierca, 2010; Malmierca, Cristaudo, Perez-Gonzalez, & Covey, 2009; Ulanovsky, Las, & Nelken, 2003; Zhao, Liu, Shen, Feng, & Hong, 2011). It is possible that when the gap between precursor and target is small, enhancement is an accumulated effect from the adaptation at multiple processing levels, whereas the enhancement at long gaps might stem solely from more central loci. A study by Carcagno et al. (2012) showed different time courses for ipsilateral and contralateral enhancement: enhancement decayed more rapidly for ipsilateral precursors, again suggesting multiple sources for enhancement. Different task designs and stimulus parameters could engage different sources, which might explain some of the discrepancy in the findings of previous studies, such as the effect of perceptual similarity between precursor and masker (Byrne et al., 2013; Carcagno et al., 2013; Carlyon, 1989; Summerfield et al., 1987; Viemeister et al., 2013) or the enhancement from a contralateral precursor (Carlyon, 1989; Erviti et al., 2011; Kidd et al., 2011; Richards et al., 2004; Summerfield et al., 1984; Viemeister, 1980).

We observed that thresholds were somewhat lower (better) in the presence of the bandpass noise than with no precursor. This improvement cannot be due to enhancement, as there was no spectral notch in the noise. There are at least two possible mechanisms for the improvement in performance observed with bandpass-noise precursors. The possibility that inspired the experiment is that the precursor (even without the spectral notch) could provide a spectral and temporal cue for listeners to focus on the target tone. Another possibility is related to the "overshoot" effect: the detection of a brief tone is improved if the tone is delayed from the onset of a broadband masker (Carlyon & White, 1992; Hicks & Bacon, 1992; Jennings, Heinz, & Strickland, 2011; Strickland, 2004; Zwicker, 1965). This phenomenon may be related to signal enhancement involving the gain control by activation of medial olivo-cochlear efferent reflex (MOCR). The time course of MOCR is thought to be on the scale of 100–200 milliseconds (Backus & Guinan, 2006), however, which probably cannot fully account for the long-lasting effects observed here with the 1000-ms gap in Experiment 2.

As mentioned in the introduction, people with hearing impairment and with cochlear implants have generally shown less enhancement than normal-hearing listeners in traditional enhancement tasks. Testing these populations on the tasks described here may provide another path towards identifying the multiple sources that contribute towards the overall enhancement effects observed here and elsewhere. Aside from providing more scientific insights into the origins of the effect, differences in performance between these groups and normal-hearing listeners may provide important input in the design of algorithms that seek to restore normal enhancement effects in hearing-impaired listeners and cochlear-implant users.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Antunes FM, Nelken I, Covey E, Malmierca MS. Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. PLoS One. 2010; 5(11):e14071. [PubMed: 21124913]

Backus BC, Guinan JJ Jr. Time-course of the human medial olivocochlear reflex. J Acoust Soc Am. 2006; 119(5 Pt 1):2889–2904. [PubMed: 16708947]

Barlow, HB. Possible principles underlying the transformation of sensory messages. In: Rosenblith, W., editor. Sensory Communication. Cambridge, MA: MIT Press; 1961. p. 217-234.

Bregman, AS. Auditory Scene Analysis. Cambridge, MA: MIT Press; 1990.

Byrne AJ, Stellmack MA, Viemeister NF. The enhancement effect: evidence for adaptation of inhibition using a binaural centering task. J Acoust Soc Am. 2011; 129(4):2088–2094. [PubMed: 21476664]

Byrne AJ, Stellmack MA, Viemeister NF. The salience of enhanced components within inharmonic complexes. J Acoust Soc Am. 2013; 134(4):2631–2634. [PubMed: 24116400]

Carcagno S, Plack CJ, Portron A, Semal C, Demany L. The auditory enhancement effect is not reflected in the 80-Hz auditory steady-state response. J Assoc Res Otolaryngol. 2014; 15(4):621–630. [PubMed: 24845402]

Carcagno S, Semal C, Demany L. Auditory enhancement of increments in spectral amplitude stems from more than one source. J Assoc Res Otolaryngol. 2012; 13(5):693–702. [PubMed: 22766695]

Carcagno S, Semal C, Demany L. No Need for Templates in the Auditory Enhancement Effect. PLoS One. 2013; 8(6):e67874. [PubMed: 23826348]

Carlyon RP. Changes in the masked thresholds of brief tones produced by prior bursts of noise. Hear Res. 1989; 41(2–3):223–235. [PubMed: 2808151]

Carlyon RP, Long CJ, Deeks JM, McKay CM. Concurrent sound segregation in electric and acoustic hearing. J Assoc Res Otolaryngol. 2007; 8(1):119–133. [PubMed: 17216383]

Carlyon RP, White LJ. Effect of signal frequency and masker level on the frequency regions responsible for the overshoot effect. J Acoust Soc Am. 1992; 91(2):1034–1041. [PubMed: 1556305]

Dean I, Harper NS, McAlpine D. Neural population coding of sound level adapts to stimulus statistics. Nat Neurosci. 2005; 8(12):1684–1689. [PubMed: 16286934]

Demany L, Carcagno S, Semal C. The perceptual enhancement of tones by frequency shifts. Hear Res. 2013; 298:10–16. [PubMed: 23376551]

Demany, L.; Clément, S.; Semal, C. Does auditory memory depend on attention; Proceedings of the Physiological and Psychophysical Bases of Auditory Function; Shaker, Masstricht, The Netherlands. 2001. p. 461-467.

Demany L, Ramos C. On the binding of successive sounds: perceiving shifts in nonperceived pitches. J Acoust Soc Am. 2005; 117(2):833–841. [PubMed: 15759703]

Duifhuis H. Level effects in psychophysical two-tone suppression. J Acoust Soc Am. 1980; 67(3):914–927. [PubMed: 7358916]

Durlach NI, Mason CR, Kidd G Jr, Arbogast TL, Colburn HS, Shinn-Cunningham BG. Note on informational masking. J Acoust Soc Am. 2003; 113(6):2984–2987. [PubMed: 12822768]

Erviti M, Semal C, Demany L. Enhancing a tone by shifting its frequency or intensity. J Acoust Soc Am. 2011; 129(6):3837–3845. [PubMed: 21682407]

Ewert, SD. A modular framework for running psychoacoustic experiments and computational perception models; Proceedings of the International Conference on Acoustics AIA-DAGA; Merano, Italy. 2013. p. 1326-1329.

Glasberg BR, Moore BC. Derivation of auditory filter shapes from notched-noise data. Hear Res. 1990; 47(1–2):103–138. [PubMed: 2228789]

Hartmann WM, Goupell MJ. Enhancing and unmasking the harmonics of a complex tone. J Acoust Soc Am. 2006; 120(4):2142–2157. [PubMed: 17069312]

Helmholtz, H. On the Sensations of Tone as a Physiological Basis for the Theory of Music. 2nd English ed.. Ellis, AJ., translator. Whitefish, MT: Reprinted by Kessinger Publishing, 2005; 1859. 1885

Hicks ML, Bacon SP. Factors influencing temporal effects with notched-noise maskers. Hear Res. 1992; 64(1):123–132. [PubMed: 1490895]

Jennings SG, Heinz MG, Strickland EA. Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot. J Assoc Res Otolaryngol. 2011; 12(3):345–360. [PubMed: 21267622]

Kidd G Jr, Richards VM, Streeter T, Mason CR, Huang R. Contextual effects in the identification of nonspeech auditory patterns. J Acoust Soc Am. 2011; 130(6):3926–3938. [PubMed: 22225048]

Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am. 1971; 49(2):467–477. [PubMed: 5541744]

Lutfi RA, Kistler DJ, Oh EL, Wightman FL, Callahan MR. One factor underlies individual differences in auditory informational masking within and across age groups. Percept Psychophys. 2003; 65(3): 396–406. [PubMed: 12785070]

Malmierca MS, Cristaudo S, Perez-Gonzalez D, Covey E. Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. J Neurosci. 2009; 29(17):5483–5493. [PubMed: 19403816]

Menard M, Gallego S, Berger-Vachon C, Collet L, Thai-Van H. Relationship between loudness growth function and auditory steady-state response in normal-hearing subjects. Hear Res. 2008; 235(1–2):105–113. [PubMed: 18093768]

Moore BC, Huss M, Vickers DA, Glasberg BR, Alcantara JI. A test for the diagnosis of dead regions in the cochlea. Br J Audiol. 2000; 34(4):205–224. [PubMed: 10997450]

Neff DL, Dethlefs TM. Individual differences in simultaneous masking with random-frequency, multicomponent maskers. J Acoust Soc Am. 1995; 98(1):125–134. [PubMed: 7608391]

Nelson PC, Young ED. Neural correlates of context-dependent perceptual enhancement in the inferior colliculus. J Neurosci. 2010; 30(19):6577–6587. [PubMed: 20463220]

Okada M, Kashino M. The role of spectral change detectors in temporal order judgment of tones. Neuroreport. 2003; 14(2):261–264. [PubMed: 12598742]

Oxenham AJ, Fligor BJ, Mason CR, Kidd G Jr. Informational masking and musical training. J Acoust Soc Am. 2003; 114(3):1543–1549. [PubMed: 14514207]

Palmer AR, Summerfield Q, Fantini DA. Responses of auditory-nerve fibers to stimuli producing psychophysical enhancement. J Acoust Soc Am. 1995; 97(3):1786–1799. [PubMed: 7699160]

Richards VM, Huang R, Kidd G Jr. Masker-first advantage for cues in informational masking. J Acoust Soc Am. 2004; 116(4 Pt 1):2278–2288. [PubMed: 15532659]

Richards VM, Neff DL. Cuing effects for informational masking. J Acoust Soc Am. 2004; 115(1): 289–300. [PubMed: 14759022]

Sachs MB, Kiang NY. Two-tone inhibition in auditory-nerve fibers. J Acoust Soc Am. 1968; 43(5): 1120–1128. [PubMed: 5648103]

Scutt, MJ.; Palmer, AR. Physiological enhancement in cochlear nucleus using single tone precursors; Paper presented at the Assoc. Res. Otolaryngol. Abs.; 1998.

Shannon RV. Two-tone unmasking and suppression in a forward-masking situation. J Acoust Soc Am. 1976; 59(6):1460–1470. [PubMed: 939879]

Stilp CE, Alexander JM, Kiefte M, Kluender KR. Auditory color constancy: calibration to reliable spectral properties across nonspeech context and targets. Atten Percept Psychophys. 2010; 72(2): 470–480. [PubMed: 20139460]

Strickland EA. The temporal effect with notched-noise maskers: analysis in terms of input-output functions. J Acoust Soc Am. 2004; 115(5 Pt 1):2234–2245. [PubMed: 15139634]

Summerfield Q, Haggard M, Foster J, Gray S. Perceiving vowels from uniform spectra: phonetic exploration of an auditory aftereffect. Percept Psychophys. 1984; 35(3):203–213. [PubMed: 6728618]

Summerfield Q, Sidwell A, Nelson T. Auditory enhancement of changes in spectral amplitude. J Acoust Soc Am. 1987; 81(3):700–708. [PubMed: 3584677]

Thibodeau LM. Performance of hearing-impaired persons on auditory enhancement tasks. J Acoust Soc Am. 1991; 89(6):2843–2850. [PubMed: 1918628]

Thibodeau LM. Evaluation of auditory enhancement and auditory suppression in listeners with normal hearing and reduced speech recognition in noise. J Speech Hear Res. 1996; 39(5):947–956. [PubMed: 8898249]

Ulanovsky N, Las L, Nelken I. Processing of low-probability sounds by cortical neurons. Nat Neurosci. 2003; 6(4):391–398. [PubMed: 12652303]

Viemeister, NF. Adaptation of masking. In: van den Brink, G.; Bilsen, FA., editors. Psychophysical, Physiological and Behavioural Studies in Hearing. Delft, The Netherland: Delft University Press; 1980. p. 190-198.

Viemeister NF, Bacon SP. Forward masking by enhanced components in harmonic complexes. J Acoust Soc Am. 1982; 71(6):1502–1507. [PubMed: 7108024]

Viemeister NF, Byrne AJ, Stellmack MA. Spectral and level effects in auditory signal enhancement. Adv Exp Med Biol. 2013; 787:167–174. [PubMed: 23716221]

Vliegen J, Moore BC, Oxenham AJ. The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. J Acoust Soc Am. 1999; 106(2):938–945. [PubMed: 10462799]

Wang N, Kreft H, Oxenham AJ. Vowel enhancement effects in cochlear-implant users. J Acoust Soc Am. 2012; 131(6):EL 421–EL 426.

Zhao L, Liu Y, Shen L, Feng L, Hong B. Stimulus-specific adaptation and its dynamics in the inferior colliculus of rat. Neuroscience. 2011; 181:163–174. [PubMed: 21284952]

Zwicker E. Temporal Effects in Simultaneous Masking and Loudness. J Acoust Soc Am. 1965; 38:132–141. [PubMed: 14347604]
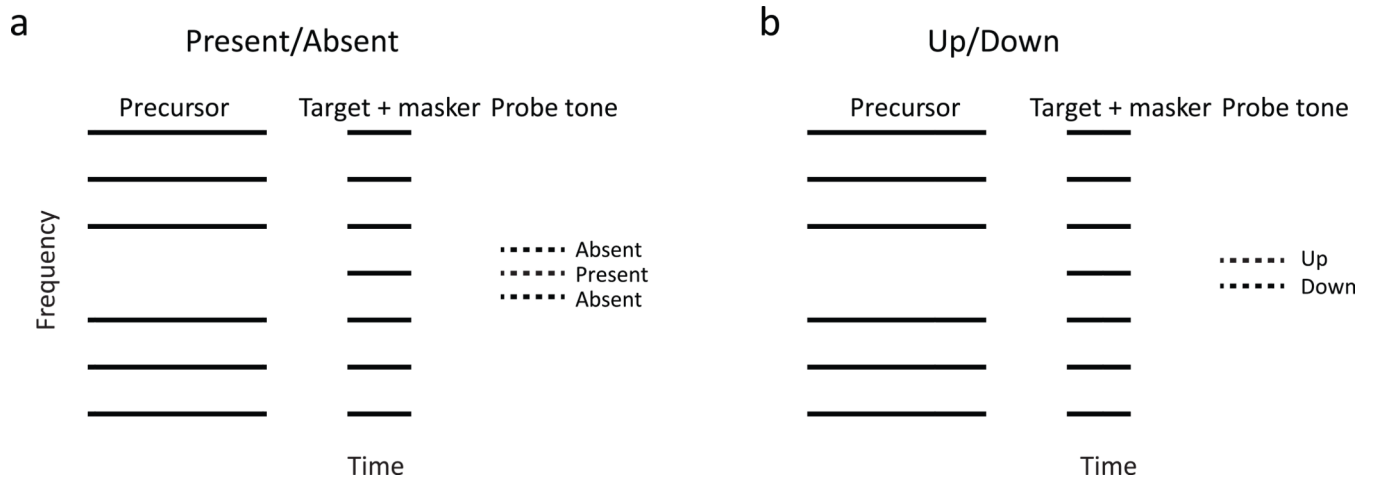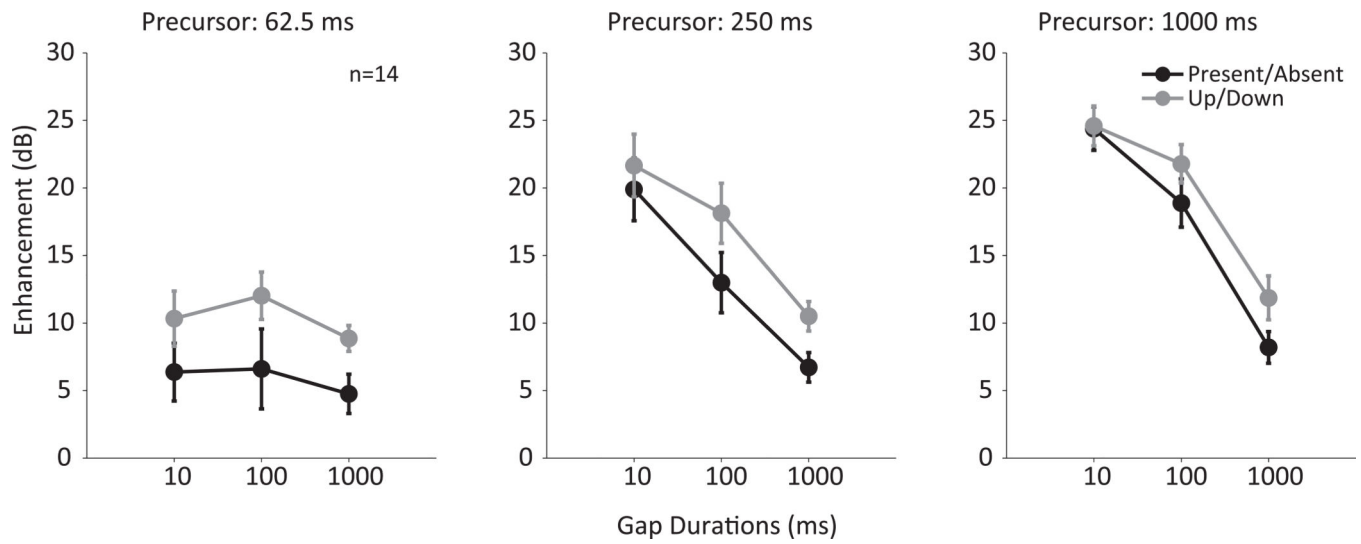
Figure 1.
Schematic representation of the stimuli. Each interval includes a seven-component inharmonic complex comprising a target tone and a six-tone masker, followed by a probe tone. The dashed lines indicate all possible positions of the probe tone relative to the target in different tasks. Only one probe tone is presented in each trial. The target tone can be the 3rd, 4th or 5th component of the complex tone (4th in this example). The complex and probe tone are 50 ms long with a delay of 100 ms between complex offset and probe onset. The duration of the precursor can be 62.5, 250, or 1000 ms. The gap between precursor and complex can be 10, 100, or 1000 ms. a, In the Present/Absent task, the probe tone is either the same as the target tone (present) or centered in frequency between the target tone and one of the most adjacent component (absent). b, In the Up/Down task, the probe tone is either a semitone higher (up) or lower (down) than the target tone.
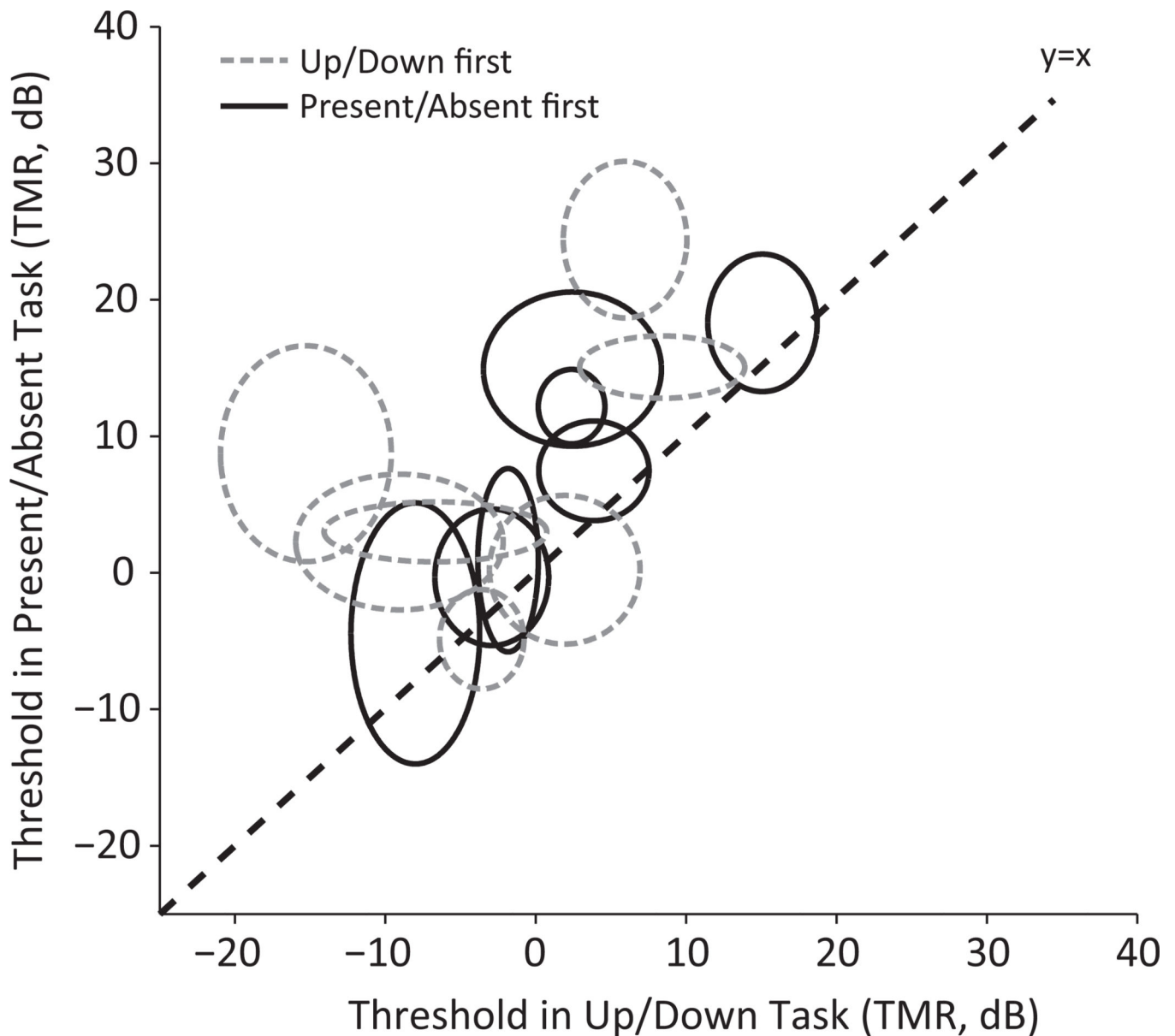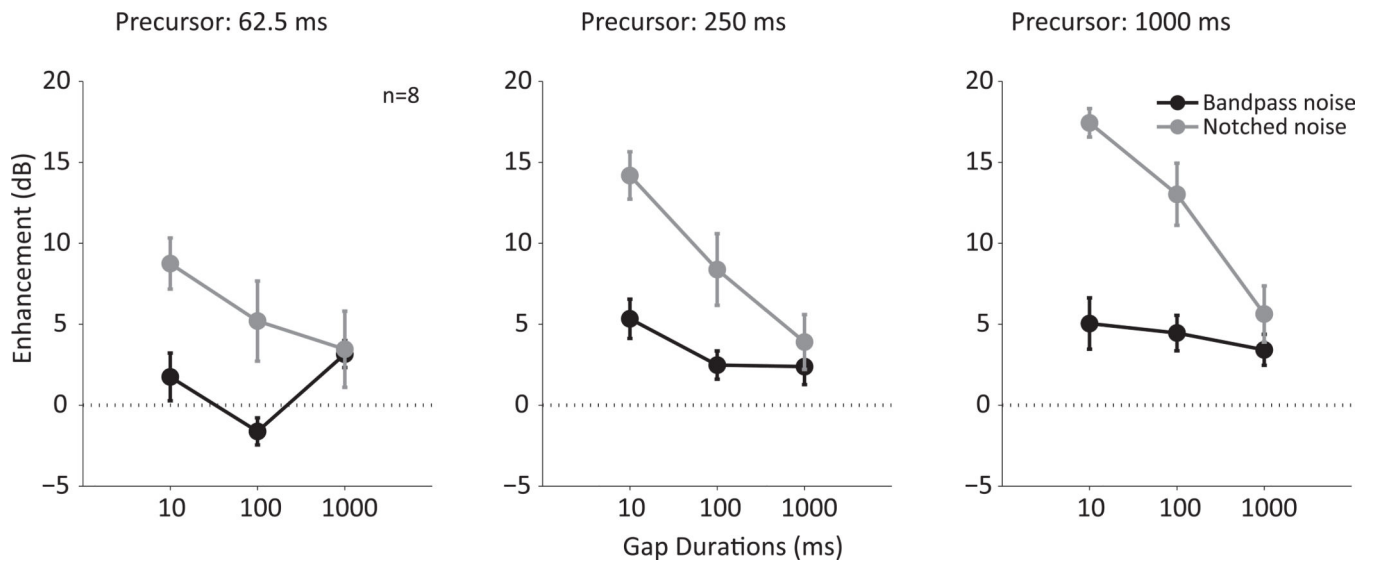
Figure 2.
Thresholds measured in the no-precursor condition in two tasks in terms of target-to-masker ratio (TMR) for individual listeners. The level of 0 dB TMR is equivalent to a target level of 45 dB SPL, the same level as the masker components. Each ellipse represents the data from one participant. The coordinates of each ellipse center are the mean thresholds in the two tasks and the height and width of each ellipse represent the standard deviations across the last four runs for each participant for each task. Black solid ellipses represent participants who started with the Present/Absent task and then completed the Up/Down task, whereas the grey dashed ellipses represent participants who completed the two tasks in the opposite order.
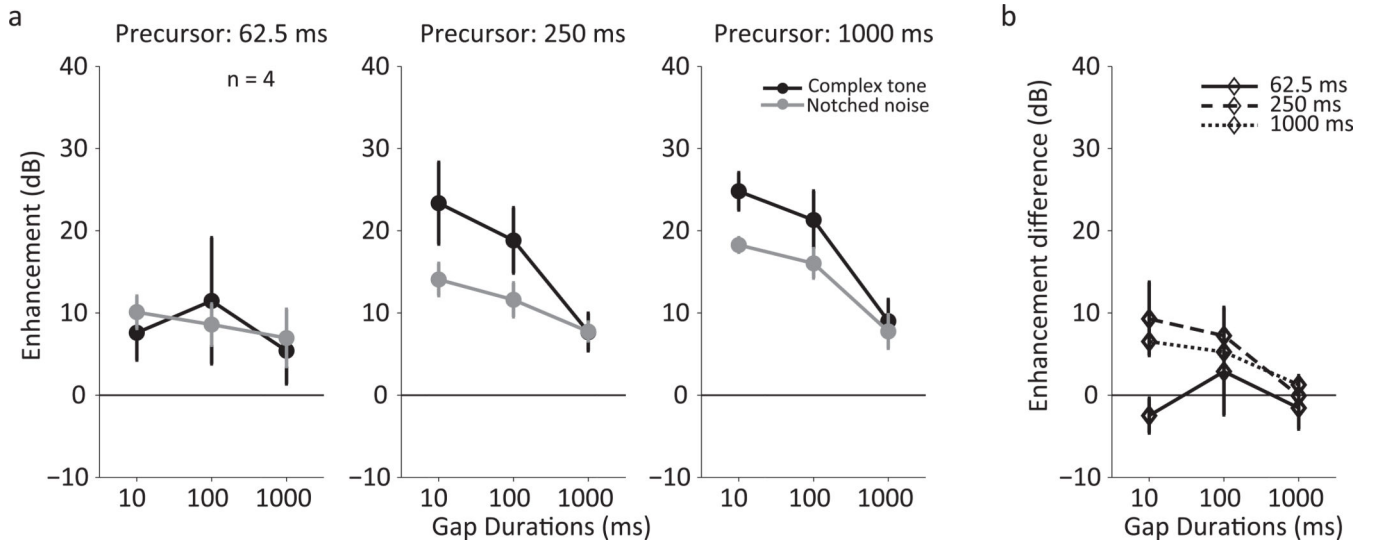
Figure 3.
Time course of enhancement in the Present/Absent and Up/Down tasks. Each panel shows the mean enhancement, defined as the difference between the thresholds measured with and without the precursor, plotted as a function of the gap duration for one of the three precursor durations. Black and grey symbols represent data from the Present/Absent task and Up/Down task, respectively. Error bars represent ±1 standard error.
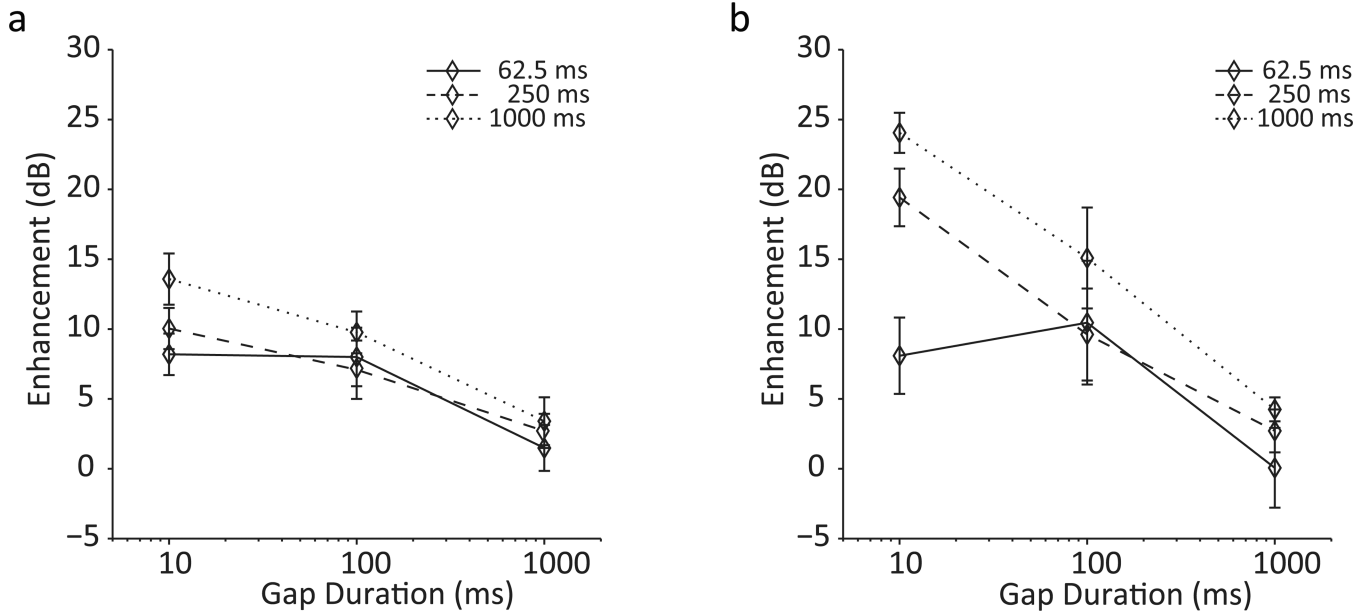
Figure 4.

Time courses of the mean enhancement of eight listeners for two precursor types (notched noise and bandpass noise). In each panel, the enhancement is plotted as a function of the gap duration for one of the three precursor durations, defined as the difference between the thresholds measured with and without the precursor. Black and grey symbols represent data from the bandpass-noise precursor and the notched-noise precursor, respectively. Error bars represent ±1 standard error of the mean.

Figure 5.
Time courses of the mean enhancement of four listeners for two precursor types (inharmonic complex and notched noise). a, In each panel, the enhancement is plotted as a function of gap duration for one precursor duration, defined as the difference between the thresholds measured with and without the precursor. Different colors represent the notched-noise precursor and complex-tone precursor separately. b, The enhancement difference between the complex-tone precursor and the notched-noise precursor in each conditions. Different symbols represent different precursor durations. Error bars represent ±1 standard error of the mean.

a

b

Figure 6.
Time courses of the mean enhancement defined as the difference between the thresholds of the precursor without spectral notch and the precursor with spectral notch. Different precursor durations are distinguished by different line styles. Error bars represent ±1 standard error. a, Time courses of enhancement for noise precursors in Experiment 2. The enhancement is defined as the difference between thresholds of notched-noise precursors and bandpass-noise precursors. The mean was calculated across 8 participants. b, Time courses of enhancement for inharmonic complex-tone precursors in Experiment 1. The enhancement here is defined as the difference between thresholds of inharmonic complex-tone precursors with or without the target tone. The mean was averaged across 5 participants.

Table 1

Slopes at threshold of the estimated psychometric functions for individual participants. The slopes with the precursor were derived from the condition with the 1000-ms precursor and 10-ms gap.

Subject Number	Present/Absent		Up/Down	
	No precursor	With precursor	No precursor	With precursor
1	2.24	1.03	.21	1.20
2	.99	1.50	2.86	1.02
3	.75	2.08	1.25	2.00
4	1.54	1.60	2.30	2.88
5	1.14	1.42	1.10	1.70
6	1.20	1.57	.57	1.45
7	3.40	.75	.91	1.15
8	1.62	1.12	4.54	1.76
9	1.50	.57	.22	1.18
10	1.24	.74	2.08	.42
11	1.34	2.20	.44	1.97
12	.44	1.14	1.26	1.76
13	1.48	1.61	2.53	1.67
14	.79	.73	1.57	1.07
Mean (SD)	1.41 (0.72)	1.29 (0.51)	1.56 (1.21)	1.52 (0.59)