Research Articles

Auditory Enhancement of Increments in Spectral Amplitude Stems from More Than One Source

SAMUELE CARCAGNO[1], CATHERINE SEMAL[1], AND LAURENT DEMANY[1]

[1]University Bordeaux and CNRS, INCIA, 146 rue Léo Saignat, 33076, Bordeaux, France

ABSTRACT

A component of a test sound consisting of simultaneous pure tones perceptually "pops out" if the test sound is preceded by a copy of itself with that component attenuated. Although this "enhancement" effect was initially thought to be purely monaural, it is also observable when the test sound and the precursor sound are presented contralaterally (i.e., to opposite ears). In experiment 1, we assessed the magnitude of ipsilateral and contralateral enhancement as a function of the time interval between the precursor and test sounds (10, 100, or 600 ms). The test sound, randomly transposed in frequency from trial to trial, was followed by a probe tone, either matched or mismatched in frequency to the test sound component which was the target of enhancement. Listeners' ability to discriminate matched probes from mismatched probes was taken as an index of enhancement magnitude. The results showed that enhancement decays more rapidly for ipsilateral than for contralateral precursors, suggesting that ipsilateral enhancement and contralateral enhancement stem from at least partly different sources. It could be hypothesized that, in experiment 1, contralateral precursors were effective only because they provided attentional cues about the target tone frequency. In experiment 2, this hypothesis was tested by presenting the probe tone before the precursor sound rather than after the test sound. Although the probe tone was then serving as a frequency cue, contralateral precursors were again found to produce enhancement. This indicates that contralateral enhancement

cannot be explained by cuing alone and is a genuine sensory phenomenon.

Keywords: auditory enhancement, perceptual pop-out, neural adaptation, spectral processing, intensity discrimination

INTRODUCTION

In a test sound consisting of a sum of simultaneous tones, one tone can be made to "pop out" perceptually by preceding the test sound with a copy of itself without the target tone (Schouten 1940; Viemeister 1980). This phenomenon, known as the "enhancement" effect, occurs more generally whenever spectrally local energy is added to a sound with energy in multiple spectral bands (Summerfield et al. 1987; Hartmann and Goupell 2006; Serman et al. 2008; Erviti et al. 2011). Enhancement is therefore helpful for the detection of changes in complex auditory environments. The effect is strongest when the test sound and its "precursor" are contiguous in time and declines gradually as a function of the duration of the inter-stimulus interval (ISI) between the precursor and test sounds (Wilson 1970; Viemeister 1980; Summerfield et al. 1984; Carlyon 1989). Using rather long precursors (>2 s), significant enhancement has been measured for ISIs up to 6.4 s (Viemeister 1980). Enhancement has often been interpreted in terms of some form of selective neural adaptation: the precursor may adapt neural units responding to the non-target components of the test sound more than neural units responding to the target component (Viemeister 1980); the benefit of the precursor may also originate from a reduction in the inhibition of the target

Correspondence to: Samuele Carcagno · University Bordeaux and CNRS, INCIA · 146 rue Léo Saignat, 33076, Bordeaux, France. Telephone: +33-5-57571651 email: samuele.carcagno@u-bordeaux2.fr

component by the non-target components (Viemeister and Bacon 1982; Nelson and Young 2010; Byrne et al. 2011). An alternative hypothesis is that the internal level of the non-target components, and as a consequence their ability to inhibit the target component, are reduced by the activation of the medial olivocochlear reflex (MOCR), which acts directly at the cochlear level (Strickland 2004, 2008). Non-sensory mechanisms may also contribute to enhancement: the precursor could provide cues that help segregate the target component from the non-target components (Richards et al. 2004).

Several studies have failed to find enhancement effects when the precursor was presented contralaterally to the test sound, that is, when the two sounds were presented to opposite ears (Viemeister 1980; Summerfield et al. 1984; Carlyon 1989; Serman et al. 2008). However, Richards et al. (2004), Erviti et al. (2011), and Kidd et al. (2011) have obtained enhancement effects for contralateral precursors. Contralateral precursors have also been shown to shift phonetic boundaries depending on their frequency content, a phenomenon that may reflect enhancement (Holt and Lotto 2002). The contralateral effects observed in these studies were weaker than the ipsilateral effects. Nonetheless, ipsilateral and contralateral enhancement may be based on the same neural mechanism, occurring at a single locus in the auditory system: at this locus, contralateral input could be simply less represented than ipsilateral input. A second possibility is that ipsilateral and contralateral enhancement are based on qualitatively different processes (e.g., adaptation vs. cuing). A third and final possibility is that enhancement results from a single mechanism (e.g., adaptation) taking place at multiple levels of the auditory pathways; it could be that at some levels this mechanism operates exclusively on monaural input while at other levels it operates on input from both ears. The present study was intended to test the first of these three hypotheses, i.e., the idea that ipsilateral and contralateral enhancement originate from a single mechanism at the same neural locus. To this end, we measured the decay of ipsilateral and contralateral enhancement as a function of the ISI between precursor and test sounds. If the decay differs between these two forms of enhancement, then the first of the three hypotheses stated above must be wrong. An additional experiment tested whether the contralateral enhancement that we observed could be explained by cuing effects, resulting from non-sensory processes.

METHODS

Participants

Ten listeners, including authors LD (listener L2) and SC (listener L3), were tested in experiment 1. Two of them did not show much contralateral enhancement in a preliminary phase of the experiment (see "Methods" below) and were not tested further. Six listeners took part in experiment 2; five of these six listeners had also taken part in experiment 1. The listeners ranged in age between 20 and 58 years (mean=28). They all had absolute pure-tone thresholds below 20 dB HL for both ears at octave frequencies from 250 to 4,000 Hz. All listeners, except one that took part only in the first experiment, had prior experience in psychoacoustic tasks. Each listener (except authors LD and SC) gave written informed consent and was paid an hourly wage.

Experiment 1

The aim of this experiment was to determine if the decay of enhancement as a function of the ISI between the precursor and test sounds is the same for ipsilateral and contralateral precursors. A spectrographic representation of the stimuli is displayed in Figure 1A. The precursor and test sounds were inharmonic "chords" composed of ten synchronous pure tones spaced by intervals of 350 cents (1 cent=1/100 semitone=1/1200 octave; an interval of 350 cents thus represents an upward frequency change of about 22 %). On each trial, the ten tones had identical frequencies in both chords. However, the chords were randomly transposed from trial to trial within a 4-octave range, from 250 to 4,000 Hz (frequency being scaled logarithmically). The durations of the precursor and test chords were, respectively, 500 and 100 ms, including 10-ms raised-cosine onset and offset ramps. In the *No-Change* conditions, each component of the precursor and test chords had a nominal level of 55 dB SPL. In the *Change* conditions, one component of the precursor was attenuated by an amount of ΔL dB, in order to enhance the corresponding component of the test chord. This component was randomly selected among the eight "inner" components of the precursor chord.

Enhancement was measured by means of a "present/absent" task (Demany and Ramos 2005). On each trial, 500 ms after the test chord, listeners were presented with a single probe tone. On "present" trials, the probe tone was identical to one of the test chord components; this component was the target of enhancement in the *Change* conditions, and it was selected randomly among the eight inner components of the test chord in the *No-Change* conditions. On "absent" trials, the frequency of the probe tone was geometrically centered between the frequencies of two randomly chosen neighboring components of the test chord. "Present" and "absent" trials had equal a priori probabilities. Listeners had to indicate if the test chord contained or not a tone identical to the probe. Performance in the task depends on the ability to hear out the components of the test chord. The
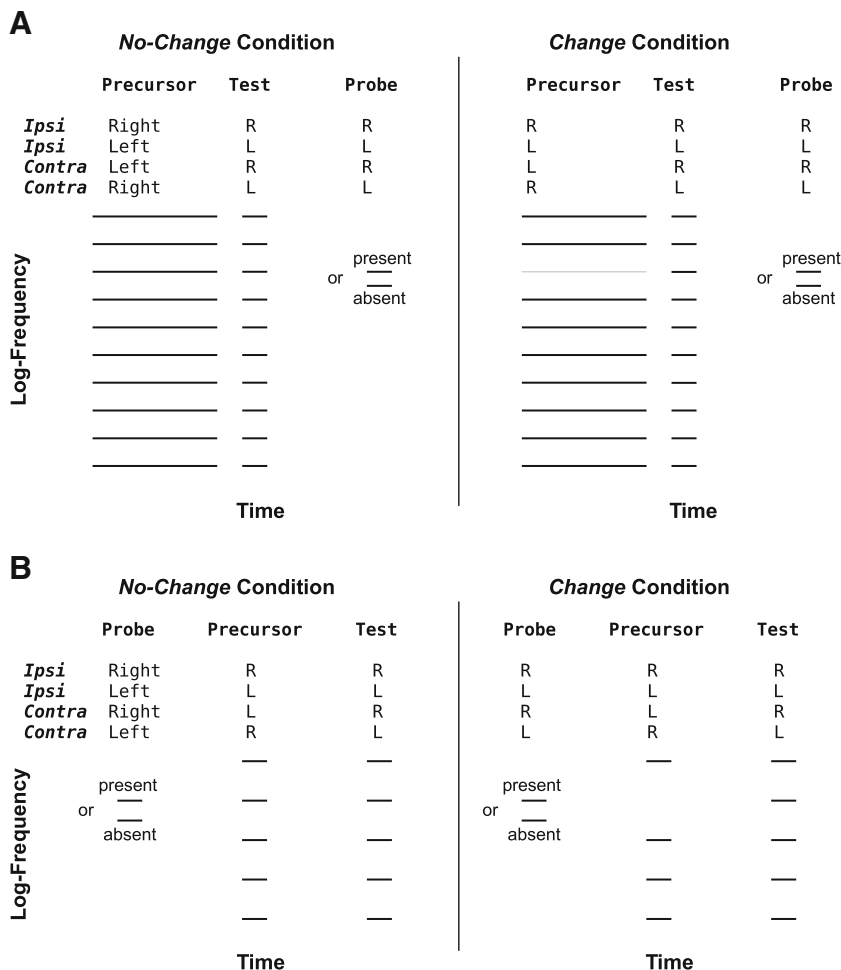
FIG. 1. **A** Spectrographic representation of the stimuli used in experiment 1. In the *No-Change* conditions, all components of the precursor chord had the same intensity. In the *Change* conditions, one randomly chosen component (indicated by the *gray line* in this example), had a lower intensity. The test chord followed the precursor chord after a 10, 100, or 600 ms silent ISI, in different blocks of trials. Listeners had to indicate whether the probe, presented 500 ms after the end of the test chord, had the same pitch as one of the test chord components or not. The precursor and test chords were either presented to the same ear (*Ipsi*) or to opposite ears (*Contra*); the four possible laterality combinations are indicated above the stimuli. The precursor-test sequence as well as the position of the probe relative to the test chord were varied across trials (see text for details). Possible positions of the probe for both "present" and "absent" trials are illustrated in the Figure. Note that in the experiment only one probe was presented on each trial. **B** Spectrographic representation of the stimuli used in experiment 2. The task for the listener was the same as in experiment 1. The main difference with experiment 1 was that this time the probe was presented before the precursor and test sounds.

ability to hear out the relevant component of the test chord in the *Change* conditions should depend on the degree of enhancement. Note that the "present/absent" task has been successfully used before to measure enhancement (Erviti et al. 2011). As this task requires an explicit identification of the enhancement target in the test sound, it may assess the perceptual pop-out more directly than tasks in which the listener is only required to discriminate between test sounds with and without the enhancement target.

The test chord and the probe were always presented monaurally to the same ear (either the right or the left ear). The precursor was also presented monaurally, either to the same ear as the test chord (conditions *Ipsi*) or to the opposite ear (conditions *Contra*). Each time a sound (precursor, test or probe) was presented to one ear, a synchronous burst of pink noise filtered between 10 and 5,000 Hz was presented to the opposite ear; this noise had a spectrum level of 10 dB at 1 kHz. The noise served two purposes: first, it masked potential interaural cross-talk; second, in the *Contra* conditions, it made the task subjectively more comfortable by decreasing the shift in the subjective lateralization of the chords (Warren and Bashford 1976). Notice that contralateral enhancement has

been obtained also without using such noise (Richards et al. 2004; Kidd et al. 2011).

Performance in the "present/absent" task was measured in terms of the sensitivity index d' (Green and Swets 1974), and enhancement was defined as the d' difference between the *Change* and *No-Change* conditions. In order to assess the temporal decay of enhancement, three ISIs between the precursor and test chords were used: 10, 100, and 600 ms. It was desirable to obtain similar magnitudes of enhancement for ipsilateral and contralateral precursors at the shortest ISI. To this aim, we used different ΔL values in the *Ipsi* and *Contra* conditions. For each listener, they were adjusted during a preliminary experimental phase. This preliminary phase, lasting for about 3–4 h, served also to familiarize the listeners with the stimuli and procedures of the experiment; at the end of this phase, listeners had achieved stable performance in the task. The selected ΔL values are displayed in Table 1.[1] For all listeners, the *Ipsi* ΔL value needed to

---

[1] During the experiment proper, in the *Ipsi* conditions, we actually used two slightly different ΔL values for each listener. We then selected a posteriori the ΔL value that minimized the difference in enhancement between the *Ipsi* and *Contra* conditions for the 10-ms ISI.

TABLE 1

Attenuation of the precursor component corresponding in
frequency to the target of enhancement, in experiment 1

Listener	ΔL Ipsi (dB)	ΔL Contra (dB)
L1	3	40
L2	5	30
L3	3	9
L4	3.5	40
L5	2.5	∞
L6	2.5	40
L7	2.5	∞
L8	3	∞

achieve a certain level of performance was consider-
ably smaller than the corresponding *Contra* ΔL value.
This is consistent with the results of previous studies
that found contralateral enhancement to be weaker
than ipsilateral enhancement (e.g., Richards et al.
2004; Erviti et al. 2011; Kidd et al. 2011).

There were 12 conditions in all, resulting from the
combination of four precursor types (*No-Change Ipsi,
No-Change Contra, Change Ipsi, Change Contra*) and
three ISIs (10, 100, and 600 ms). Each of the 12
conditions was divided in two sub-conditions: one in
which the test chord was presented to the left ear and
another in which the test chord was presented to the
right ear. In each experimental session, listeners
completed one block of 40 trials per sub-condition;
these 24 blocks were randomly ordered. The experi-
ment was run in five sessions, for a total of 200 trials
per sub-condition.

Experiment 2

In the *Change* conditions of experiment 1, the precursor
chord potentially provided information about the target
tone frequency because this frequency always coincided
with the frequency of the attenuated precursor compo-
nent. Although the attenuated component in the precur-
sor was probably inaudible, the notch in the precursor
spectrum resulting from the attenuation may have pointed
the listeners to the target frequency. It is conceivable that
the *Change* conditions were advantageous due to the
corresponding "cuing" rather than due to sensory en-
hancement. In experiment 2, which was a simplified
variant of experiment 1, we wished to test this cuing
hypothesis. A crucial difference between the two experi-
ments was that, in experiment 2, the probe tone was
presented *before* the precursor-test sequence rather than
after it (see Fig. 1B). We reasoned that, in the "present/
absent" task, the probe tone would then serve in itself as a
cue, not less informative than the attenuation of one of the
precursor chord components. As a result, according to the
cuing hypothesis, no difference in performance should be
observed between *Change* and *No-Change* conditions.

The "present/absent" task is easier when the probe
tone precedes, rather than follows, the test chord. In
order to avoid ceiling effects, the stimuli used in
experiment 2 were shorter than in experiment 1. All
sounds had a duration of 100 ms, including 10-ms
raised-cosine onset and offset ramps. The ISI between
probe and precursor was 400 ms while the ISI
between precursor and test was 500 ms. The precursor
and test chords were formed by summing five pure
tones spaced by intervals of 500 cents (corresponding
to upward frequency changes of about 33 %). Within
a trial, these five tones had the same frequencies in
both chords, but the chords were randomly trans-
posed from trial to trial within a 3.1-octave range,
from 250 to 2200 Hz. In the *No-Change* conditions,
each component of the precursor and test chords had
an intensity of 55 dB SPL, while in the *Change*
conditions one component of the precursor was
infinitely attenuated.

The "present/absent" task was set up in the same
manner as in experiment 1. Again, only the inner
components of the test chord could be enhanced. On
"absent" trials, as before, the probe was always positioned
halfway between two neighboring components of the test
chord; this time, however, the target of enhancement was
always one of the two components of the test chord which
was closest to the probe in frequency. As in experiment 1,
the test chord and the probe were always presented
monaurally to the same ear; the precursor was also
presented monaurally, either to the same ear as the test
chord (conditions *Ipsi*) or to the opposite ear (conditions
Contra). The stimuli were combined with bursts of pink
noise exactly as before. There were in total four
conditions: *No-Change Ipsi, No-Change Contra, Change Ipsi*,
and *Change Contra*. Each of these four conditions was
divided in two sub-conditions: one in which the test
sound was presented to the left ear and another in which
the test sound was presented to the right ear. In each
experimental session, listeners completed one block of
50 trials per sub-condition; the eight sub-conditions were
randomly ordered. The experiment was run in four
sessions, for a total of 200 trials per sub-condition.
Listeners were given at least 1 h of practice before data
collection began. Additional practice sessions were run if
d' in the *Change Ipsi* condition was below one.

Apparatus and procedures

Listeners were seated in a double-walled sound
attenuating booth (Gisol, Bordeaux). The stimuli
were generated digitally with 16-bit resolution and a
44.1-kHz sampling rate on a PC housed outside the
booth. They were played through an ECHO Gina24
digital-to-analog converter and presented via Sennhe-
iser HD650 headphones. Listeners gave their
responses by pressing a button on the numeric keypad

of the PC. Feedback was provided at the end of each trial by means of a colored light (green if the response was correct, red otherwise).

RESULTS

Experiment 1

Figure 2 displays the mean d' value obtained in each condition of experiment 1. In the *No-Change* conditions, performance was poor but above chance for each ISI, both when the precursor was presented ipsilaterally to the test chord and contralaterally to the test chord. Performance in the *Change* conditions was generally higher than in the *No-Change* conditions and strongly dependent on ISI. Figure 3 shows enhancement magnitude, computed as the d' difference between the *Change* and *No-Change* conditions, as a function of ISI and precursor laterality (*Ipsi* or *Contra*). The mean data (top panel) indicate that while enhancement magnitude was similar for the *Ipsi* and *Contra* conditions at an ISI of 10 ms, the decay of enhancement as a function of ISI was sharper in the *Ipsi* conditions than in the *Contra* conditions. In the *Ipsi* case, on average, enhancement decreased by about 50 % as the ISI increased from 10 to 100 ms, and by 90 % as the ISI increased from 10 to 600 ms. In the *Contra* case, enhancement decreased by about 10 % as the ISI increased from 10 to 100 ms and by about 50 % as the ISI increased from 10 to 600 ms. An inspection of the individual listeners' data (bottom panels) reveals that enhancement decayed more rapidly in the *Ipsi* case than in the *Contra* case for the majority of listeners. However, for three listeners (L3, L7, and L8), enhancement decayed with a similar time course in the *Ipsi* and *Contra* cases. Such inter-individual differences are not uncommon in studies of enhancement, even in
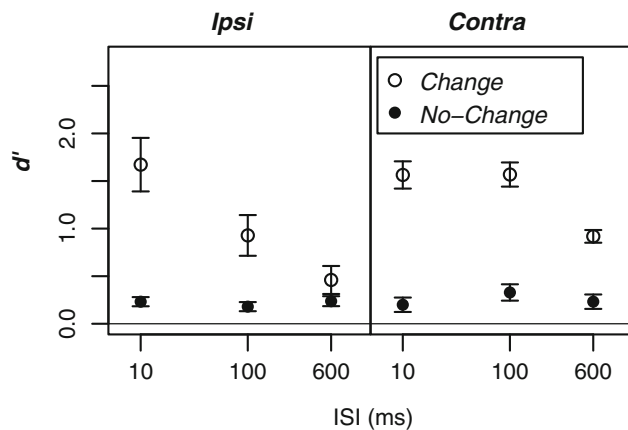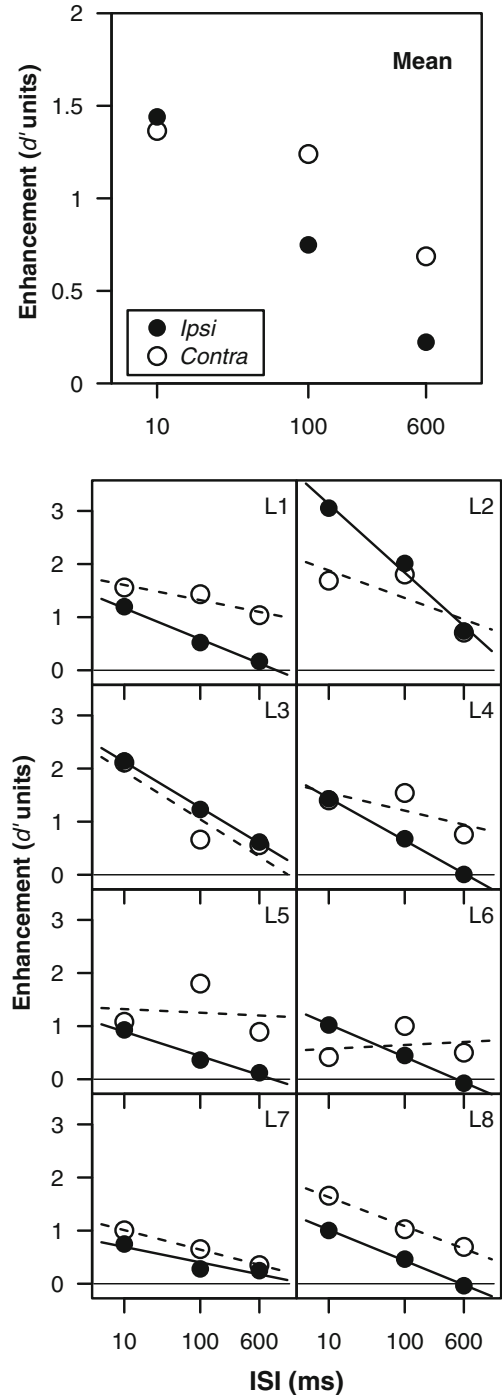

FIG. 3. Enhancement values obtained in experiment 1 as a function of ISI for ipsilateral and contralateral precursors. The *upper panel* presents the mean data and the *lower panels* present the data of the individual listeners. For each data point, enhancement is defined as the difference between the d' scores obtained in the corresponding *Change* and *No-Change* conditions.

the absence of stimulus roving and for listeners with tens of hours of practice in the task (e.g., McFadden and Wright 1990).

Enhancement has been previously modeled as decreasing linearly as a function of log ISI (Wilson 1970; Viemeister 1980). Accordingly, we quantified the decay

FIG. 2. Mean d' values obtained in experiment 1 as a function of ISI, for the *Change* and *No-Change* conditions. The *left panel* shows the results for the *Ipsi* conditions, and the *right panel* shows the results for the *Contra* conditions. *Error bars* represent ±1 standard error of the mean.

of enhancement by measuring the slopes of least-squares lines fitted to the individual listeners' data using a log scale for the ISI. A paired t test revealed that the slopes were significantly different for the *Ipsi* and *Contra* conditions [$t(7)=2.768$, $p=0.028$]. However, while the log ISI function provided excellent fits for the *Ipsi* data (mean goodness-of-fit $R^2=0.97$), it was less successful in the *Contra* case (mean $R^2=0.61$). The values of R^2 differed significantly between the two conditions [$t(7)=2.461$, $p=0.043$], suggesting that the decay of enhancement has a different *shape* for ipsilateral and contralateral precursors.

We also analyzed enhancement magnitude using a repeated-measures analysis of variance (ANOVA) with precursor laterality (*Ipsi* or *Contra*) and ISI (10, 100, 600 ms) as within-subject factors. This analysis did not assume an exponential decay function for enhancement. The main effect of laterality was not significant [$F(1, 7)=3.064$, $p=0.123$], while the main effect of ISI [$F(2, 14)=24.251$, $p<0.001$] and most importantly the interaction between precursor laterality and ISI [$F(2, 14)=5.589$, $p=0.016$] were both significant.

These results are consistent with the idea that ipsilateral and contralateral enhancement stem from different sources. However, an alternative interpretation of our data had to be considered. It might be hypothesized that, for the shortest ISI, contralateral precursors failed to produce good performance in the *Change* condition because listeners' attention could not be focused optimally on the test sounds due to the rapid stimulus shift between the two ears. As this putative attentional difficulty should have decreased as the ISI increased, our data for contralateral precursors would not reflect the "true" temporal decay of contralateral enhancement. That does not seem to be a likely hypothesis, because trials were organized in blocks during which the precursor and test chords were presented at fixed ears; consequently, shifts in stimulus lateralization could be systematically anticipated by the listeners. In addition, the duration of the precursor chords (500 ms) was such that the time at which a test chord would be presented was optimally predictable. However, the precursor may have exogenously captured the listeners' attention, and this automatic attentional capture may take time to subside. In the auditory domain, Spence and Driver (1994) found exogenous attentional effects of that kind only in tasks requiring spatial judgments, not in tasks requiring pitch judgments. Nonetheless, in order to completely rule out the hypothesis that performance in the *Contra* condition at the 10-ms ISI was limited by the need to swiftly shift attention between the ears, we ran a control experiment.

In this control experiment, listeners performed a "present/absent" task in which they had to compare a monaural probe tone with a test chord presented to the opposite ear. The procedure was identical to that used in the *Contra* condition of experiment 1, except for the following differences. On each trial, only two stimuli were presented: a 500-ms probe tone and a 100-ms test chord. The probe preceded the test chord and was presented at a level of 45 dB SPL. The ISI between the probe and the test chord was either 10 or 100 ms, in different blocks of trials. The components of the test chord had a level of 55 dB SPL, except for one component (T), which was presented at $55+\Delta L$ dB SPL (information on ΔL is provided below). T was chosen randomly among the eight inner chord components, and on "present" trials T had the same frequency as the probe. The task thus mimicked the *Contra* condition of experiment 1, but in the absence of enhancement (which was replaced by a physical increase in the level of the target component). As in experiment 1, the most salient component of the test chord had to be identified among the other components soon after the presentation of a 500-ms long contralateral sound. If listeners had difficulties in performing this task because it requires a sudden switch of attention between the ears, they should have performed better when the ISI between the probe and the test chord was 100 ms than when it was 10 ms.

Six of the listeners who had participated in experiment 1 were tested in the control experiment; the remaining two listeners were not available at the time of testing. The experiment proper was preceded by a preliminary session during which, for each listener, ΔL was adjusted in order to obtain a d' of about 2- for the 10-ms ISI. The selected ΔL values had a mean of 9.5 dB and ranged from 7 to 12 dB. Each listener performed 400 trials in each of the two ISI conditions.

The results of the control experiment are displayed in Figure 4, using digits to represent the performance of individual listeners and dots for the average
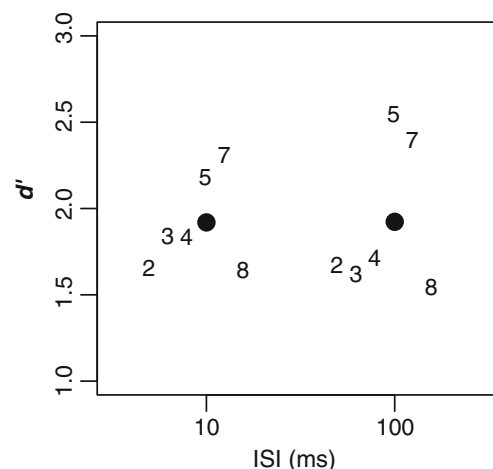

FIG. 4. Results of the control experiment. Digits represent individual listeners (labeled as in Fig. 3). The *filled symbols* represent the mean d' values across listeners.

performance at each ISI. Performance was almost exactly the same in the 10-ms condition (average d'= 1.919) and the 100-ms condition (average d'=1.923). This strongly suggests that, in the *Contra* condition of experiment 1, our measurement of enhancement decay was not biased by attentional factors unrelated to enhancement itself.

Experiment 2

Figure 5 displays the d' values obtained in the four conditions of experiment 2. Despite large interindividual differences in the overall level of performance, a clear pattern of results emerges with respect to the effect of precursor type. All listeners performed better in the *Change* than in the *No-Change* condition for both the ipsilateral and the contralateral precursors, except listener L9, who showed no contralateral enhancement. A repeated-measures ANOVA with precursor type (*Change, No-Change)* and precursor laterality (*Ipsi, Contra*) as within-subjects factors revealed a significant main effect of precursor type [$F(5, 1)$=17.13, p=0.009], but no main effect of precursor laterality or interaction between precursor type and laterality [p>0.15]. Paired t tests confirmed that performance was significantly higher in the *Change* condition than in the *No-Change* condition for both ipsilateral [$t(5)$=4.25, p=0.009] and contralateral [$t(5)$=3.33, p=0.021] precursors.

DISCUSSION

In experiment 1, we found that the time course of enhancement was different for ipsilateral and contralateral precursors. Enhancement appears to decay more rapidly, as a function of ISI, for an ipsilateral precursor than for a contralateral precursor. This result strongly suggests that different sources contribute to ipsilateral and contralateral enhancement: either different neural processes or the same neural process occurring at different sites with different time courses. Before discussing in more detail the possible mechanisms of enhancement, it is important to consider the potential role of the MOCR in the perceptual effects reported here.

There is evidence that the MOCR is involved in a phenomenon akin to enhancement, "overshoot": the improvement in detectability of a brief tone as the tone is delayed from the onset of a broadband masker (Strickland 2004, 2008; Jennings et al. 2011). The MOCR is mediated by the efferent system and operates directly at the level of the cochlea by modifying the action of the outer hair cells. The MOCR may cause enhancement by decreasing the cochlear gain in the non-target frequency regions

following the presentation of the precursor. Since the MOCR is activated by both ipsilateral and contralateral sounds (Guinan 2011), it could potentially explain both ipsilateral and contralateral enhancement. However, in order to produce enhancement, the cochlear gain reduction caused by the MOCR should be frequency specific; that is, it should be maximal for the frequency regions stimulated by the precursor. Recent experiments that have investigated the frequency tuning of the MOCR in humans using otoacoustic emissions do not provide much evidence that the MOCR is frequency specific. These studies indicate that the MOCR is broadly tuned and that maximum MOCR effects are often produced by elicitors well below or above the probe frequency (Lilaonitkul and Guinan 2009a; Walsh et al. 2010; Lilaonitkul and Guinan 2012; see Guinan 2011, for a review). At a given frequency, the MOCR reflects spatial summation of responses to elicitors over most of the cochlea (Lilaonitkul and Guinan 2009b). In light of these data, MOCR activation cannot account for the frequency-specific enhancement of the target tone in a chord composed of equal-amplitude tones. Further evidence that the MOCR does not play a crucial role in enhancement comes from studies demonstrating that cochlear-implant users, for whom acoustic stimulation bypasses the cochlea, show enhancement effects (Wang et al. 2012; Goupell and Mostardi 2012). Moreover, the time courses of the ipsilateral and contralateral MOCR are similar (Backus and Guinan 2006). Therefore, even if the MOCR were involved in enhancement, it would be unlikely to account for the fact that contralateral enhancement decays more slowly than ipsilateral enhancement, as shown by experiment 1.

The role of cuing in enhancement

It is customary to distinguish two forms of simultaneous masking: *energetic* masking and *informational* masking. Energetic masking arises from peripheral limitations, in particular the frequency selectivity of the cochlea. Informational masking arises from more central limitations and may be based on a completely different mechanism or set of mechanisms (see Durlach et al. 2003; Durlach 2006, and Kidd et al. 2007, for reviews). Erviti et al. (2011) hypothesized that contralateral enhancement reflects a release from informational masking only while ipsilateral enhancement can reflect a release from both energetic and informational masking. This would explain why only some of the attempts to find contralateral enhancement were successful. In the studies that found contralateral enhancement (Richards et al. 2004; Erviti et al. 2011; Kidd et al. 2011; and the current study), the target tones and/or the surround-

ing masking tones were randomly varied in frequency across trials and the target tones were well separated in frequency from the masking tones. These circumstances increase the contribution of informational masking relative to that of energetic masking (Kidd et al. 2007).

Several factors can potentially contribute to informational masking (Durlach et al. 2003; Kidd et al. 2007), notably signal or masker uncertainty (Richards and Neff 2004) and signal/masker similarity (Kidd et al. 2002). How could an appropriate precursor, then, reduce informational masking? One possibility is that

the precursor provides a cue about the target tone frequency. The precursors reducing masking contain a spectral notch. Performance in the "present/absent" task could be better because listeners focus their attention on the frequency region of the test sound that corresponds to the notch in the precursor. Crucially, this explanation does not require that the target component stand out perceptually among the other components of the test sound; in other words, masking could be reduced in the absence of a truly sensory enhancement.

However, the results of experiment 2 are inconsistent with this idea. In the *No-Change* conditions of experiment 2, the probe tone was providing a task-relevant cue which was presumably even more precise than the cue provided by the precursor notch in the *Change* conditions. Nevertheless, listeners' performance was still better in the *Change* conditions than in the *No-Change* conditions.

Another possibility is that the precursor improves performance by reducing masker uncertainty. This reduction in masker uncertainty could then be exploited for "masker minimization" (Durlach et al. 2003), a putative strategy in which the listener uses his/her knowledge of the masker components to filter them out. However, according to the results of Richards et al. (2004), knowledge of the masker component frequencies is not sufficient to explain the benefit of a notched precursor. These authors measured the detectability of a 1-kHz signal presented in a multitone random masker using two possible contralateral precursors: a preview of the masker alone or a preview of the masker-plus-signal. Since the signal frequency was fixed in their study, these two types of precursors contained the same information about the masker component frequencies. Nonetheless, listeners performed better in the task when they were provided with a preview of the masker alone. Overall, these results indicate that cuing of the signal frequency or the masker component frequencies cannot explain the benefit provided by a precursor with a spectral notch at the signal frequency. It should be noted that in the study of Richards et al. (2004), the advantage provided by a preview of the masker alone disappeared when the signal frequency varied from trial to trial. This is in contrast with the results of the current experiments and the study of Erviti et al. (2011), in which the advantage was present despite the fact that both the target and non-target frequencies varied across trials. The reason for this discrepancy is not clear. While the "present/absent" task requires the identification of the frequency region where the precursor and test sounds differ, the "same/different" task employed by Richards et al. (2004) requires only the detection of a change between precursor and test.
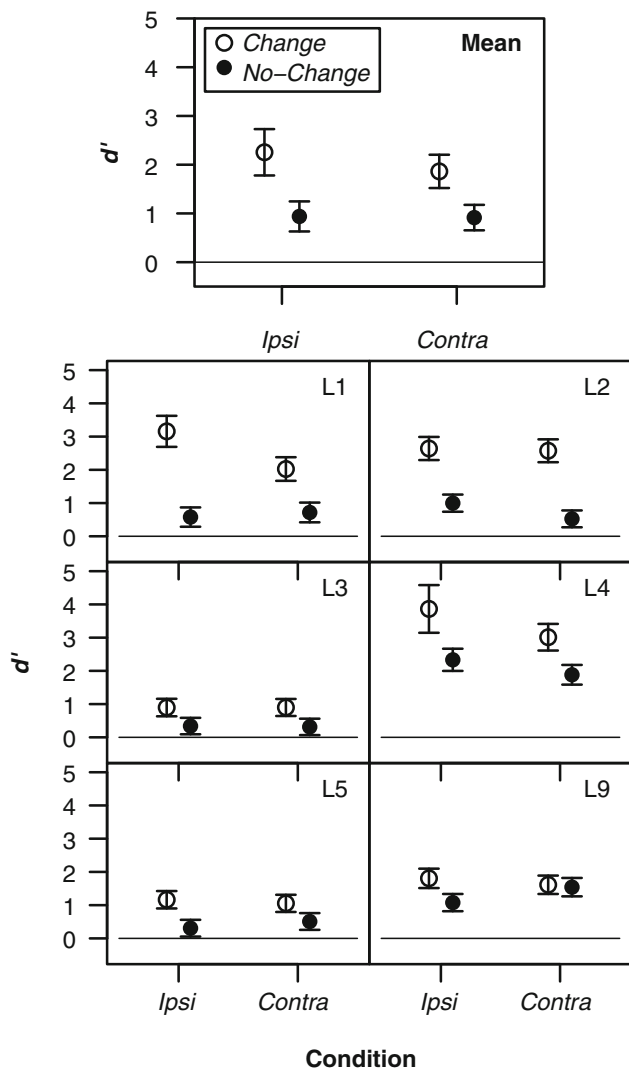

FIG. 5. Values of d' obtained in experiment 2 as a function of precursor laterality (*Ipsi* or *Contra*) for the *Change* and *No-Change* conditions. Listeners who participated also in experiment 1 are designated by the same labels as in Fig. 3. The *error bars* represent 95 % confidence intervals for the results of individual listeners (*bottom panels*) and ±1 standard error of the mean across listeners for the group data (*upper panel*).

Enhancement may reflect adaptation processes at multiple levels of the auditory pathways

The neurophysiological substrates of enhancement have been sought at different levels of the auditory pathways, including the auditory nerve (Palmer et al. 1995), the cochlear nucleus (Scutt and Palmer 2000), and the inferior colliculus (Nelson and Young 2010). Adaptation of the non-target components at the level of the auditory nerve may contribute to enhancement but cannot account for the gain in the target frequency region that has been demonstrated psychophysically (Viemeister and Bacon 1982; Wright et al. 1993; Byrne et al. 2011). Small gains in the target frequency region have been observed at the level of the cochlear nucleus (Scutt and Palmer 2000); they were limited to the initial portion (∼25 ms) of the response to the target. More robust gains in the target frequency region have been recently measured in the inferior colliculus (Nelson and Young 2010). In all these studies, the test sound immediately followed the precursor; thus, it is not known what is the decay of the neurophysiological effects as a function of the ISI between precursor and test sounds. This question is particularly important since enhancement has been measured psychophysically for ISIs up to 6.4 s, and by extrapolation based on curve fitting its decay may not be complete before 30 s in some cases (Viemeister 1980). Remarkably, stimulus-specific adaptation effects with longer time constants have been recently uncovered at several levels of the auditory system, including the inferior colliculus (Malmierca et al. 2009), the thalamus (Antunes et al. 2010), and the primary auditory cortex (Ulanovsky et al. 2003, 2004). These adaptation effects could potentially account for the enhancement observed psychophysically at long ISIs. Moreover, since they take place in central sites, they could in principle contribute to both ipsilateral and contralateral enhancement.

ACKNOWLEDGMENTS

REFERENCES

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

Backus BC, Guinan JJ (2006) Time-course of the human medial olivocochlear reflex. J Acoust Soc Am 119:2889–2904

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

Carlyon RP (1989) Changes in the masked thresholds of brief tones produced by prior bursts of noise. Hear Res 41:223–235

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

Durlach N (2006) Auditory masking: need for improved conceptual structure. J Acoust Soc Am 120:1787–1790

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

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

Goupell MJ, Mostardi MJ (2012) Evidence of the enhancement effect in electrical stimulation via electrode matching. J Acoust Soc Am 131:1007–1010

Green DM, Swets JA (1974) Signal detection theory and psychophysics. Krieger, New York

Guinan JJ (2011) Physiology of the medial and lateral olivocochlear systems. In: Ryugo DK, Fay RR (eds) Auditory and vestibular efferents. Springer, New York

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

Holt LL, Lotto AJ (2002) Behavioral examinations of the level of auditory processing of speech context effects. Hear Res 167:156–169

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

Kidd G Jr, Mason CR, Arbogast TL (2002) Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. J Acoust Soc Am 111:1367–1376

Kidd G Jr, Mason C, Richards VM, Gallun F, Durlach N (2007) Informational Masking. In: Yost WA, Popper AN, Fay RR (eds) Auditory perception of sound sources. Springer, New York, pp 143–189

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

Lilaonitkul W, Guinan JJ Jr (2009a) Reflex control of the human inner ear: a half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking. J Neurophysiol 101:1394–1406

Lilaonitkul W, Guinan JJ Jr (2009b) Human medial olivocochlear reflex: effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. J Assoc Res Otolaryngol 10:459–470

Lilaonitkul W, Guinan JJ (2012) Frequency tuning of medial-olivocochlear-efferent acoustic reflexes in humans as functions of probe frequency. J Neurophysiol 107:1598–1611

Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. J Neurosci 29:5483–5493

McFadden D, Wright BA (1990) Temporal decline of masking and comodulation detection differences. J Acoust Soc Am 88:711–724

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

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

Richards VM, Huang R, Kidd G Jr (2004) Masker-first advantage for cues in informational masking. J Acoust Soc Am 116:2278–2288

Richards VM, Neff DL (2004) Cuing effects for informational masking. J Acoust Soc Am 115:289–300

Schouten JF (1940) The residue: a new component in subjective sound analysis. Proc Koninkl Ned Akad Wetenschap 43:356–365

Scutt MJ, Palmer AR (2000) Physiological enhancement in cochlear nucleus using single tone precursors. Assoc Res Otolaryngol Abs:381

Serman M, Semal C, Demany L (2008) Enhancement, adaptation, and the binaural system. J Acoust Soc Am 123:4412–4420

Spence CJ, Driver J (1994) Covert spatial orienting in audition: exogenous and endogenous mechanisms. J Exp Psychol Human Percept Perform 20:555–574

Strickland EA (2004) The temporal effect with notched-noise maskers: analysis in terms of input-output functions. J Acoust Soc Am 115:2234–2245

Strickland EA (2008) The relationship between precursor level and the temporal effect. J Acoust Soc Am 123:946–954

Summerfield Q, Haggard M, Foster J, Gray S (1984) Perceiving vowels from uniform spectra: phonetic exploration of an auditory after effect. Percept Psychophys 35:203–213

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

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440–10453

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

Viemeister NF (1980) Adaptation of masking. In: van den Brink G, Bilsen FA (eds) Psychophysical, physiological and behavioural studies in hearing. Delft University Press, Delft, pp 190–199

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

Walsh KP, Pasanen EG, Mcfadden D (2010) Properties of a nonlinear version of the stimulus-frequency otoacoustic emission. J Acoust Soc Am 127:955–969

Wang N, Kreft H, Oxenham AJ (2012) Vowel enhancement effects in cochlear-implant users. J Acoust Soc Am 131:EL421–EL426

Warren RM, Bashford JA (1976) Auditory contralateral induction: an early stage in binaural processing. Percept Psychophys 20:380–386

Wilson JP (1970) An auditory after image. In: Plomp R, Smoorenburg GF (eds) Frequency analysis and periodicity detection in hearing. Sijthoff, Leiden, pp 303–315

Wright BA, Mcfadden D, Champlin CA (1993) Adaptation of suppression as an explanation of enhancement effects. J Acoust Soc Am 94:72–82