

Research Article

The Auditory Enhancement Effect is Not Reflected in the 80-Hz Auditory Steady-State Response

SAMUELE CARCAGNO,^{1,2} CHRISTOPHER J. PLACK,^{2,3} ARTHUR PORTRON,¹ CATHERINE SEMAL,¹
AND LAURENT DEMANY¹

¹*Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, Université de Bordeaux and CNRS, Bordeaux, France*

²*Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK*

³*School of Psychological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, M13 9PL, UK*

Received: 7 January 2014; Accepted: 4 April 2014; Online publication: 21 May 2014

ABSTRACT

The perceptual salience of a target tone presented in a multitone background is increased by the presentation of a precursor sound consisting of the multitone background alone. It has been proposed that this “enhancement” phenomenon results from an effective amplification of the neural response to the target tone. In this study, we tested this hypothesis in humans, by comparing the auditory steady-state response (ASSR) to a target tone that was enhanced by a precursor sound with the ASSR to a target tone that was not enhanced. In order to record neural responses originating in the brainstem, the ASSR was elicited by amplitude modulating the target tone at a frequency close to 80 Hz. The results did not show evidence of an amplified neural response to enhanced tones. In a control condition, we measured the ASSR to a target tone that, instead of being perceptually enhanced by a precursor sound, was acoustically increased in level. This level increase matched the magnitude of enhancement estimated psychophysically with a forward masking paradigm in a previous experimental phase. We found that the ASSR to the tone acoustically increased in level was significantly greater than the ASSR to the tone enhanced by the precursor sound. Overall, our results suggest that the enhancement effect cannot be explained by an amplified neural response at the level of the brainstem. However, an alternative possibility is that brainstem

neurons with enhanced responses do not contribute to the scalp-recorded ASSR.

Keywords: auditory enhancement, perceptual pop-out, ASSR, intensity coding

INTRODUCTION

The addition of a target tone in the second of two successive presentations of a multitone complex causes the added tone to “pop-out” perceptually, even if its level is not higher than the level of the other tones. This auditory “enhancement effect” has been demonstrated psychophysically in several ways. It has been shown that the threshold for detecting a target tone within a multitone background is lower when a precursor consisting of the multitone background alone precedes the target-plus-background mixture (Viemeister 1980; Byrne et al. 2011). Identification of the target frequency is also facilitated by the presentation of the precursor (Erviti et al. 2011; Carcagno et al. 2012; Demany et al. 2013). Furthermore, a target tone that is enhanced by the precursor has been shown to produce more forward masking of a subsequent short tone (Viemeister and Bacon 1982; Byrne et al. 2011). Overall, these findings suggest that within the auditory system the level of an enhanced tone is effectively amplified following the presentation of the precursor. Viemeister and Bacon (1982) hypothesized that this amplification occurs because the precursor, adapting neurons tuned to the background components, reduces the lateral inhibition

Correspondence to: Samuele Carcagno · Department of Psychology · Lancaster University · Lancaster, LA1 4YF, UK. email: s.carcagno@lancaster.ac.uk

that these neurons exert on neurons tuned to the target component (“adaptation of inhibition¹” hypothesis).

Several studies examined, at different stations of the auditory pathways, the responses of single neurons to target-plus-background mixtures that were preceded either by a precursor consisting of the background alone (notched precursor), or by silence. Palmer et al. (1995) failed to find evidence that the notched precursor affected the firing rate of guinea pig auditory nerve fibers tuned to the target tone. They found, however, that it caused a reduction in the firing rate of neurons tuned to the non-target frequencies. As a result, the notched precursor increased the neural representation of the target component *relative* to the neural representation of the non-target components. While such a relative increase may explain some aspects of enhancement, it cannot explain the increased forward masking effect of an enhanced tone, and also has difficulties explaining the reduction of the detectability threshold of an enhanced tone (Viemeister and Bacon 1982; McFadden and Wright 1990). Enhanced neural responses in the target frequency region following the presentation of the notched precursor were observed by Scutt and Palmer (1997, 1998) at the level of the cochlear nucleus in guinea pigs; these effects were generally small. More robust neural gains in the target frequency region following the presentation of the notched precursor were found by Nelson and Young (2010) at the level of the inferior colliculus (IC) in marmoset monkeys. Additionally, these authors found that during the presentation of the precursor, many fibers tuned to the target frequency region exhibited a buildup response consistent with a progressive release from lateral inhibition.

Little is known about the neurophysiological correlates of enhancement in humans. Zhang and Viemeister (2012) recorded cortical auditory evoked potentials (AEPs) in response to a target-plus-background mixture preceded by an exact copy of the mixture (no enhancement), or by a notched precursor causing enhancement. They found that the target-plus-background mixture elicited the P300 wave of the AEPs only when it was preceded by the notched precursor and listeners were actively attending to the stimuli. In a passive listening condition, some listeners showed a mismatch

¹ Viemeister and Bacon (1982) referred to the phenomenon as “adaptation of suppression,” suggesting that what was adapting was cochlear suppression. However, in the same paper, the authors also speculated that the adaptation could be occurring in the central nervous system. By using the phrase “adaptation of inhibition,” we refer here to this second hypothesis. Although cochlear suppression may “adapt” through the action of the medial olivo-cochlear efferent reflex (MOER) (Strickland 2004), its involvement in the enhancement effect remains unclear because of the lack of response enhancement at the level of the auditory nerve (Palmer et al. 1995), as well as the presence of enhancement effects in cochlear implant listeners (Wang et al. 2012).

negativity (MMN) to the target-plus-background mixture when it was preceded by the notched precursor, but this effect was not statistically significant at the group level. While transient AEPs, such as the P300 and the MMN, can answer some important questions (e.g., whether enhancement is affected by attention), their usefulness in elucidating the mechanisms of enhancement is limited by their inability to differentiate the response to the target tone from the response to the background tones. Steady-state evoked potentials, on the other hand, can be used to differentiate the response to tones with different frequencies presented simultaneously. For example, the auditory steady-state response (ASSR) can be used to estimate the magnitude of the neural response to each of a set of tones with different carrier frequencies. The technique works by “tagging” each tone with a signature amplitude modulation (AM) frequency. The magnitude of the neural response to each tone can then be estimated by measuring the power of the electroencephalographic (EEG) response at its signature AM frequency, by means of spectral analyses (Picton 2007). The ASSR magnitude correlates well with loudness (Ménard et al. 2008). Its neural generators depend on the modulation frequency, with modulation frequencies around 80 Hz reflecting mainly brainstem responses and modulation frequencies around 40 Hz reflecting mainly cortical responses (Herdman et al. 2002). Potentially, the ASSR thus represents an excellent technique to study the mechanisms and the locus of the auditory enhancement effect in humans.

In this study, we tested the hypothesis that enhancement originates from adaptation of inhibition. To this aim, we compared the ASSRs to a target tone presented with a simultaneous multitone background in the presence, and in the absence, of a precursor consisting of the multitone background alone. Given that the neurophysiological studies on non-human animals suggest that enhanced neural responses can be recorded at the level of the brainstem, we used signature AM frequencies close to 80 Hz to record the ASSR generated in brainstem nuclei. To assess whether ASSR changes caused by enhancement could account for the enhancement measured psychophysically, we compared them to changes in the ASSR caused by an acoustical increase in the level of the target tone; the magnitude of this acoustical increase was equivalent to the magnitude of enhancement estimated psychophysically.

METHODS

Participants

Fifty-nine listeners, who were mostly students in their 20s, with no self-reported hearing loss, took part in

the experiment. All listeners gave written informed consent and were paid an hourly wage for their participation in the experiment.

Stimuli

Enhancement was estimated psychophysically using a forward masking paradigm. The stimuli for the psychophysical sessions are illustrated in Figure 1. They consisted of a precursor (which could correspond with a period of silence), a masker, and a signal, and were presented in succession without silent delays. The signal was a 1.5-kHz pure tone. Its duration was 20 ms, including 5-ms raised-cosine onset and offset ramps. The masker was composed of 17 tones. The central component of the masker had the same frequency as the signal; the remaining 16 components formed two frequency sidebands placed symmetrically above and below the central component. The spacing between the central component and the closest component of each sideband was 500 cents (1 cent=1/1,200 octave). The spacing between consecutive components of the sidebands was 150 cents. All components of the masker were sinusoidally amplitude modulated with a 100 % modulation depth. The central component was modulated at a rate of 77.71 Hz, while the other components were all modulated at a rate of 73.14 Hz. The masker duration was 437.5 ms, including 10-ms onset and offset ramps.

There were three experimental conditions. In the *Prec* condition, the precursor was an exact copy of the masker, except for the central component which was not present (this precursor will be referred to as the

“notched precursor”). The notched precursor should elicit an enhancement of the central component of the masker, increasing as a consequence its forward masking effect. In the *Sil* condition, the precursor was a 437.5-ms silent interval. This condition served as a baseline to estimate the forward masking effect of the masker in the absence of enhancement. In the *Prec* and *Sil* conditions, each component of the masker was presented at 40 dB SPL. In the third experimental condition, the *Sil+10* condition, the masker component centered on the signal frequency was presented at 50 dB SPL, while the other masker components were presented at 40 dB SPL. The precursor consisted of a 437.5-ms silent interval, as in the *Sil* condition. A comparison of the signal thresholds in the *Sil* and *Sil+10* conditions allowed us to measure, for each individual listener, how much the forward masking threshold increased with a 10-dB increase in the level of the central masker component. This measure was used to estimate the slope (S) of the forward masking function using the following equation: $S=(T_{Sil+10}-T_{Sil})/10$, where T_{Sil+10} is the signal threshold in the *Sil+10* condition and T_{Sil} is the signal threshold in the *Sil* condition. The slope of the forward masking function allowed us to estimate, for each listener, the increase in level ΔE of the central masker component which was necessary to obtain, in the *Sil* condition, the same forward masking threshold as in the *Prec* condition. The following equation was used: $\Delta E=(T_{Prec}-T_{Sil})/S$, where T_{Prec} is the signal threshold in the *Prec* condition. According to the model of Viemeister and Bacon (1982), ΔE represents the amplification of the central masker component caused by the precursor in the *Prec* condition. We used

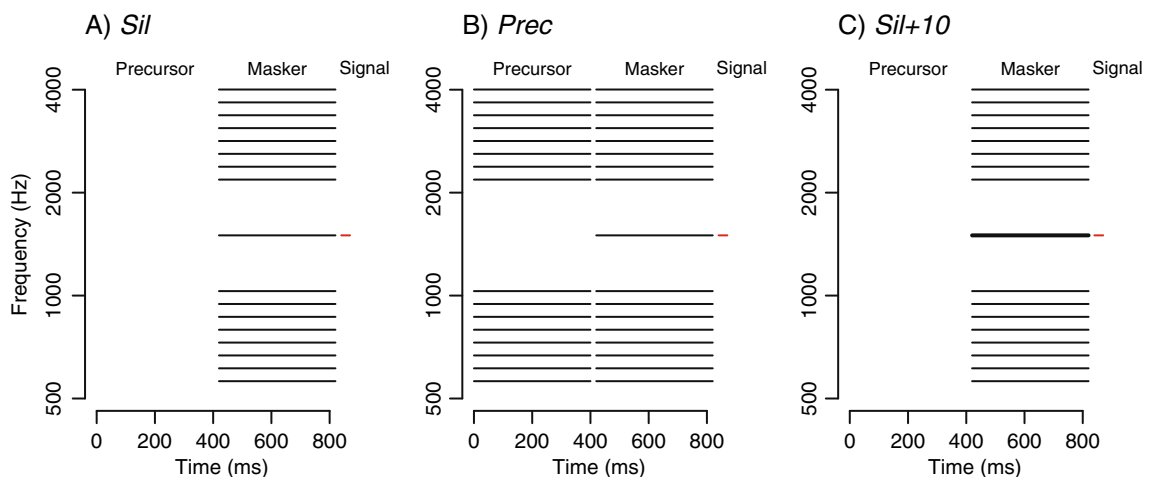


FIG. 1. Spectrographic representation of the stimuli used in the experiment. The signal (in red) was a 1.5-kHz tone. The masker consisted of 17 tones, and its central component was at the signal frequency. **A** In the *Sil* condition, the precursor corresponded with a 437.5-ms silent interval. **B** In the *Prec* condition, the precursor was

an exact copy of the masker, except for the omission of the central component. **C** In the *Sil+10* condition, the precursor corresponded with a 437.5-ms silent interval, and the central masker component (thick line) was 10 dB higher than in the other two conditions.

ΔE as a measure of enhancement. Byrne et al. (2011) have shown that the forward masking technique that we used in the current study produces similar estimates of enhancement magnitude as those obtained with a simultaneous masking technique in which the target component is used as a signal and its level is varied to find its audibility threshold.

In the EEG sessions, there were also three conditions: *Prec*, *Sil*, and *Sil ΔE* . The stimuli used in the *Prec* and *Sil* EEG conditions were identical to those employed in the corresponding psychophysical conditions, except that the signal was not presented. Given that the “masker” did not play a masking role in the EEG sessions, it will be referred to as the “test” sound. The *Sil ΔE* EEG condition was identical to the *Sil* EEG condition, except that the central component of the test sound was presented at $40 + \Delta E$ dB instead of 40 dB. In other words, the central component of the test sound was increased by the amount of enhancement estimated psychophysically for each listener. The purpose of the *Sil ΔE* condition was to test whether ASSR changes caused by enhancement in the *Prec* condition would be equivalent to the ASSR changes caused by an acoustical increase in the level of the central masker component.

Procedures

There are large differences in the magnitude of enhancement effects between different listeners, even after many hours of practice (McFadden and Wright 1990). In order to more readily observe potential enhancement effects with the ASSR, we selected those listeners showing the largest amounts of enhancement. During the selection phase, listeners performed several psychophysical sessions lasting about 1 h each. In each session, four thresholds per condition were obtained, and the magnitude of enhancement for each listener was estimated from these four thresholds. Listeners whose thresholds were relatively stable across two to three consecutive sessions and who showed at least 5 dB of enhancement were selected and went on to complete the remaining parts of the experiment. Fourteen out of the 59 listeners that were tested were selected. After the selection phase, the selected listeners performed two additional psychophysical sessions during which six threshold estimates for each experimental condition were obtained. The final estimate of enhancement magnitude, ΔE , for these listeners was computed from these six thresholds. Thus, the final enhancement magnitude estimate was independent of the estimates obtained during the selection phase. Pure-tone thresholds in quiet for the selected listeners were also measured at octave frequencies from 0.25 to 8 kHz to ascertain normal hearing. For each selected listener, ASSRs in

the three EEG conditions were then recorded in three separate sessions lasting about 2 h and 20 min each. For one listener, the EEG recordings showed abnormal levels of noise during the first two EEG sessions. This listener was not tested in the third EEG session, and her data were discarded from both the psychophysical and EEG analyses. Therefore, only the data of the 13 listeners who completed the entire experimental protocol (four males; mean age=21 years) will be presented.

Psychophysical Measurements

Signal thresholds were measured with a two-interval, two-alternative forced-choice task, using an adaptive procedure. On each trial, two observation intervals, separated by a 750-ms silent interval, were presented. Each observation interval contained the precursor (which could be silent), followed by the masker. In one observation interval, selected at random on each trial, the signal followed the masker, while in the other observation interval the signal was replaced by a 20-ms silence. Listeners had to indicate, by means of a button press on a computer keyboard, which observation interval contained the signal. At the end of each trial, feedback was provided by means of a colored light on a computer screen. Signal level varied according to a two-down, one-up adaptive rule tracking the 70.7 % correct point on the psychometric function (Levitt 1971). At the beginning of a block of trials, the signal level was set at least 10 dB above its forward masked threshold. Then, the signal level was increased (after an incorrect response) or decreased (after two consecutive correct responses) by 2 dB until the fourth reversal and by 1 dB thereafter. The adaptive procedure stopped at the 16th reversal. The mean signal level in the last 12 reversals was used to estimate the threshold for each block. In successive series of three blocks, the three experimental conditions were randomly ordered.

Listeners were seated in a double-walled, sound-insulated booth (Gisol, Bordeaux). The stimuli were generated digitally with a 32-bit resolution and a 48-kHz sampling rate in Python, on a PC housed outside the booth. The stimuli were sent to a 24-bit digital-to-analog converter (RME Hammerfall DSP Multiface) and played binaurally via Sennheiser HD650 headphones.

Electrophysiological Measurements

EEG responses were recorded using a Biosemi ActiveTwo system with a 2,048-Hz sampling rate. Electrodes were placed at Cz, Fz, and on the right mastoid (RM) following the 10-20 system. An additional electrode was placed on the neck just

below the hairline. During the recording, listeners reclined comfortably in the soundbooth and were asked to relax and refrain from extraneous body movements. The apparatus for generating and delivering the stimuli was the same apparatus used in the psychophysical sessions, except that the stimuli were played via mu-metal shielded Etymotic ER2 insert earphones. Triggers marking the start of a stimulus were sent to the Biosemi receiver from additional channels of the soundcard after being transformed to discrete pulses by a custom-built device. The stimuli were played binaurally in blocks of 50 trials. The inter-trial silent interval was 500 ms. In successive series of three blocks, the three experimental conditions were randomly ordered. During each EEG session, 1,200 responses for each experimental condition were collected. The final ASSRs were thus based on a total of 3,600 trials per condition.

The EEG responses were filtered offline with a zero-phase-shift, finite-impulse-response, high-pass digital filter with a 60-Hz cutoff. Four configurations of active/passive electrode were obtained by referencing the Cz and Fz electrodes either to the RM or to the Neck electrode: Cz-RM, Fz-RM, Cz-Neck, Fz-Neck. For each of these combinations, the recordings were segmented into discrete epochs and baseline corrected using a 200-ms window before the onset of the precursor. Epochs with electric potential values exceeding ± 20 μV were discarded. The average percentage of discarded trials across listeners was between 4 and 6 % for any of the active/passive electrode configurations. The recording segments corresponding to the presentation of the test sound (from 437.5 to 875 ms) were then extracted and concatenated into longer sweeps containing 20 segments each. Abrupt modulation phase shifts across the segments linked together should have been

minimal given that within a 437.5-ms segment there was an integer number of modulation cycles (32 cycles for the 73.14-Hz modulation frequency and 34 cycles for the 77.71-Hz modulation frequency) and the starting modulation phase was the same for all segments. These sweeps were then averaged. The spectral magnitudes of the EEG responses were obtained by applying a fast Fourier transform to the averaged sweeps for each condition. The spectral resolution was 0.114 Hz. The ASSRs to the central test sound component (target) and to the ensemble of the other components (background) were measured by taking the power of the EEG spectra at their respective modulation frequencies (target=77.71 Hz, background=73.14). The quality of the EEG signal, as assessed by the ASSR signal-to-noise ratio (SNR) for the target component across experimental conditions, was not equivalent across the four electrode configurations. A repeated-measure analysis of variance (ANOVA) on the ASSR SNR showed a significant effect of electrode configuration [$F(3,36)=6.59$, $p=0.001$]. For this reason, we will present only the data from the Fz-RM configuration, which showed the best ASSR SNR for the target component across experimental conditions.

RESULTS

Psychophysical Enhancement

Figure 2A displays the signal thresholds for the 13 listeners who showed the largest enhancement magnitudes during the selection phase and then completed the entire experimental protocol. The estimated enhancement magnitude ΔE for each of these listeners is displayed in Figure 2B. Although the selected listeners had shown 5 or more decibels of

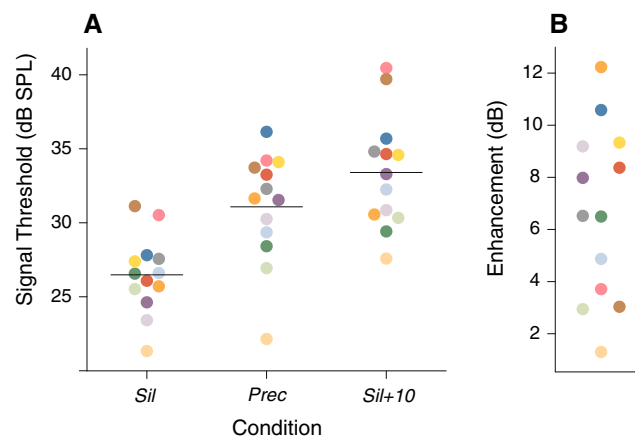


FIG. 2. Psychophysical results. **A** Signal thresholds in the *Sil*, *Prec*, and *Sil+10* conditions. Each colored point represents an individual listener. The color assigned to a given listener is the same across

panels. The *horizontal segments* represent the across-listener averages. **B** Enhancement magnitude, ΔE .

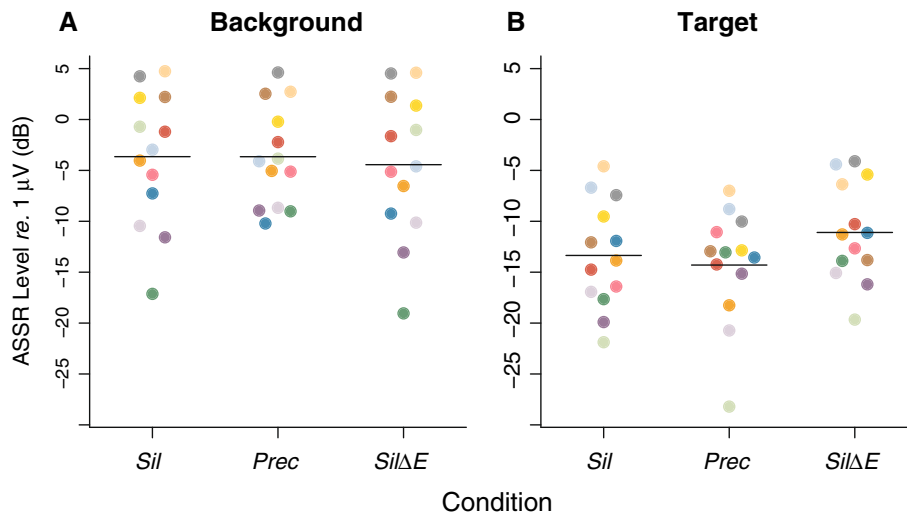


FIG. 3. **A** ASSR levels for the background components as a function of precursor type. **B** ASSR levels for the target component as a function of precursor type. Each *colored point* represents an

individual listener. The color code is the same as in Figure 2. The *horizontal segments* represent the across-listener averages.

enhancement during the last sessions of the selection phase, ΔE was less than 5 dB for some of them. The average ΔE value was 6.66 dB.

ASSRs

Figure 3 shows, for each experimental condition, the level of the ASSR to the test sound at the modulation frequencies of the target and background components. The responses at the background modulation frequency were larger than the responses at the target modulation frequency. This likely reflects the fact that the background included 16 tones while the target consisted of a single tone. The responses to the target and to the background components appear to differ

little between the *Sil* and *Prec* conditions. This can be seen better in Figure 4, which shows the difference between the ASSRs in the *Prec* and *Sil* conditions, a difference that was not statistically significant either for the target [$t(12)=-0.917$, $p=0.377$] or for the background [$t(12)=-0.009$, $p=0.993$] components. Additionally, as can be seen in Figure 5, there was no significant correlation between ASSR target enhancement (difference between *Prec* and *Sil* conditions) and psychophysical enhancement (ΔE) [$\rho=-0.1$, $p=0.746$]. These results suggest that the precursor had little or no effect on the ASSRs to the test sound. In contrast, acoustically increasing the level of the target component had a significant effect on the responses to both the target and the background

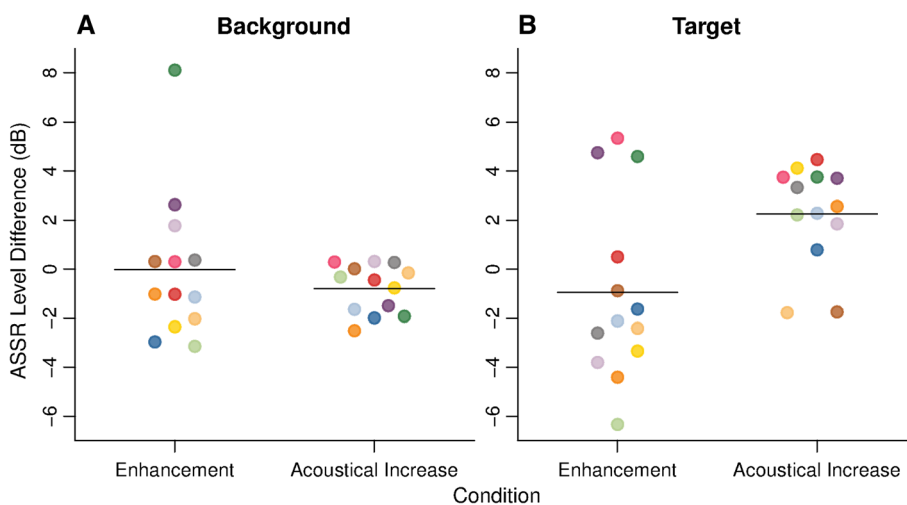


FIG. 4. **A** ASSR level difference between the *Prec* and *Sil* conditions (Enhancement), and between the *Sil* ΔE and *Sil* conditions (Acoustical Increase) for the background components. **B** Same as **A**, but for the target component. Each *colored point*

represents an individual listener. The color code is the same as in Figures 2 and 3. The *horizontal segments* represent the across-listener averages.

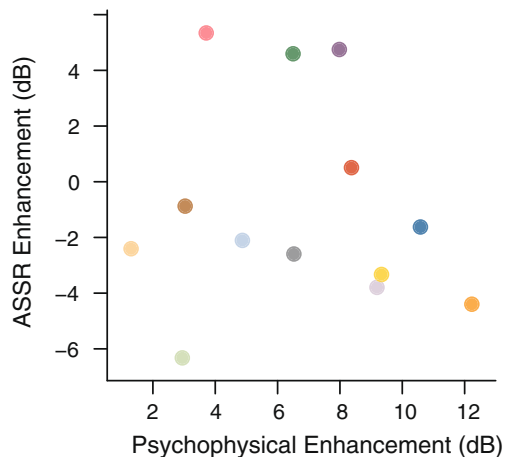


FIG. 5. Scatterplot showing the relationship between psychophysical enhancement (ΔE) and ASSR enhancement (difference of the ASSRs to the target tone in the *Prec* and *Sil* conditions). Each colored point represents an individual listener. The color code is the same as in Figures 2, 3, and 4.

components: The responses to the target were significantly larger in the *Sil* ΔE condition than in the *Sil* condition [$t(12)=3.959$, $p=0.002$], while the responses to the background components were significantly smaller in the *Sil* ΔE condition than in the *Sil* condition [$t(12)=-2.872$, $p=0.014$]. The reduction of the response to the background components when the acoustic level of the target was higher is likely to be a consequence of cochlear two-tone suppression (Shannon 1976). A direct comparison of the *Prec* and *Sil* ΔE conditions reveals that the ASSRs to the target were significantly larger in the *Sil* ΔE than in the *Prec* condition [$t(12)=3.194$, $p=0.008$]. The ASSRs to the background, on the other hand, did not differ significantly between the *Sil* ΔE and *Prec* conditions [$t(12)=-0.858$, $p=0.407$].

Because the enhancement effect is larger immediately after the end of the precursor, we also analyzed only the first half of the response to the test sound (the first 218.75 ms of the response). In each experimental condition, the magnitude of the ASSRs for the first half of the test sound was very similar to that obtained for the full duration of the test sound. Again, for the target component, the difference between the ASSRs in the *Prec* and *Sil* conditions was not statistically significant [$t(12)=-1.449$, $p=0.173$]. On the other hand, the difference between the *Sil* ΔE and *Sil* conditions [$t=2.668$, $p=0.02$], as well as the difference between the *Sil* ΔE and *Prec* conditions [$t=3.9$, $p=0.002$], were statistically significant.

DISCUSSION

We investigated the auditory enhancement effect psychophysically, by means of a forward masking

technique, and electrophysiologically, using the 80-Hz ASSR. While we obtained considerable enhancement effects psychophysically, the ASSR to an enhanced tone was not significantly greater than the ASSR to a tone that was not enhanced. Acoustically increasing the level of the target tone by the magnitude of enhancement estimated psychophysically elicited stronger ASSRs than enhancing the tone by means of a notched precursor. Before discussing the possible implications of these results with respect to the neural locus of enhancement, we will discuss two other incidental findings: the apparent insensitivity of the ASSR to adaptation and the interindividual differences in the magnitude of enhancement effects.

ASSR Adaptation

The ASSR to the background components did not appear to be affected by the precursor: Remarkably, it did not show any signs of reduction, which could be expected as a result of neural adaptation. We are not aware of any studies on the adaptation of the 80-Hz ASSR. However, our results are consistent with other experiments showing a lack of adaptation of the 40-Hz ASSR (Forss et al. 1993; Ross et al. 2002; Pantev et al. 2004; Okamoto et al. 2004; Kuriki et al. 2013). The reasons for the lack of adaptation effects on the 40-Hz ASSR are not well understood. The lack of adaptation of the ASSR, however, does not in itself rule out the ASSR as a neural metric of enhancement. Models of enhancement based exclusively on adaptation of the background components cannot explain the increased forward masking effect of an enhanced tone. Moreover, these models have difficulties explaining the reduction of the detectability threshold of an enhanced tone (Viemeister and Bacon 1982; McFadden and Wright 1990). A reduction of the response outside the critical band of the signal, on its own, would not change the signal-to-noise ratio within the critical band of the signal. A reduction of across-channel inhibitory effects is necessary to explain the increased detectability of an enhanced tone. Therefore, adaptation of the background components is, on its own, a poor neural correlate of the enhancement effect. A better neural correlate of the enhancement effect is an increased neural response to the enhanced target tone. The insensitivity of the ASSR to adaptation has no bearing on its ability to detect an increased neural response to the target tone.

Interindividual Differences in Enhancement Effects

Only 14 of the 59 listeners tested met the criterion of having at least 5 dB of enhancement to proceed to the main phase of the experiment. For the remaining 45

participants, the average enhancement measured in their last session was 2.3 dB. This estimate is likely to be biased downward because we stopped testing these listeners as soon as the estimated magnitude of their enhancement effect fell below our criterion threshold. The average enhancement estimate across all listeners (both the selected and the discarded listeners) was 3.3 dB. This estimate is about 2 dB lower than the estimate obtained by Byrne et al. (2011) using a similar methodology, but different stimuli. The lower enhancement estimate obtained in our study is probably due to the fact that the duration of the masker that we used was longer (437.5 ms) than the duration of the masker used by Byrne et al. (250 ms). In the forward masking paradigm, enhancement decreases as the masker duration increases (Wright and McFadden 1992). The longer masker duration used in our study was necessary to obtain ASSRs to the test sound of sufficient overall time length, while limiting the duration of the recording sessions.

The interindividual differences that we observed in the magnitude of the enhancement effect were large, although not as large as those observed previously by McFadden and Wright (1990) in a simultaneous masking task. In their study, listeners showing the greatest enhancement effects also had the lowest sensitivity. In our study, the 14 selected participants had, on average, lower thresholds (thus a higher sensitivity) in the *Sil* and *Sil+10* conditions (by 1.7 and 2.2 dB, respectively) than the 45 discarded participants. The average threshold in the *Prec* condition was 1.1 dB higher in the 14 selected participants than in the 45 discarded participants. Overall, these threshold differences were small and none of them was statistically significant. These data indicate that, across conditions, the selected participants did not perform worse than the discarded participants. The generalizability of our results to the overall population rests on the assumption that the neurophysiological mechanism underlying enhancement in the selected participants is the same as in the discarded participants. Overall, the data above suggest that this assumption is correct, and that the difference between the selected participants and the discarded participants was quantitative (enhancement magnitude) rather than qualitative (enhancement mechanism). Further studies will be necessary to better understand the origin of the large interindividual differences in enhancement effects.

Neural Locus of Enhancement

Our electrophysiological recordings show that the 80-Hz ASSR was able to index differences in the acoustical level of the target tone equivalent to the magnitude of enhancement. However, we found no evidence that

enhancement of the target tone affected the amplitude of the 80-Hz ASSR. Overall, these results suggest that the enhancement effects that we measured psychophysically cannot be fully explained by an amplified neural response at the level of the brainstem to the enhanced tone (Viemeister and Bacon 1982). Our results differ from those of Nelson and Young (2010): using similar stimuli, they found that a population of neurons of the IC in marmoset monkeys had greater firing rates in response to enhanced tones than in response to unenhanced tones. While it is possible that the neural gains found by Nelson and Young were too small to be detected by the ASSR, we were able to detect ASSR gains caused by an acoustical increase that was psychophysically equivalent to the enhancement effect. Nelson and Young did not measure enhancement psychophysically; thus, it is not clear whether the neural gains they measured would be sufficient to explain psychophysical enhancement. This leaves open the possibility that the main contribution to the enhancement effects measured psychophysically comes from further neural gains at the level of the auditory cortex. This hypothesis could be tested in humans by employing the 40-Hz ASSR, which reflects the contribution of both cortical and subcortical generators (Herdman et al. 2002).

An alternative interpretation of our results is that the putative neural populations showing enhanced responses in the brainstem did not contribute, or contributed only little, to the 80-Hz ASSR. Source localization studies on human subjects indicate that the main generators of the 80-Hz ASSR are located in the brainstem (Herdman et al. 2002), but the precise brainstem structures involved are not known. In non-human animals, the results of a selective-lesion study (Kiren et al. 1994) and the results of a study recording local field potentials (Kuwada et al. 2002) are consistent with the hypothesis that the IC is the dominant generator of the scalp-recorded 80-Hz ASSR. Thus, it seems likely that the ASSRs that we recorded had a dominant source in the IC, the same structure studied by Nelson and Young (2010). However, we cannot exclude that other neural populations outside of the IC contributed significantly to our ASSRs. In the same vein, we cannot exclude the possibility that neural units sensitive to enhancement do not exhibit phase-locked responses to AM stimuli (Shadduck Palombi et al. 2001), and hence do not contribute to the ASSR. Finally, for some of the IC neurons studied by Nelson and Young (2010), the notched precursor caused a *suppression*, rather than an enhancement, of their firing rate to the target tone. These suppressive responses may thus have reduced the enhanced neural response at the level of the global neural population indexed by the ASSR. However, in the study of Nelson and Young (2010), the proportion of IC neurons showing suppression

was much smaller than the proportion of neurons showing enhancement. Therefore, it is unlikely that these suppressive responses could account for the lack of enhancement effect on the ASSR.

Another possible interpretation of our findings is that enhancement is not mediated by an increase in the neural representation of the target tone. Alternative hypotheses regarding the origins of the phenomenon exist (see for example Viemeister and Bacon 1982; Summerfield et al. 1987; Carcagno et al. 2013 for discussion of alternative hypotheses), and there is evidence that multiple mechanisms at different levels of the auditory system can contribute to enhancement (Carcagno et al. 2012; Byrne et al. 2013). However, none of these alternative hypotheses predicts an absolute gain of the internal representation of an enhanced tone, and this gain is thought to be necessary to account for the increased forward masking effect of an enhanced tone. An alternative account of the increased forward masking effect of an enhanced tone is that the salient spectral change in the transition from the precursor to the masker “distracts” the listener from the subtle change associated with the signal, and therefore raises the detection threshold of the latter (see commentary by Moore in Wright and McFadden 1992). This potential distraction effect is likely to be larger the closer the precursor-masker spectral transition is to the signal. Given that in our experiment the signal started 400 ms after this spectral transition, it is unlikely that such a distraction effect could account for our forward masking results.

CONCLUSIONS

Our study suggests that enhancement does not affect the amplitude of the brainstem 80-Hz ASSR and shows that enhancement-induced changes of the 80-Hz ASSR (if any) cannot account for the magnitude of the enhancement effect measured psychophysically. These results could be seen as evidence that enhancement is not occurring at the level of the brainstem. However, they are also consistent with the possibility that enhancement does occur in the brainstem but the putative brainstem neural populations responsible for enhancement do not contribute to the 80-Hz ASSR.

ACKNOWLEDGMENTS

We would like to thank Thierry Lafon for technical help in the setup of the EEG apparatus. We also thank the Associate Editor and two anonymous reviewers for constructive comments on an earlier version of the manuscript. This work was supported by grants from the Agence Nationale de la Recherche (LEAP, AMUSIA).

Conflict of Interest The authors declare that they have no conflict of interest.

REFERENCES

- 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
- BYRNE AJ, STELLMACK MA, VIEMEISTER NF (2013) The salience of enhanced components within inharmonic complexes. *J Acoust Soc Am* 134:2631–2634
- CARCAGNO S, SEMAL C, DEMANY L (2012) Auditory enhancement of increments in spectral amplitude stems from more than one source. *J Assoc Res Otolaryngol* 13:693–702
- CARCAGNO S, SEMAL C, DEMANY L (2013) No need for templates in the auditory enhancement effect. *PLoS One* 8:e67874
- DEMANY L, CARCAGNO S, SEMAL C (2013) The perceptual enhancement of tones by frequency shifts. *Hear Res* 298:10–16
- ERVITI M, SEMAL C, DEMANY L (2011) Enhancing a tone by shifting its frequency or intensity. *J Acoust Soc Am* 129:3837–3845
- FORSS N, MÄKELÄ JP, McEVOY L, HARI R (1993) Temporal integration and oscillatory responses of the human auditory cortex revealed by evoked magnetic fields to click trains. *Hear Res* 68:89–96
- HERDMAN AT, LINS OG, VAN ROON P, STAPPELLS DR, SCHERG M, PICTON TW (2002) Intracerebral sources of human auditory steady-state responses. *Brain Topogr* 15:69–86
- KIREN T, AOYAGI M, FURUSE H, KOIKE Y (1994) An experimental study on the generator of amplitude-modulation following response. *Acta Otolaryngol Suppl* 511:28–33
- KURIKI S, KOBAYASHI Y, KOBAYASHI T, TANAKA K, UCHIKAWA Y (2013) Steady-state MEG responses elicited by a sequence of amplitude-modulated short tones of different carrier frequencies. *Hear Res* 296:25–35
- KUWADA S, ANDERSON JS, BATRA R, FITZPATRICK DC, TEISSIER N, D'ANGELO WR (2002) Sources of the scalp-recorded amplitude-modulation following response. *J Am Acad Audiol* 13:188–204
- LEVITT H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477
- McFADDEN D, WRIGHT BA (1990) Temporal decline of masking and comodulation detection differences. *J Acoust Soc Am* 88:711–724
- MÉNARD M, GALLÉGO S, BERGER-VACHON C, COLLET L, THAI-VAN H (2008) Relationship between loudness growth function and auditory steady-state response in normal-hearing subjects. *Hear Res* 235:105–113
- NELSON PC, YOUNG ED (2010) Neural correlates of context-dependent perceptual enhancement in the inferior colliculus. *J Neurosci* 30:6577–6587
- OKAMOTO H, ROSS B, KAKIGI R, KUBO T, PANTEV C (2004) N1m recovery from decline after exposure to noise with strong spectral contrasts. *Hear Res* 196:77–86
- PALMER AR, SUMMERFIELD Q, FANTINI DA (1995) Responses of auditory-nerve fibers to stimuli producing psychophysical enhancement. *J Acoust Soc Am* 97:1786–1799
- PANTEV C, OKAMOTO H, ROSS B, STOLL W, CIURLIA-GUY E, KAKIGI R, KUBO T (2004) Lateral inhibition and habituation of the human auditory cortex. *Eur J Neurosci* 19:2337–2344
- PICTON TW (2007) Audiometry using auditory steady-state responses. In: Burkard RF, Eggermont JJ, Don M (eds) *Auditory evoked potentials: basic principles and clinical applications*. Lippincott Williams & Wilkins, Baltimore, pp 441–462

- ROSS B, PICTON TW, PANTEV C (2002) Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res* 165:68–84
- SCUTT MJ, PALMER AR (1997) Psychophysical and physiological responses to signals which are enhanced by temporal context. *Assoc Res Otolaryngol Abs* 188 (A)
- SCUTT MJ, PALMER AR (1998) Physiological enhancement in cochlear nucleus using single tone precursors. *Assoc Res Otolaryngol Abs* 381 (A)
- SHADDUCK PALOMBI P, BACKOFF PM, CASPARY DM (2001) Responses of young and aged rat inferior colliculus neurons to sinusoidally amplitude modulated stimuli. *Hear Res* 153:174–180
- SHANNON RV (1976) Two-tone unmasking and suppression in a forward-masking situation. *J Acoust Soc Am* 59:1460–1470
- STRICKLAND EA (2004) The temporal effect with notched-noise maskers: analysis in terms of input-output functions. *J Acoust Soc Am* 115:2234–2245
- SUMMERFIELD Q, SIDWELL A, NELSON T (1987) Auditory enhancement of changes in spectral amplitude. *J Acoust Soc Am* 81:700–708
- VIEMEISTER NF (1980) Adaptation of masking. In: van den Brink G, Bilsen FA (eds) *Psychophysical, physiological and behavioural studies in hearing*. Delft University, Delft, The Netherlands, pp 190–199
- VIEMEISTER NF, BACON SP (1982) Forward masking by enhanced components in harmonic complexes. *J Acoust Soc Am* 71:1502–1507
- WANG N, KREFT H, OXENHAM AJ (2012) Vowel enhancement effects in cochlear-implant users. *J Acoust Soc Am* 131:EL421–EL426
- WRIGHT BA, MCFADDEN D (1992) Evidence that adaptation of suppression cannot account for auditory enhancement or enhanced forward masking. *Philos Trans R Soc Lond B Biol Sci* 336:325–328, discussion 328–329
- ZHANG Y, VIEMEISTER NF (2012) Cortical processing of the enhancement effect in passive and active listening conditions. *Assoc Res Otolaryngol Abs*: 39–40