Dynamic representation of spectral edges in guinea pig primary auditory cortex

Noelia Montejo¹, Arnaud J. Noreña¹*

Affiliation:

¹ Aix Marseille Université, CNRS, NIA UMR 7260, 13331, Marseille, France

*Corresponding author:

Aix-Marseille Université, CNRS, NIA UMR 7260, Centre Saint Charles, Case B

3 place Victor Hugo, 13331 Marseille, France

arnaud.norena@univ-amu.fr
Abstract

The central representation of a given acoustic motif is thought to be strongly context-dependent, i.e. to rely on the spectro-temporal past and present of the acoustic mixture in which it is embedded. The present study investigated the cortical representation of spectral edges (i.e. where stimulus energy changes abruptly over frequency) and its dependence on stimulus duration and depth of the spectral contrast in guinea pig. We devised a stimulus ensemble composed of random tone pips with or without an attenuated frequency band (AFB) of variable depth. Additionally, the multi-tone ensemble with AFB was interleaved with periods of silence or with multi-tone ensembles without AFB. We show that the representation of the frequencies near but outside the AFB is greatly enhanced whereas the representation of frequencies near and inside the AFB is strongly suppressed. These cortical changes depend on the depth of the AFB: although they are maximal for the largest depth of the AFB, they are also statistically significant for depths as small as 10 dB. Finally, the cortical changes are quick, as they occur within a few seconds of stimulus ensemble presentation with AFB, and are very labile as they disappear within a few seconds after the presentation without AFB. Overall, this study demonstrates that the representation of spectral edges is dynamically enhanced in the auditory centers. These central changes may have important functional implications, particularly in noisy environments where it could contribute to preserving the central representation of spectral edges.

Keywords: synaptic inhibition; short-term synaptic plasticity; artificial hearing loss; tinnitus
Introduction

An important question in auditory neuroscience addresses the understanding of how acoustic stimuli are represented in the central auditory system (Nelken, 2004; Escabí and Read, 2005; Sutter, 2005; Winer et al., 2005; Oswald et al., 2006; Reyes, 2011). This process is particularly challenging as sounds in the natural environment (such as conspecific vocalizations) have complex spectra that change rapidly with time, and are often embedded in background noise (Singh and Theunissen, 2003; Palmer and Shamma, 2004; Shamma and Micheyl, 2010; Theunissen and Elie, 2014). Although we are far from a complete understanding of how complex sounds are represented in the central auditory system, it appears that the representation of many sound features is context-dependent, i.e., it depends on the spectro-temporal past and present of the acoustic mixture (Brosch et al., 1999; Brosch and Schreiner, 2000; Blake and Merzenich, 2002; Brosch and Scheich, 2008; Noreña et al., 2008; Gourévitch et al., 2009; Catz and Noreña, 2013).

The sensitivity of central auditory neurons to the spectro-temporal history of the acoustic environment has traditionally been examined using simple stimulus paradigms such as two-tone sequences (Shamma and Symmes, 1985; Rhode and Greenberg, 1994; Brosch and Schreiner, 1997, 2000; Brosch et al., 1999; Kadia and Wang, 2003; Brosch and Scheich, 2008). Namely, a “target” tone presented at a neuron’s best frequency (BF) is preceded by a “modulating” tone at variable frequency, sound pressure level (SPL), and interstimulus interval (typically between 0 and a few hundred msec). Such studies showed that neural responses to the target tone could be strongly suppressed, or in some cases enhanced, by the prior presentation of the modulating tone; that is, they could be strongly context-dependent. Other studies derived neural properties, such as spectro-temporal receptive fields (STRFs), from a variety of more complex stimuli, such as white noise, spectro-temporally modulated (“rippled”) noise, ensembles of random tone bursts, and conspecific vocalizations (Eggermont et al., 1983; deCharms et al., 1998; Depireux et al., 2001; Theunissen et al., 2001; Blake and Merzenich, 2002; Valentine and Eggermont, 2004; Escabí and Read, 2005; Noreña et al., 2006, 2008; Gourévitch et al., 2009; Rabinowitz et al.,...
2011, 2012; Catz and Noreña, 2013; Theunissen and Elie, 2014). It was shown that while the shapes of STRFs are relatively insensitive to some stimulus characteristics (i.e., SPL; Valentine and Eggermont 2004; Pienkowski and Eggermont 2011), they are sensitive to others (i.e. stimulus spectro-temporal density; Blake and Merzenich, 2002; Noreña et al., 2008).

Animal vocalizations, including human speech, contain sharp spectral peaks and transitions (i.e., vowel formants) which differentiate them from each other and from the noise background (Moore and Glasberg, 1983; Summerfield and Assmann, 1989; Baer et al., 1993; Assmann and Summerfield, 2004; Palmer and Shamma, 2004). We recently investigated the auditory cortical representation of spectral edges or contours, and how this representation varies with their physical properties (i.e. width, depth and slope) (Catz and Noreña, 2013). In order to investigate the respective representation of frequencies near and remote to the spectral edges, we devised a stimulus ensemble composed of random tone bursts (between 500 Hz and 32 kHz, 1/8 octave step) with or without an attenuated frequency band (AFB), within which tone levels are reduced to create edge contrasts. It was shown that the representation of tone frequencies adjacent to the AFB is enhanced, whereas the representation of frequencies in the AFB, especially near the spectral edges, is reduced. These cortical changes are very sensitive to the characteristics of the spectral notch. These results are broadly consistent with the general idea that the representation of acoustic features is context-dependent, and suggest that some mechanisms act to enhance the central auditory representation of spectral edges (Von Békésy, 1967, 1969; Gourévitch et al., 2009; Shamma et al., 2011).

The present study is aimed at extending the results of our previous work. First, we further investigated the effects of spectral notch depth to determine whether edge enhancement could be seen for spectral contrasts as small as 10 dB. Second, the edge enhancements described in our previous study were observed after a 180-sec presentation of the AFB stimulus. This is relatively fast but still too slow to dynamically shape the representation of real acoustic environments which can vary on much smaller (fractions of a
second) time scales (Singh and Theunissen, 2003). We thus investigated whether edge enhancement also occurs during a much shorter (2-10 sec) presentation of the AFB stimulus. Third, it was unknown whether edge enhancement persisted for some time after the presentation of the AFB stimulus. This question was addressed by interleaving 10-sec AFB stimulus presentations with 10-sec periods of the control multi-tone stimulus (without AFB). Fourth, the edge enhancement reported in our previous study was observed in anesthetized animals. In order to assess the effects of anesthesia on the neural changes at spectral edges, a subset of experiments were carried out in a few awake animals.
Methods

Animal preparation

The care and use of animals used in this study were approved by the Animal Care Committee of Bouches du Rhones, France (# A 13-504). A total of 18 guinea pigs weighing between 400 and 800 g were used for this study. Animals were deeply anesthetized with the administration of 50 mg/kg of ketamine hydrochloride (Imalgene 1000) and 3 mg/kg of xylazine (Rompun 2%), injected intramuscularly; 0.1 ml of Atropine methyl nitrate were also administered to reduce mucus secretion in the throat. Throughout the experiment, anesthesia was maintained with half the dose of ketamine and xylazine administrated every hour. The tissue overlying the frontal lobe was opened, four small screws were fixed to the top of the skull and dental cement was used to fix a large screw. This screw provided a mechanically stable point of attachment to a metal head post. The tissue, skull and dura overlying the primary auditory cortex were removed over a surface of roughly one cm² (see (Grimsley et al., 2012), for precise location). The body temperature was maintained at around 37°C with a thermostatically controlled heating blanket. After the experiment, a lethal dose of sodium pentobarbital was administered.

Three additional animals were implanted chronically with micro-electrodes to investigate whether cortical changes reported in anesthetized animals were also present in the awake preparation. In these animals, a cylindrical plastic chamber (9 mm diameter and 7 mm height) was fixed on the skull with dental cement. The center of the chamber was positioned roughly on the center of the craniotomy. The micro-electrodes were then advanced perpendicularly to the cortex through the chamber and inserted into the primary auditory cortex. The micro-electrodes were fixed to the chamber with a silicone sealant. The animals were allowed to recover for one week before cortical recordings. Animals’ head was fixed with the same method described above for anesthetized preparation.

Acoustic stimulation
Stimuli were generated in MATLAB and transferred to an RP2.1-based sound delivery system (Tucker Davis Technologies). Acoustic stimuli were presented in a sound booth room from a headphone (Sennheiser HD595) placed 10 cm in front of the ear contralateral to the cortex where the recordings were carried out. The amplitude of each tone pip was adjusted to the transfer function of the sound delivery system so that they were presented at the desired level in dB SPL.

Spectro-Temporal Receptive Fields (STRFs) were obtained from a 180-second multi-tone pip stimuli (deCharms et al., 1998; Blake and Merzenich, 2002; Valentine and Eggermont, 2004; Noreña et al., 2008; Gourévitch et al., 2009; Catz and Noreña, 2013). Tone pips (49 frequencies, 8 frequencies per octave covering 6 octaves between 500 Hz and 32 kHz) were presented randomly over time (independent Poisson process for each frequency with a rate of 2 Hz and a 50-msec dead time designed to prevent tones of the same frequency from overlapping in time). However, tone pips of different frequencies could occasionally overlap in time. The envelope of the tone pips is given by $\gamma(t) = (t/4)^2e^{-t/4}$ with $t$ in milliseconds (the stimulus duration is 50 msec and the maximum amplitude is reached at 8 msec). The average rate of tone pip presentation was around 16 Hz/octave (considering the number of tone frequencies present per octave, along with the average presentation rate of each). This pip rate per octave of the stimulus was chosen as a trade off between spectro-temporal density and neural adaptation (so as to preserve cortical activity) (Blake and Merzenich, 2002; Noreña et al., 2008). Control STRFs were obtained from multi-tone stimuli with tone pips presented at 70 dB SPL (ctrl-70), 60 dB SPL (ctrl-60), 50 dB SPL (ctrl-50), 40 dB SPL (ctrl-40) (Figure 1A).

A first experiment was carried out to study the effects of spectral contrast depth on the representation of acoustic spectral edges. Multi-tone pip stimuli (as described above) with an attenuated frequency band (AFB) were used: all pure tones were presented at 70 dB, except those corresponding to the frequency band of the AFB where pure tones were omitted (AFB 70-0), or presented at 40 dB SPL (AFB 70-40), 50 dB SPL (AFB 70-50), or 60 dB SPL (AFB 70-60) (Figure 1A). The frequencies within half-an octave above or below the AFB were
called the edge-out frequencies. On the other hand, the frequencies within half-an octave
above or below the lower or upper edge, respectively, were called the edge-in frequencies. In
this experiment the width of the AFB was fixed to 1 octave. The center frequency of the AFB
was set as follows: the BF for each cortical site was derived from the control stimulus (ctrl-70). The center frequency of the AFB was then set to the BF of a given cortical site. Cortical
responses were obtained for all stimulus conditions (different depths) for that specific center
frequency of the AFB. Once a set of recordings was completed, another set of recordings
was carried out with a different AFB stimulus (centered on the BF of another cortical site).
One notes that, as we recorded from many cortical sites simultaneously, the BFs could
correspond to the center frequency of the AFB, an edge frequency of the AFB, or a remote
frequency from the AFB.

In a second experiment, we investigated whether the cortical changes produced by
multi-tone pip stimuli presented for 180 sec could also be induced by a shorter presentation
(10 sec). Three stimuli were constructed to address this question. A control stimulus
consisted of an alternation of 10-sec presentation of ctrl-70 and 10-sec silence (ctrl-sil)
(Figure 1B). One test stimulus consisted of an alternation between 10-sec AFB 70-0 stimulus
and 10-sec silence (AFB-sil). This tests stimulus was specifically designed to address the
cortical changes during the 10-sec presentation of the stimulus with AFB. The other test
stimulus consisted of an alternation of 10-sec AFB 70-0 stimulus and ctrl-70 (ctrl-AFB). This
test stimulus was specifically designed to address whether the cortical changes produced by
the stimulus with AFB produces some persistent effects, i.e. during the following presentation
of the ctrl-70 stimulus. In this experiment, two widths of the AFB (½ or 1 octave) were
investigated.

MUA and LFP recording procedure

Each set of recordings was obtained with 1 array of 16 electrodes (Alpha-Omega Eng,
Nazareth, Israel) arranged in an 8 by 2 pattern with 0.25 mm electrode separation within the
long row and 0.5 mm separation between columns. The electrodes had impedances between
0.8-1.4 MOhm. The array was manually advanced using a Narishige microdrive into the primary auditory cortex (Wallace et al., 2000). The signals were then amplified 10,000 times with filter cutoff frequencies set at 2 Hz and 5 kHz. The amplified signals were processed by a TDT-System 3 multichannel data acquisition system. Multi-unit activity was sampled at 24414 Hz and was extracted from the 300-Hz high-pass filtered signal. Local field potentials (LFPs) were sampled at 1061 Hz and were extracted from the 300-Hz low-pass filtered signal. In this way, we were able to record spikes and LFPs simultaneously.

At an initial stage of the experiments, a “search procedure” was used and consisted of recording cortical activity induced by clicks, noise bursts and tone pips (from 500 Hz to 32 kHz, 1/8 octave step). This procedure provided a rough estimate of the tonotopy and the amplitude of LFPs. Electrodes were placed at a depth where the (negative) amplitude of stimulus-induced LFPs was near maximal (region of the border between layer III and IV - (Szymanski et al., 2011).

**Data analysis**

All results were computed using custom MATLAB routines. Multi-unit activity or “spike events” were detected by using an amplitude threshold on the high-pass filtered data. The median was calculated on the negative values of the filtered signal; the threshold was then set to six times the median (Quiroga et al., 2004).

The methodology for computing STRFs was similar to that used in previous studies (Valentine and Eggermont, 2004; Noreña et al., 2008; Catz and Noreña, 2013). Briefly, STRFs for MUA were determined by constructing poststimulus time histograms (PSTHs), with time bins of 1 msec for each tone pip frequency. In other words, spikes falling in the averaging time window (starting at the stimulus onset and lasting 100 msec) are counted. Because the average interstimulus interval in the stimulus ensemble (~10 msec) is smaller than the averaging time window, a spike can be counted in the PSTH of several pip frequencies. STRFs for LFPs were obtained by a similar procedure, except that the LFP waveforms (0–100 msec after stimulus onset) were averaged for each tone pip frequency.
The build-up and break down of the neural changes produced by AFB stimulus was investigated by alternating 10-sec AFB 70-0 stimulus with control stimulus or silence. STRFs were obtained by averaging PSTHs or LFP waveforms over the time periods of interest (either ctrl or AFB). Moreover, the 10-sec stimulation periods (AFB or ctrl) were divided into three time periods, i.e. [0-2.5 sec], [2.5-5 sec] and [5-10 sec] and STRFs were derived from each of these time windows. This more detailed analysis of the neural changes produced by AFB stimulus was made possible by the fact that 9 samples of each multi-tone ensemble were available to derive the STRFs.

The maximal MUA response (or the minimal LFP amplitude) within the 10-30 msec time window after stimulus onset and over all frequencies was obtained from the ctrl-70 STRF. All STRFs (including those obtained from the ctrl-70 condition) were then normalized by dividing the mean neural activity by this single value. This normalization was aimed at minimizing the firing rate variability across recording sites. By definition, the maximum neural activity for the ctrl-70 condition was 1 (at the best frequency). One notes that values above 1 are frequently observed in the AFB conditions (i.e. at the edge frequencies of the AFB); this indicates that the maximum of absolute firing rate in the AFB conditions is larger than the maximum of absolute firing rate in the ctrl-70 condition. This normalized mean neural activity is the dependent variable displayed in the STRFs.

In order to compare the STRFs obtained from control and AFB stimuli (and for display purpose), the differences between their frequency profiles were computed (Catz and Noreña, 2013). The frequency profiles were obtained from the normalized STRFs by taking the maximum neural activity within a time window of 10-30 msec after stimulus onset for each tone pip frequency. For the frequencies outside the AFB, which were presented at 70 dB, the responses were compared to the corresponding frequencies obtained from the ctrl-70. For the frequencies inside the AFB, which were presented at 40, 50 or 60 dB, the responses were compared to the corresponding frequencies obtained from the ctrl-40, ctrl-50 or ctrl-60, respectively.
The main purpose of this study was to investigate whether the cortical responses at or near the acoustic spectral edges were modified compared to the neural responses induced by control stimuli (without spectral edges). As neural responses were not normally distributed in all conditions and frequencies tested, we chose to use non-parametric tests. The Wilcoxon signed-rank test was used to compare the same recording sites over different conditions (AFB condition vs. control condition, or AFB conditions between them). The threshold for the significance value was Bonferroni corrected. The differences between AFB stimuli and the control condition were compared at all frequencies over 4 octaves (±2 octaves on either side of the AFB center, 33 frequencies overall). A difference was considered significant if $p<0.05/33=0.0015$ at one or more frequencies. The statistical analysis was performed for MUA and LFP data separately. The Mann-Whitney test was used to compare independent samples, i.e. neural responses obtained at different cortical sites (comparison of neural responses for sites with BF corresponding to the upper edge of the AFB and those with BF corresponding to the lower edge of the AFB).
Results

Effects of contrast depth on the representation of spectral edges

For this experiment, 139 cortical sites were recorded from 8 animals. Figure 2 shows a representative example of MUA and LFP responses obtained for the different control and AFB conditions. In this example, the BF of the STRF derived from the LFPs and MUA was near 2000 Hz and the center frequency of the AFB corresponded to the BF. This example shows the two main results of this study. First, the responses at the edge-out frequencies were increased in all AFB conditions compared to the responses obtained in control conditions. Remarkably, there was a clear and sharp increase of neural responses at the edge-out frequencies even for the smallest depths tested (10 and 20 dB). Second, the responses within the notch frequency band were suppressed compared to the respective control conditions.

Figure 3 shows the medians of the effects on neural activity of AFB stimuli, relative to the control stimuli, for three positions of BF relative to the AFB center: when the AFB is centered on BF (at ± 1/8 octave, n=82 sites, middle panels), when the lower edge of the AFB is centered on BF (at ± 1/8 octave, n=25 sites, left panels) or when the upper edge of the AFB is centered on BF (at ± 1/8 octave, n=32 sites, right panels). The effects of the AFB stimuli relative to the control stimuli were tested statistically for both MUA and LFPs (see methods). In all relevant comparisons, the results were identical for both MUA and LFPs. Instead of reporting the identical effects for MUA and LFPs, we reported the results for MUA only. The increase in neural responses at the edge-out frequencies of the AFB were statistically significant in all AFB conditions and for the three positions of AFB center relative to the BF (p<0.0015). We also compared the AFB 70-60 condition to all other AFB conditions (3 comparisons) at the first edge-out frequencies of the AFB (for the three positions of AFB center relative to the BF). Neural responses at edge-out frequencies were significantly larger for large spectral contrasts (AFB 70-40 and AFB 70-0) compared to small spectral contrast (AFB 70-60) (p<0.017, after Bonferroni correction). The decrease in neural responses at the edge-in frequencies of the AFB were statistically significant in all AFB conditions and
positions of the AFB center relative to the BF (p<0.0015). We compared the neural responses at the lower edge-in frequency to those at the upper edge-in frequency. Consistent with our previous study (Catz and Noreña, 2013), we found that the neural suppression of responses at the lower edge-in frequency was significantly larger compared to the suppression at the upper edge-in frequency (p<0.05).

The percentages of recording sites showing at least 20% increase or decrease of neural activity for the three groups of neural responses (whether BF corresponded to the center of the AFB, or the lower or upper edge of the AFB) are shown in Figure 4. The percentages of the recording sites showing an increase of neural responses at edge-out frequencies ranged between nearly 100% for the AFB 70-00 condition and 50% for the AFB 70-60 condition. The percentages of the recording sites showing a decrease of neural responses at the edge-in frequencies ranged between 90% for the AFB 70-00 condition and 40% for the AFB 70-60 condition. These results suggest that the cortical changes induced by the AFB stimuli are very systematic for all the depths tested in this study, although less cortical sites are affected by spectral contrast with small depths. Remarkably, the percentage of sites showing a decrease of neural responses at edge-in frequencies was larger when BF corresponded to the lower edge of the AFB (50-90% of the cortical sites) than when BF corresponded to the upper edge of the AFB (30-70% of the cortical sites). This result is also consistent with our previous study (Catz and Noreña, 2013).

The time course of the edge enhancement

All the changes reported in the previous section have been induced by the 180-sec presentation of the AFB stimulus. While these changes were relatively rapid, it is unknown whether these changes occur over a shorter time period. Moreover, it is unclear whether these changes can persist after the end of the AFB stimulus presentation. We designed specific stimuli for answering these questions (see methods, Figure 1). For this experiment, 145 cortical sites were recorded from 10 animals. Figure 5 shows an example at a given cortical site for two widths of the AFB (0.5 and 1 octave, respectively) and for which the AFB
was centered on the BF (near 10 kHz). First, neural responses were globally enhanced in the
ctrl-sil condition compared to the ctrl-70 condition (Figure 5, columns 1-2). Second, AFB
stimuli presented in alternation with silence or control stimulus produced a sharp neural
enhancement at both edges of the AFB (Figure 5, columns 3-4 and 6-7). On the other hand,
the STRF derived from the control condition presented in alternation with the AFB stimulus
was unchanged compared to the control STRF (Figure 5, columns 5 and 8). These results
indicate that 10-sec presentation of the AFB stimulus is sufficient to produce a neural
enhancement at the spectral edges of acoustic stimuli and that this effect is short lived as
neural enhancement has disappeared within the 10 sec of ctrl-70 stimulus presentation (in
the ctrl-AFB condition).

Neural recordings were grouped according to the position of BF relative to the AFB
center: BF at the lower edge of the AFB (left panel, n=41 and 35 sites for the 1 and ½
octave-width AFB, respectively), BF at the center of the AFB (middle panel, n=72 and 65
sites for the 1 and ½ octave-width AFB, respectively), BF at the upper edge of the AFB (right
panel, n=32 and 28 sites for 1 and ½ octave-width AFB, respectively). The neural responses
obtained from the ctrl-sil and ctrl-AFB conditions were significantly increased at frequencies
of the AFB (and at adjacent frequencies in the ctrl-sil condition) compared to the ctrl-70
condition (p<0.0015) (Figure 6). More interestingly, neural responses obtained from AFB
stimuli in the ctrl-AFB and AFB-sil conditions were significantly increased at the edge-out
frequencies of the AFB (p<0.0015). Finally, we tested whether the edge enhancement
reported during the presentation of the AFB stimulus in the ctrl-AFB condition (AFB period,
Figure 1) was also present during the presentation of the control stimulation (ctrl period,
Figure 1): we did not find any significant increase of neural responses at spectral edges
(p>0.0015). In summary, these results indicated that the edge enhancement occurs within 10
sec of acoustic stimulation and disappears very shortly after the presentation of the AFB
stimulus (Figure 6).

The build-up and break down of neural changes during and after stimulation with AFB
stimulus, respectively, were further investigated with a finer temporal resolution, i.e. within
the following time windows: [0-2.5 sec], [2.5-5 sec] and [5-10 sec] (see methods). Our results indicate that the neural enhancement at the edge-out frequencies build-up and break down within a few seconds ([0-2.5 sec]) of stimulation with AFB stimulus and the consecutive ctrl stimulus, respectively (p<0.0015) (Figure 7).

Occasionally, we could get single-unit activity over an entire set of stimulus conditions. This provides an opportunity to apprehend whether the central changes observed for MUA is comparable to the putative central changes for single-unit activity. Figure 8 shows an example of single-unit activity for different stimulus conditions. This example shows a very clear increase of neural activity at the edge-out frequencies within 10-sec of stimulus presentation. This result suggests that the effects observed at the level of MUA are likely also present at the level of single-unit activity.

**Effects of anesthesia**

Three animals were implanted chronically with microelectrodes to investigate the effects of anesthesia on the spectral edge enhancement (see methods). Figure 9 shows the STRFs obtained simultaneously from 5 recording sites in an awake animal (channels 2, 4, 6, 13 and 14). The STRF is unchanged when it is remote or adjacent from the AFB (columns 1 and 3, respectively). On the other hand, for STRFs with BF close to the AFB center, the neural responses at the spectral edges of the AFB are greatly enhanced (columns 2 and 4-5). Remarkably, the neural enhancement at the lower spectral edge was particularly strong for the recording site shown in column 2 (as there was no neural response at these frequencies produced by the control condition). Figure 10 shows the medians of the effects on neural activity of AFB stimuli, relative to the control stimuli, for three positions of BF relative to the AFB center. The number of recording sites was 15, 30 and 19 for recordings with BF at the lower edge, the center or the upper edge of the AFB, respectively. The percentages of recording sites showing at least 20% neural increase are shown in the second row of Figure 10. As MUA and LFP have been documented above and since they show very similar results in awake animals as compared to anesthetized animals, LFPs are not shown. The neural
responses at the spectral edges of the AFB stimuli were significantly enhanced (p<0.0015, except when BF is at the lower edge of the AFB where statistical significance does not survive the Bonferroni correction: p=0.004).

Finally, the time course of the spectral edge enhancement was also investigated in two awake animals. Figure 11 shows an example of cortical site where the BF corresponded to the AFB center. One notes the very clear enhancement of neural responses at the spectral edges of the AFB in the AFB-sil and ctrl-AFB conditions. On the other hand, there was no enhancement at spectral edges in the ctrl-AFB condition (ctrl period). The results on this example are representative of the few other examples obtained in the two animals. Altogether, these results suggest that anesthesia plays only a minor role, if any, in the cortical changes reported in this study and our previous one (Catz and Noreña, 2013).
Discussion

In a previous study, we showed that the cortical representation of tone-pip frequencies near spectral edges is greatly “distorted”: the representation of frequencies outside a spectral notch is greatly enhanced, while the representation of frequencies inside a spectral notch is reduced. These cortical changes were exquisitely sensitive to the characteristics of the spectral notch, including the width of the notch and the slope of the spectral edges (Catz and Noreña, 2013). The present study extends on these initial findings. We found that the neural enhancement at spectral edges are larger for deeper spectral notches, but that it is statistically significant for depths as small as 10 dB. Moreover, the cortical changes at the spectral edges of the notch are rapid and labile as they build-up and break down within 2.5-sec of the multi-tone stimulus presentation with or without AFB, respectively. These results imply that the neural mechanisms (discussed below) involved in these central changes are necessarily rapid, labile and highly sensitive to small changes in the physical characteristic (depth, slope and width) of the spectral notches.

Mechanisms of the dynamic changes in spectral contrast representations

The rapid changes of cortical responses observed in the present study preclude the involvement of mechanisms with slower time constants such as long-term potentiation or depression, synaptic scaling or intrinsic plasticity (Buonomano and Merzenich, 1998; Turrigiano et al., 1998; Desai et al., 1999; Turrigiano, 2008; Grubb and Burrone, 2010; Kuba et al., 2010; Watt and Desai, 2010). Instead, they are likely the results of one or a combination of relatively fast mechanisms occurring on the order of seconds or less.

A first mechanism that comes to mind is synaptic inhibition (including “lateral” inhibition) as it has long been suggested to play a critical role in emphasizing the representation of edges and contours in audition and vision (Hartline et al., 1956; Hartline and Ratliff, 1957; Ratliff and Hartline, 1959; Ratliff et al., 1967; Von Békésy, 1967, 1969; Carterette et al., 1969a; Shamma and Symmes, 1985; Catz and Noreña, 2013). Synaptic inhibition has been proposed to be involved in the neural changes produced by two-tone sequences and after
hearing loss (Shamma and Symmes, 1985; Calford et al., 1993; Rhode and Greenberg, 1994; Rajan, 2001; Calford, 2002; Wang et al., 2002). Although synaptic inhibition may contribute to shape the tuning of cortical neurons, its effects are however limited in frequency range (cortical excitation and inhibition are approximatively co-tuned) and time, i.e. up to 100 msec after the stimulus onset (Wehr and Zador, 2003, 2005; Tan et al., 2004). One notes, however, that the respective tuning of excitation and inhibition may be different at the sub-cortical than at the cortical level. In particular, lateral (or side band) inhibition may play a major role in type “O” and “I” patterns of responses in the cochlear nucleus and inferior colliculus (Ramachandran et al., 1999; Ropp et al., 2014). It is conceivable that part of the results reported here result from neural changes produced at the sub-cortical level (see below). Disinhibition, i.e. inhibitory neurons inhibiting other inhibitory neurons (Pi et al., 2013), might also play a role in our results, producing a response enhancement at the edge-out frequencies. Disinhibition has been postulated as a possible mechanism accounting for the response enhancement in two-tone sequences (Brosch et al., 1999).

Another mechanism that may be involved in the dynamic representational changes reported in this study is short-term synaptic plasticity, including synaptic depression and facilitation. Synaptic transmission is a dynamic and context-dependent process that can contribute to shape the representation of rapidly evolving acoustic features (Brenowitz et al., 1998; Fortune and Rose, 2001, 2002; Zucker and Regehr, 2002; Abbott and Regehr, 2004; Oswald et al., 2006; Reyes, 2011). Synaptic depression is usually attributed to depletion of some readily releasable vesicle pools due to repeated stimulations at a rate above the vesicle turn-over rate. Synaptic depression can occur in the central auditory system but also as early as the synapses between inner hair cells and cochlear fibers (Goutman and Glowatzki, 2007; Zilany et al., 2009). In general, synapses with high probability release are subjected to synaptic depression and those with low probability release can demonstrate synaptic facilitation (Zucker and Regehr, 2002; Regehr, 2012). Other post-synaptic mechanisms can contribute to synaptic depression such as receptor desensitization (Jones and Westbrook, 1996; Otis et al., 1996).
Interestingly, inhibition and short-term synaptic plasticity may interact in a complex way to shape the spectro-temporal properties of cortical neurons. At the end-bulb glutamatergic synapse, for example, it has been shown that while GABA<sub>B</sub> pre-synaptic inhibition reduces the initial synaptic current evoked during a stimulation train, it can also greatly reduce the synaptic depression produced over time by the train (Brenowitz et al., 1998). By reducing the probability of neurotransmitter release, presynaptic inhibition produces less receptor desensitization and therefore causes a net increase of the synaptic transmission during a prolonged stimulation (Brenowitz and Trussell, 2001).

Any discussion relative to the mechanisms involved in our results is necessarily speculative. At this stage, we suggest that complex interactions between inhibition and short-term synaptic plasticity can account for the neural changes we observe at the spectral edges of an acoustic stimulus. Another important question relative to the mechanisms enhancing spectral contrasts is whether they operate at the cortical level or are inherited from lower levels of the auditory pathway. In broad agreement with our previous study (Catz and Noreña, 2013), the similar pattern of responses for MUA and LFPs suggests that the enhancement of spectral contrasts observed in the cortex is largely inherited from lower levels. Finally, our study also shows that anesthesia plays only a modest role, if any, in our results. This corroborates with an earlier study showing that anesthesia marginally affects post-stimulation suppression and facilitation produced by tone sequences (Brosch and Scheich, 2008).

**Functional implications**

Given the demonstration in the present study that the representation of spectral contrasts is rapidly enhanced in the auditory cortex for spectral depth as small as 10 dB, one can wonder whether these neural changes have functional implications. Importantly, the spectral profile of the stimuli used in the present study shows a clear contrast only when they are time-averaged over a few hundreds of milliseconds. This stimulus design (overlapping but asynchronous tone pips) was used to assess the representation of each tone pip
frequency for different spectral profiles. One can therefore wonder whether the results obtained from this particular stimulus can apply to other (more ecological) stimuli. Many natural sounds, such as vocalizations, usually last a few hundreds of milliseconds or more and often consist of a succession of similar acoustic motifs or syllables (Aubin and Jouventin, 2002; Huetz et al., 2009). Guinea pig vocalizations, for instance, while they may vary over time, usually maintain their harmonic structure over the course of the sequence (Berryman, 1976). The enhancement observed in the present study, implying temporal integration over a few hundreds of milliseconds, could therefore well apply to some natural temporal sequences.

Moreover, ecologically relevant stimuli are often composed of frequency-specific information such as spectral edges (Moore and Glasberg, 1983; Assmann and Summerfield, 2004; Palmer and Shamma, 2004). It seems that the auditory system relies heavily on these spectral cues. For instance, the detection of spectral peaks caused by vocal tract resonances (formants) can play a major role in speech recognition (Darwin, 1984; Assmann and Nearey, 1987; Henry et al., 2005). These acoustic features may also help differentiate relevant acoustic features from background noise (Darwin, 1984; Roberts and Moore, 1990). This hypothesis that spectral contrasts are important for speech recognition in noise has been further corroborated by psychoacoustic studies. They showed an improvement of intelligibility in noise for processed speech with enhanced spectral contrasts (increase of peak-to-valley ratio) (Simpson et al., 1990; Baer et al., 1993; Alexander et al., 2011). The perceptual effect of spectral enhancement was equivalent to a speech-to-noise ratio improvement of 4.2 dB (Baer et al., 1993). The increase and decrease of edge-out and edge-in frequency representation, respectively, may contribute to emphasize the representation of spectral edges and/or to prevent the loss of the spectral information degraded in a noisy environment (Rhode and Greenberg, 1994).

The neural enhancement at the spectral edges may also account for “Mach bands” in hearing, i.e. perceptual enhancement at the spectral edges of acoustic stimuli (Carterette et al., 1969b; Houtgast, 1972). In particular, this may account for the pitch induced by noise
bands at their spectral edges (Small and Daniloff, 1967; Bilsen, 1977). Indeed, low and high-pass noises have been shown to produce a pitch up to 10 kHz which is correlated with the cutoff frequency (Small and Daniloff, 1967; Bilsen, 1977). One notes that the high cutoff frequency of the noise (> 10 kHz), which can produce a pitch, precludes a mechanism based on a temporal enhancement, as no temporal information is available above 3.5-5 kHz (Rose et al., 1967; Palmer and Russell, 1986). The enhancement at spectral edges could also account for the dominant role played in pitch perception by the lowest and highest partials of a harmonic complex, especially when the low-numbered (resolved) partials are removed from the complex (Dai, 2000; Moore and Gockel, 2011).

Stimulation with notched stimuli (filtered noise or harmonic complex) is known to produce interesting auditory phenomena. The Zwicker tone is a tonal illusory percept produced immediately after the noise presentation. The pitch of this sensation lies within the frequency range of the notch (Zwicker, 1964; Wiegrebete et al., 1996; Noreña et al., 2000, 2002; Parra and Pearlmutter, 2007). Additionally, auditory enhancement can be produced immediately after the presentation of a broadband signal. A deleted frequency component from a harmonic complex perceptually “pops out” when it is reinserted. The detection thresholds for the component that was deleted are also lowered (Viemeister and Bacon, 1982; Summerfield and Assmann, 1989; Wiegrebete et al., 1996; Byrne et al., 2011). It is unknown whether the stimuli used here, which are composed of tone pips with asynchronous onsets, can produce Zwicker tone and/or auditory enhancement immediately after the presentation of the AFB stimuli. Our results suggest, however, that these two auditory phenomena are preceded (during stimulation with the notched stimulus) by a reduction of neural activity within the frequency range of the spectral notch. One speculates that this neural suppression might be followed by a rebound of activity immediately after the AFB stimulus presentation, that could account for the Zwicker tone and/or auditory enhancement (see Figure 7, ctrl-AFB (ctrl) condition with one-octave AFB, and Noreña and Eggermont, 2003a).

Artificial hearing loss, acoustic environment and long-term central changes
We and others have previously suggested that notched stimuli, such as our stimulus ensemble with AFB or notched noise, mimics the discontinuity in sensory inputs over frequency produced by sharp hearing loss (Pantev et al., 1999; Norena et al., 2000; Okamoto et al., 2007; Catz and Noreña, 2013). In this context, the stimulus ensemble used in the present study can be interpreted as being an equivalent of those producing an artificial scotoma in vision (Ramachandran and Gregory, 1991; Pettet and Gilbert, 1992; Kapadia et al., 1994; Das and Gilbert, 1995; DeAngelis et al., 1995; Weil et al., 2007, 2008; Parks and Corballis, 2012). The neural enhancement at the edge-frequencies of the spectral notch is reminiscent of the long-term central changes triggered by cochlear hearing loss, i.e. neural enhancement or even unmasking at the edge-frequency of hearing loss (Robertson and Irvine, 1989; Calford et al., 1993; Noreña and Eggermont, 2005). In case of an extensive exposure (for a few weeks) to an acoustic environment with spectral notch it can be speculated that the short-term changes reported in this study may be gradually converted into long-term changes (Noreña et al., 2006; Pienkowski and Eggermont, 2010; Pienkowski et al., 2013). Our results further suggest that chronic exposure to an acoustic environment with a spectral contrast as small as 10 dB and presented at a moderate level (~70 dB SPL) may be sufficient to produce chronic central changes. These chronic central changes, including changes in the pattern of spontaneous firing (Noreña et al., 2006), may have functional consequences, such as tinnitus and hyperacusis (Noreña, 2011; Noreña and Farley, 2013).

Acknowledgements

This work was supported by L’Agence Nationale de la Recherche Grant ANR-2010-JCJC-1409-1. The authors wish to thank Martin Pienkowski for valuable comments on a previous version of the manuscript and Nicolas Catz and Sophie Cavé for their contribution to data collection.


Depireux DA, Simon JZ, Klein DJ, Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex.


Pettet MW, Gilbert CD. Dynamic changes in receptive-field size in cat primary visual cortex.


Simpson AM, Moore BC, Glasberg BR. Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners.


**Figure captions**

**Figure 1**

Schematic representation of the acoustic stimuli used in the present study. A. Time onsets (black dots) and long-term spectrum of control (left) and AFB stimuli (right). B. Time onsets (black dots) of tone pips for stimulus sequences alternating control stimulus and silence (ctrl-sil), AFB stimulus and silence (AFB-sil) or control stimulus and AFB stimulus (ctrl-AFB).

**Figure 2**

Neural tuning of individual recordings obtained from a selected example at a given location in the primary auditory cortex. Each column corresponds to a stimulus condition (indicated at the top of panels). First and second rows: spectro-temporal receptive fields obtained from control stimuli (multi-tone ensemble without AFB) and from MUA and LFPs, respectively. Third and Fourth rows: spectro-temporal receptive fields obtained from multi-tone ensemble with AFB and from MUA and LFPs, respectively. Horizontal dotted lines represent the edge frequencies of the AFB. Neural activity in the spectro-temporal receptive field is represented with a color continuum from blue (minimum values) to red (maximum values). Fifth and sixth rows: frequency profiles obtained by taking the maximum firing rate of MUA, or by taking the minimal amplitude of LFPs, in the 10-30 msec time window, respectively. The red and black lines represent the neural responses for the stimulus ensemble with AFB and without AFB, respectively. Control conditions at 60, 50 and 40 dB are represented with a blue line. Vertical dotted lines represent the edge frequencies of the AFB. Neural responses are greatly enhanced at the edges of the AFB and decreased within the AFB.

**Figure 3**
Medians of the difference between the AFB conditions and the control conditions for three positions of the AFB relative to the neural BF, as a function of frequency. Each column represents a position of the AFB relative to the neural BF. The first and second rows represent the average MUA and LFP, respectively. First column: average data for neurons with BF corresponding to the lower edge of the AFB. Second column: average data for neurons with BF corresponding to the AFB center. Third column: average data for neurons with BF corresponding to the upper edge of the AFB. Vertical dotted lines represent the edge frequencies of the notch. Circles indicate statistically significant differences between control and AFB conditions for MUA (results are identical for LFPs and are not shown). Neural responses are enhanced near and outside the notch, while they are reduced near and within the notch.

<table>
<thead>
<tr>
<th>Position relative to neural BF</th>
<th>Average MUA</th>
<th>Average LFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower edge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFB center</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper edge</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4**

Percentage of sites showing an increased response (first and second rows, for MUA and LFPs, respectively), or a decreased response by at least 20%. (third and fourth rows, for MUA and LFPs, respectively). Otherwise as in Figure 3.

**Figure 5**

Neural changes produced by 10-sec AFB stimulus alternated with 10-sec silence or control stimulus. Stimulus condition and time period (between parentheses) are indicated at the top of each panel. First and second rows: spectro-temporal receptive fields obtained from MUA and LFPs, respectively. Third and fourth rows: frequency profiles of neural responses for MUA and LFPs, respectively. AFB bandwidth (Nw, indicated at the top of panels) was 0.5 octave in columns 3-5 and 1 octave in columns 6-8. The black and red lines represent the neural responses for the stimulus ensemble without AFB and with AFB, respectively. The black lines from column 1 (for MUA and LFPs) are replicated in columns 2-8 for comparison. Neural responses are greatly enhanced at the edges of the notch and decreased within the notch. The neural changes at the spectral edges are fast, i.e. they
occur within 10-sec of the stimulus with AFB presentation (columns 3-4 and 6-7), and labile, i.e. they disappear within the following 10-sec of the stimulus ensemble without AFB presentation (columns 5 and 8).

**Figure 6**

Medians of the difference between the various stimulus conditions and the control condition (ctrl-70) for three positions of the AFB relative to the neural BF, as a function of frequency and for the different stimulus conditions. Each column represents a position of the AFB relative to the neural BF (see Figure 3). The first and second rows represent the results for MUA and for the 0.5 and 1-octave wide notch (Nw), respectively. The third and fourth rows represent the results for LFPs and for the 0.5 and 1-octave wide notch, respectively. Vertical dotted lines represent the edge frequencies of the AFB. Circles indicate statistically significant differences for MUA (results are identical for LFPs and are not shown). Neural responses at both edges of the AFB are largely enhanced, while they are reduced within the AFB.

**Figure 7**

Build-up and break down of neural changes produced by the AFB stimulus alternated with control stimulus or silence. Stimulus condition and time period (between parentheses) are indicated at the top of each panel. The 10-sec stimulation periods (ctrl or AFB) have been divided into three time windows, i.e. [0-2.5 sec], [2.5-5 sec] and [5-10 sec]. First and second rows show the neural changes for stimulus sequence with AFB of 1 or 0.5 octave bandwidth, respectively. The neural enhancement at the edge-out frequencies build-up and break down within a few seconds ([0-2.5 sec]) of stimulation.

**Figure 8**

Cortical responses evoked by stimulus ensembles with or without AFB for individual recordings where a single-unit activity over all the stimulus conditions could be
obtained. Stimulus condition and time period (between parentheses) are indicated at the top of each panel. First row: individual (and averaged – red line) waveforms of the single-unit activity. Second row: Spectro-temporal receptive fields in all stimulus conditions. Third row: frequency profiles of cortical responses in all stimulus conditions. One observes a clear response enhancement at the edge-out frequencies, in particular on the upper edge of the spectral notch.

Figure 9
Effects of anesthesia on the neural enhancement produced by AFB stimulus. Five individual examples of cortical responses evoked by the stimulus ensemble without (first and second rows, MUA and LFPs, respectively) or with AFB (third and fourth rows, MUA and LFPs, respectively) obtained simultaneously in an awake animal. Fifth and sixth rows: frequency profiles of cortical responses for MUA and LFPs, respectively. One observes a clear response enhancement at the edge-out frequencies, suggesting that anesthesia did not play a major role in our results.

Figure 10
Medians of the cortical changes induced by the stimulus ensemble with AFB on awake animals. Each column corresponds to a position of the cortical BF relative to the AFB (see figures 3 and 6). First row: Medians of the difference between the AFB conditions and the control conditions. Circles indicate statistically significant differences between control and AFB conditions. Second row: Percentage of sites showing an increased response by at least 20%.

Figure 11
Time-course of the neural changes produced by AFB stimulus with one-octave wide spectral notch in an awake animal. Each column corresponds to a stimulus condition and period (indicated at the top of panels). First and second rows: spectro-temporal receptive
fields obtained from MUA and LFPs, respectively. Third and fourth rows: frequency profiles of
neural responses for MUA and LFPs, respectively. The black and red lines represent the
neural responses for the stimulus ensemble without AFB and with AFB, respectively. The
enhancement of neural activity at the spectral edges are fast, i.e. they occur within the 10-
sec presentation of the stimulus ensemble with AFB (column 3-4), and labile, i.e. it
disappears within the following 10-sec presentation of the stimulus ensemble without AFB
(column 5). These results are comparable to those reported in anesthetized animals.