Frequency tagging to track the neural processing of contrast in fast, continuous sound sequences

Article *in* Journal of Neurophysiology · April 2017

DOI: 10.1152/jn.00971.2016

3 authors, including:

Sylvie Nozaradan
Université Catholique de Louvain
21 PUBLICATIONS **525** CITATIONS

FREQUENCY-TAGGING TO TRACK THE NEURAL PROCESSING OF CONTRAST IN FAST CONTINUOUS SOUND SEQUENCES

Sylvie Nozaradan[1,2,3], André Mouraux[1], Marion Cousineau[3]

[1] Institute of Neuroscience (Ions), Université catholique de Louvain (UCL), 53 (UCL 53.75) Mounier avenue, B-1200, Belgium
[2] The MARCS Institute for Brain, Behavior and Development (WSU), Sydney, Australia
[3] International laboratory for Brain, Music and Sound Research (Brams), 1430 Mont Royal boul., H3C 3J7, (QC) Montreal, Canada

Corresponding author:
Dr. Sylvie Nozaradan
Institute of Neurosciences
Université catholique de Louvain
53, Avenue Mounier – UCL 53.75
B-1200 Bruxelles
Belgium
Phone: +32-2-764-9349
Email: sylvie.nozaradan@uclouvain.be

Running title: Contrast-specific neural responses in fast sound sequences

Keywords: EEG, frequency-tagging, auditory contrast, sensory discrimination

ABSTRACT

The human auditory system presents a remarkable ability to detect rapid changes in fast continuous acoustic sequences, as best illustrated in speech and music. However, the neural processing of rapid auditory contrast remains largely unclear, probably due to the lack of methods to objectively dissociate the response components specifically related to the contrast from the other components in response to the sequence of fast continuous sounds. To overcome this issue, we tested a novel use of the frequency-tagging approach allowing contrast-specific neural responses to be tracked based on their expected frequencies. The EEG was recorded while participants listened to 40-s sequences of sounds presented at 8Hz. A tone or inter-aural time contrast was embedded every fifth sound (AAAAB), such that a response observed in the EEG exactly at 8Hz/5 (1.6Hz) or harmonics should be the signature of contrast processing by neural populations. Contrast-related responses were successfully identified, even in the case of very fine contrasts. Moreover, analysis of the time course of the responses revealed a stable amplitude over repetitions of the AAAAB patterns in the sequence, except for the response to perceptually salient contrasts which showed a buildup and decay across repetitions of the sounds. Overall, this new combination of frequency-tagging with an oddball design provides a valuable complement to the classic transient evoked potentials approach, especially in the context of rapid auditory information. Specifically, we provide objective evidence on the neural processing of contrast embedded in fast continuous sound sequences.

NEW & NOTEWORTHY

Recent theories suggest that the basis of neurodevelopmental auditory disorders such as dyslexia might be an impaired processing of fast auditory changes, highlighting how the

encoding of rapid acoustic information is critical for auditory communication. Here, we present a novel electrophysiological approach to capture in humans neural markers of contrasts in fast continuous tone sequences. Contrast-specific responses were successfully identified, even for very fine contrasts, providing direct insight on the encoding of rapid auditory information.

INTRODUCTION

The human auditory system presents a remarkable ability to detect changes even when embedded in fast continuous sequences, as best illustrated in speech and music. For example, understanding speech or enjoying a musical piece requires the ability to process acoustic contrasts occurring at a fast rate, often low within the sub-second range (Bregman 1990). Psychoacoustic investigations of the phenomenon have shown that the rate of the sequence in which the contrast is embedded is a critical parameter influencing contrast processing (Freyman and Nelson 1985; Albouy et al. 2016). In normal listeners, the ability to discriminate between two successive sounds of distinct frequencies is positively correlated with the duration of the sound and is also enhanced by the presence of a silent gap between the sounds (Freyman and Nelson 1985; Albouy et al. 2016). Specifically, as duration increases, the sensitivity to frequency contrast between two successive tones increases rapidly over a range of short durations (starting at 5 ms duration), increases more gradually for longer durations (100 ms), to reach an asymptote around 200 ms (Freyman and Nelson 1985). This observation has been interpreted as an effect of backward masking through which the perceptual analysis of a sound, or accumulation of evidence, is stopped by a subsequent sound presented soon after the first one, thus explaining the reduced discrimination abilities observed with reduced tone duration (Massaro and Idson 1977; Demany and Semal, 2005, 2007). Relatedly, difficulties in processing brief, rapidly changing, acoustic information have been described in auditory neurodevelopmental disorders such as dyslexia, specific language impairments and congenital amusia (Tallal et al. 1993; Tallal 2004; Goswami 2015; Albouy et al. 2016), highlighting the importance of processing auditory information that arrives rapidly and sequentially.

To investigate the underlying neural mechanisms, electroencephalographic (EEG) recordings appear as an optimal means to obtain measures of contrast processing with high temporal resolution. Moreover, these measures present the advantage of being independent from behavioral performance, thus preventing contamination by irrelevant decisional processes that could possibly bias these measures, especially in impaired individuals. Typically, in these EEG studies, the detection of auditory contrasts is associated with a particular brain response referred to as the mismatch negativity (MMN), and identified in the human EEG around 150-200 ms from contrast onset (for recent reviews, see e.g. Näätänen et al. 2007; May and Tiitinen 2010; Escera and Malmierca 2014). More recently, a number of studies have shown that earlier components in the range of mid-latency auditory-evoked potentials (i.e. a series of positive and negative deflections occurring between 10 and 50 ms after stimulus onset) can also be modulated by a contrast in acoustic features such as a contrast in the frequency of tones (further referred to as "tone contrast" in the current study; Kraus et al. 1994; Escera and Malmierca 2014; Nelken 2014; see Grimm et al. 2012, 2016 for a review).

However, most of these EEG studies have investigated deviant-related neural responses using relatively long inter-stimulus intervals (usually longer than 150 ms, between 200 and 500 ms on average, including silent gaps between successive items). In contrast, there is surprisingly few studies investigating the neural processing of contrasts embedded in fast continuous sequences, although this range seems critical to understand the brain mechanisms underlying the processing of brief, rapidly changing, acoustic information and its impairment (e.g. in neurodevelopmental disorders such as dyslexia; Albouy et al. 2016). Moreover, previous studies measuring MMN potentials have failed to capture significant deviance responses especially when the contrast was regular and thus predictable and was

presented at high rate, as is most often the case in speech and music for instance (see Sussman et al. 1998; Grimm et al. 2016).

The current study aimed to investigate the neural processing of contrasts especially when embedded in fast continuous sound sequences, that is, when the contrast is presented at a rate that appears to be critical for auditory communication in daily-life activities and its impairment. To this aim, we used a new alternative approach which combines an oddball design and a frequency-tagging method (see e.g. Liu-Shuang et al. 2014; Rossion 2014; Norcia et al. 2015; de Heering and Rossion 2015 for the first studies using this approach in high-level vision). This approach was proposed as a means to objectively dissociate contrast-specific responses from the other response components to the continuous sequence of inputs based on the frequency one would expect given the structure of the sequence, irrespective of the rate with which stimuli are presented. In the current study, the continuous sequence of inputs consisted in pure tones of 125 ms duration repeated during 40 s. Hence, the response to the envelope of acoustic energy was expected to elicit a periodic EEG response at 8 Hz (and harmonics), corresponding to the frequency of the tone repetition in the sequence (Fig. 1). Most importantly, a contrasting stimulus was introduced every 5^{th} tone in the sequence according to an AAAAB pattern, such that an EEG response observed exactly at 8Hz/5 (1.6Hz) and harmonics would be the signature of contrast detection and/or its consequence by neural populations.

In two experiments, we tested two different types of auditory contrasts, tone or inter-aural time contrast. In the first experiment, the tone contrast was either undetectable, or small or large according to the Western musical scale of tones. With the rate of presentation used here, the large contrast was also expected to induce a perceptual effect of “stream segregation”, in which the A and B tones would be perceived as distinct auditory streams

(van Noorden 1975; Bregman 1990). Furthermore, because contrast responses observed in this first experiment could be due to effects occurring at a peripheral level — as stimuli of different tones do not activate the same peripheral cochlear afferents — we conducted a second experiment in which a difference in inter-aural timing was used to induce a contrast in the perceived spatial location of A and B tones. In this second experiment, the peripheral cochlear channels activated by the A and B tones were strictly identical, and the observation of a contrast-specific EEG response would thus demonstrate that the processing of auditory contrast embedded in fast continuous sequences does not rely solely on a process of peripheral cochlear channeling.

MATERIAL AND METHODS

Participants

Twenty-four healthy participants took part in the study after providing written informed consent, with twelve volunteers in Experiment 1 (5 males, all right-handed except one, mean age 27±3 years) and Experiment 2 (6 males, all right-handed, mean age 29±4 years) respectively. All participants were mostly familiar with music produced according to the Western musical scale but none were professional musicians (no training of more than 10 years). They had no history of hearing, neurological or psychiatric disorder, and were not taking any drug at the time of the experiment. The study was approved by the Ethics Committee of the Catholic University of Louvain (UCL, Belgium).

Experiment 1: tone contrast

The stimulus consisted of a pattern made of alternated A and B tones, in the form AAAAB (Fig. 1A). This pattern was continuously looped to generate 40-s sequences. A and B were pure tones of 125 ms duration (i.e. 8 Hz presentation rate), with 10 ms rise and fall cosine ramps. Given the structure of the sequence, an EEG response observed at 8 Hz and harmonics should correspond to the neural response elicited by the acoustic energy of the sequence (further referred to as the "base response"), while an EEG response observed at 8 Hz/5 and harmonics should correspond to a neural response specifically related to the *contrast* between A and B sounds ("contrast response") (Liu-Shuang et al. 2014; Rossion 2014; de Heering and Rossion 2015; Jonas et al. 2016). Note that the term "contrast response" refers to the differential response between A and B tones, which would arise if and only if the responses elicited by the A and B tones substantially differ from each other (see e.g. Lochy et al. 2016, Retter and Rossion 2016).

In condition 1 (baseline condition), the contrast between A and B was 0.02 semitone (tone A = 1000 Hz, tone B = tone A + 0.02 semitones = 1001.2 Hz). In condition 2, the contrast was 0.50 semitone (tone A = 1000 Hz, tone B = 1029.3 Hz). In condition 3, the contrast was 4 semitones (tone A = 1000 Hz, tone B = 1259.9 Hz) (Fig. 1B). These sound sequences were presented in three separate blocks for the three conditions. The order of the blocks was counter-balanced across participants. In each block, the 40-s auditory sequence was repeated 8 times (8 trials per condition). The onset of each sequence was self-paced, and preceded by a 3-s foreperiod. The experimenter remained in the recording room at all times, to monitor compliance to the procedure and instructions throughout the experiment. To ensure that participants focused their attention on the sound, they were asked to carefully listen to the sound and report at the end of each sequence any irregularity in duration, pitch and intensity of the sounds over the sequences. There was no actual irregularity in any of the sequences, nevertheless, participants did report detecting subtle changes throughout the sequences in about half of the trials on average.

Based on previous behavioral studies on contrast detection and auditory stream segregation (see e.g. van Noorden 1975; Bregman 1990), a contrast of 0.02 semitone embedded in a continuous sequence presented at 8 Hz was expected to be undetectable. Conversely, 0.50 and 4 semitone contrasts were expected to be detectable. In addition, the sequence with 4 semitone contrast was expected to induce stream segregation. To confirm this, every participant was asked at the end of each condition whether “there was one or two distinct notes repeated over the sequences”, as an index of discrimination across A and B tones. They were also asked whether they had “a feeling that an additional, high-pitched, sound was played on top of the stream of fast repeated low-pitched tone”, as an index of stream segregation between A and B tones. The auditory stimuli were created and presented using

Matlab (The MathWork, USA) with binaural presentation through headphones at a comfortable hearing level corresponding to ~70dB SPL (BeyerDynamic DT 990 PRO, Germany).

Experiment 2: inter-aural time contrast

The sequence design and task of Experiment 2 were identical to those of Experiment 1 except for the fact that the contrast between A and B tones was based on a difference in inter-aural timing, inducing an illusion of spatial contrast. Because inter-aural timing contrasts are known to be processed in the central relays of the ascending auditory pathway after the cochlea (see e.g. Yost 2000), Experiment 2 allowed testing whether contrast-related responses embedded in fast continuous auditory sequences and identified with the frequency-tagging approach could be elicited by mechanisms other than differences in peripheral cochlear channeling.

Such as in Experiment 1, an undetectable 0.02 semitone frequency contrast between A and B tones was used as baseline condition (tone A = 200 Hz, tone B = tone A + 0.02 semitones = 200.2 Hz). Lower tone frequencies were used in Experiment 2 as compared to Experiment 1, because a longer cycle period was required to introduce the inter-aural time delay. In condition 2, tones A were presented with the left/right ear preceding the right/left ear by a time lag of 600 µs while tones B were presented with the right/left ear preceding the left/right ear by 600 µs, respectively (Fig. 2). This inter-aural timing contrast induced an illusion of spatial source contrast between A and B tones, and was also expected to induce a perceptual phenomenon of stream segregation based on these perceived separate spatial sources for A and B tones (segregated auditory streams are usually associated to distinct sound sources; see van Noorden 1975; Bregman 1990). To confirm this, participants were asked at the end of each condition whether or not they had “a feeling that one sound was

coming from one side on top of the stream of fast repeated tones coming from the opposite side". The side of the perceived location of A and B tones was counter-balanced across participants.

Sound sequence analysis

The envelope of the different 40-s sound sequences of Experiments 1 and 2 was extracted using the Hilbert function implemented in Matlab, to obtain a time-varying estimate of the instantaneous amplitude of the sound envelope (Nozaradan et al. 2016a; 2016b; 2016c; Cirelli et al. 2016). The obtained waveforms were then transformed in the frequency domain using a discrete Fourier transform to obtain a frequency spectrum of envelope magnitude, as displayed in Figures 1 and 2 (for Experiments 1 and 2 respectively). Hence, this analysis yielded a model of response corresponding to the envelope of acoustic energy of the sequences.

The sequences were also analyzed using a cochlear channel model (Gammatone filter bank modeling the peripheral processing performed by cochlear channels), as implemented in the Auditory Toolbox running on Matlab (Slaney 1998). The obtained bands centered on the carrier frequencies of the A and B tones were then transformed in the frequency domain using a Fourier transform (Figs. 1 and 2). This analysis yielded a cochlear channels model or prediction of responses to the contrasts embedded in the sequences, to which the EEG data could be further compared.

EEG recording

Participants were comfortably seated in a chair with the head resting on a support. They were instructed to relax, avoid any unnecessary head or body movement and keep their eyes fixated on a point displayed on a computer screen in front of them. The EEG was recorded using 64 Ag-AgCl electrodes placed on the scalp according to the International

10/10 system. Vertical and horizontal eye movements were monitored using four additional electrodes placed on the outer canthus of each eye and on the inferior and superior areas of the left orbit. Electrode impedances were kept below 10 kΩ. The signals were amplified, low-pass filtered at 500 Hz, digitized using a sampling rate of 1000 Hz and referenced to an average reference (64-channel high-speed amplifier, Advanced Neuro Technologies, The Netherlands).

EEG analysis

The continuous EEG recordings were filtered using a 0.1-Hz high-pass FFT filter to remove very slow drifts in the recorded signals. Epochs lasting 40 s were obtained by segmenting the recordings from +0 to +40 s relative to the onset of the auditory sequence. Artifacts produced by eye blinks or eye movements were removed using a validated method based on an Independent Component Analysis applied on the entire set of EEG data (Jung et al. 2000) using the runica algorithm (Bell and Sejnowski 1995; Makeig 2002), leading to the rejection of one independent component per participant. For each subject and condition, EEG epochs were averaged across trials (i.e. 8 trials averaged per condition and participant). The time-domain averaging procedure was used to enhance the signal-to-noise ratio of EEG activities time-locked to the input sequences (and to attenuate the contribution of possible artifacts, e.g. due to heart rate). The obtained average waveforms were then transformed in the frequency domain using a discrete Fourier transform, yielding a frequency spectrum of signal amplitude (μV) ranging from 0 to 500 Hz with a frequency resolution of 0.025 Hz. These EEG processing steps were carried out using Letswave 5 and 6 (http://nocions.webnode.com/) and Matlab (The MathWorks, USA).

Within the obtained frequency spectra, the amplitudes may be expected to correspond to the sum of (i) responses elicited by the stimulus and (ii) unrelated residual background noise

due, for example, to spontaneous EEG activity, muscle activity or eye movements. Therefore, to obtain valid estimates of the responses, the contribution of this noise was removed by subtracting, at each bin of the frequency spectra, the average amplitude measured at neighboring frequency bins (2^{nd} to 12^{th} frequency bins relative to each bin), for each participant, condition and electrode. The validity of this subtraction procedure relies on the assumption that, in the absence of a periodic EEG response, the signal amplitude at a given frequency bin should be similar to the signal amplitude of the mean of the surrounding frequency bins (Mouraux et al. 2011; Nozaradan et al. 2012a, 2015, 2016a; Chemin et al. 2014). The magnitudes of the contrast and base responses were obtained for each condition and participant by averaging the amplitude over a pool of fronto-central electrodes (Fz, FCz, Cz, F1, F2, FC1 and FC2) based on the topographical maps of responses to auditory contrasts as obtained in most previous studies using oddball designs (Näätänen et al. 2007).

The magnitudes of the responses were then estimated from these averaged activities by taking the noise-subtracted amplitude measured at the exact frequency of the expected responses (8 Hz for the base response, and 8 Hz/5 = 1.6 Hz for the contrast response). Because the sound envelope itself was not a pure sine wave, and because the shape of the EEG response does not necessarily match the shape of the sound envelope, the periodic EEG response elicited by the periodic sound envelope was not expected to correspond to a pure sine wave. Therefore, the frequency-domain representation of this periodic activity was expected to appear as a series of peaks at the expected frequencies and upper harmonics (Liu-Shuang et al. 2014). To determine the number of harmonic frequencies to take into account, one-sample t-tests against zero were computed on the noise-subtracted amplitudes obtained for each participant in each condition. In the absence of a significant

response, the noise-subtracted signal amplitude may be expected to tend towards zero (Mouraux et al. 2011). Harmonics were analyzed until they were no longer significant in either condition (Liu-Shuang et al. 2014). This yielded 4 harmonics for the contrast response (1.6, 3.2, 4.8 and 6.4 Hz) and 2 harmonics for the base response (8 and 16 Hz). For each response, these amplitudes were averaged across harmonics for further tests. The average of the amplitudes over the harmonics relies on the assumption that these activities are related to the same phenomenon (see e.g. Appelbaum et al. 2006; Heinrich et al. 2009; see also Retter and Rossion 2016 for an extensive justification for grouping amplitudes across harmonics to quantify the magnitude of the EEG contrast or base responses).

Statistical evaluation

Statistical analyses were performed using SPSS Statistics 21.0 (IBM, Armonk, NY, USA). Significance level was set at $p < 0.05$. All t-tests were Bonferroni-corrected for multiple comparisons by multiplying the obtained p values by the number of performed comparisons. When relevant, the Greenhouse-Geisser correction was used to correct for violations of sphericity in the performed ANOVAs.

We first examined whether the auditory stimulus elicited EEG responses at the expected frequencies across the conditions of the two Experiments. To this aim, a one-sample t-test was used to determine whether the amplitudes of the expected base and contrast responses were significantly different from zero.

Furthermore, the base and contrast responses were also compared with each other and across conditions for each experiment. These comparisons were performed using a two-way repeated-measures ANOVA with the factors ‘response' (two levels: contrast and base response) and ‘condition’ (Experiment 1: three levels: 0.02, 0.50 and 4 semitones contrast;

Experiment 2: two levels: 0 μs and 600 μs inter-aural time contrast). When significant, partial ANOVAs and paired-sample t tests were used to perform *post hoc* comparisons.

The obtained EEG responses were also compared to the responses predicted by the cochlear model. The aim of this comparison was to test whether the relative amplitude of the contrast and base responses observed in the EEG could be explained by the peripheral processing of the sequences for tone contrasts. One-sample t-tests were used to compare across participants the ratio between the amplitude measured at 8 Hz/5 and 8 Hz with the corresponding ratio obtained from the cochlear channels model for each condition. Moreover, these ratio values were also compared across conditions using a one-way ANOVA and paired-sample t tests.

In addition to the frequency domain analysis of the 40-s epochs, we also analyzed the time course of the EEG responses within AAAAB pattern. For each participant and condition, the 40-s epochs were first filtered between 1 and 17 Hz (FFT band-pass filter, 0.1 Hz slope), thus preserving the contrast and base responses and their harmonics while filtering out the EEG activity unrelated to these responses. The 40-s epochs were then segmented into a series of 0.625-s chunks, corresponding to the temporal window between the onsets of each AAAAB pattern in the sequences, and the chunks were averaged. Note that this event-related analysis does not exclude the possibility of overlapping oscillatory activity from previous trials. Instead, it provides a time domain analysis of the EEG frequency-tagged responses to characterize the shape of these responses and their modulation across conditions at a fine temporal scale, complementary to the frequency domain analysis.

Because there was no a priori assumption regarding the latencies at which a difference could arise between these waveforms, a point-by-point comparison of the waveforms over the averaged chunk 0.625-s duration was performed. To account for the multiple

comparisons, a cluster-based permutation approach was used (cluster-level statistics, representing an analysis of continuous data based upon inference about clusters and randomisation testing; see e.g. Maris and Oostenveld 2007; van den Broeke et al. 2015). The technique assumes that true neural activity will tend to generate signal changes over contiguous time points (Groppe et al. 2011). In Experiment 1, the waveforms of the different conditions were compared by means of a one-way repeated-measures ANOVA comparing the three conditions. In Experiment 2, the waveforms were compared using a paired-sample t-test comparing the two conditions. Clusters of contiguous time points above the critical F- or t-value for a parametric two-sided test were identified, and an estimate of the magnitude of each cluster was obtained by computing the sum of the F-values or absolute t-values constituting each cluster. Random permutation testing (1000 times) of the waveforms of the different epochs (performed independently for each subject) was then used to obtain a null distribution of maximum cluster magnitude. Finally, the proportion of random partitions that resulted in a larger cluster-level statistic than the observed one (i.e. p value) was calculated. Clusters in the observed data were regarded as significant if they had a magnitude exceeding the threshold of the 95th percentile of the permutation distribution of the F-statistic and t-statistic.

Finally, the time course of the amplitude of the frequency-tagged responses *across* repetitions of the AAAAB patterns was analyzed over the 40-s sequence, thus at a longer time scale. A sliding FFT was computed on the averaged 40-s epochs for each condition and participant. A temporal window of 6.25 s (10 repetitions of the AAAAB pattern) moving by steps of 0.625 s (i.e. the duration of one AAAAB pattern) was used as a compromise between temporal resolution, spectral resolution (0.15 Hz resolution obtained in the FFT as calculated on 6.25-s windows) and signal-to-noise ratio of the responses. This analysis

yielded 54 successive values of amplitudes for the contrast and base responses. As for the unwindowed FFT spectra, the obtained values were corrected by subtracting, at each target frequency (base and contrast responses and their harmonics) the amplitude measured at a neighboring frequency bin (the third upper frequency bin relative to each target bin was taken instead of the 2^{nd} - 12^{ve} bin range as used in the 40-s unwindowed FFT spectra, because the frequency resolution was coarser in the windowed FFT as compared to the unwindowed FFT and a 2^{nd} - 12^{ve} bin range would thus overlap with the frequencies of interest). A one-way ANOVA was then used to compare these values over time. Although successive values in the series were not completely independent from each other due to the overlap between successive time windows, the ANOVA tested the null hypothesis that the values would not be significantly different across the sequence. If significant, the presence of a linear trend was tested. Importantly, the aim of this analysis was not to determine the best-fit model explaining the dynamics of amplitude across time points, but rather to test whether these amplitudes significantly decreased or increased over the sequence, as an estimation of the buildup or decay of the response over the sequence.

RESULTS

Sound sequence analysis

In Experiment 1, the spectrum of the envelope of acoustic energy consisted of a peak at 8 Hz (and harmonics) for all the different sequences, corresponding to the 125-ms duration of the sequential sounds (Figs. 1 and 2). The cochlear channels model of response also consisted of a peak at 8 Hz (and harmonics) in all conditions. However, an additional peak at 8 Hz/5 (1.6 Hz and harmonics: 3.2 Hz, 4.8 Hz and 6.4 Hz) was obtained for the sequences with 0.50 and 4 semitones contrast (conditions 2 and 3). In condition 2, this contrast

response was relatively small compared to the base response (contrast/base ratio of 0.39). In condition 3 the contrast response was prominent relative to the base response (contrast/base ratio of 4.21).

Expectedly, the spectrum obtained for the cochlear channels model in the 0.02 semitone contrast (condition 1) did not exhibit a significant peak at 8 Hz/5 and harmonics, as such a small tone contrast is not resolved by the cochlear channels. This was also the case for Experiment 2, where the cochlear channels models only exhibited a peak at 8 Hz but not at 8 Hz/5, as the inter-aural time contrast is not processed by the cochlear channels (Fig. 2).

EEG data

In both Experiments 1 and 2, the grand-average spectra (Fig. 3) exhibited a clear base response over fronto-central electrodes, demonstrating successful synchronization to the envelope of acoustic energy of the sequences in all conditions. Moreover, when the contrast was detectable (conditions 2 and 3 of Experiment 1, condition 2 of Experiment 2), contrast responses appeared as additional peaks in the EEG spectra. The one-sample t-tests comparing the noise-subtracted amplitude of these responses against zero (as detailed in Table 1) confirmed these observations. Note, however, the small amplitude obtained for the small 0.50 semitone contrast response in Experiment 1, suggesting some just above threshold response with such a small contrast (10 participants/12 with > 0.01 μV in this condition, as opposed to 0 participants/12 in condition 1). In both experiments, these contrast responses exhibited similar fronto-central topographies than those of the base response, as shown in Figure 3.

Experiment 1. The two-way ANOVA used to compare the amplitudes of the base and contrast responses (factor 'response') across the three conditions (factor ‘condition’) (Fig. 4) revealed a significant interaction between the two factors ($F(1.587, 17.462) = 20.654$, η^2 =

0.652, $p < 0.0001$). This interaction was explained by the fact that the base response was not significantly different across the three conditions (condition 1-2: $p = 0.748$, condition 2-3: $p = 0.441$, condition 1-3: $p = 0.100$), as opposed to the contrast response (condition 1-2: $p = 0.043$, condition 2-3: $p = 0.003$, condition 1-3: $p < 0.0001$).

Experiment 2. Such as in Experiment 1, the two-way ANOVA used to compare the amplitudes of the base and contrast responses across the two conditions showed a significant interaction between the two factors ($F(1, 11) = 33.879$, $\eta^2 = 0.755$, $p < 0.0001$). This interaction was also explained by the fact that the base response was not significantly different across conditions ($p = 0.253$), as opposed to the contrast response ($p<0.0001$) (Fig. 4).

*Comparison between EEG responses and the cochlear responses predicted by the cochlear channels model (*for Experiment 1, tone contrast). No significant difference was observed between the EEG responses and the responses predicted by the cochlear channels model in condition 1 with 0.02 semitone contrast, as no significant contrast response was observed in the EEG as well as in the cochlear model (values of contrast/base amplitude ratio: 0.01 for the cochlear channels model and -0.05 ± 0.26 for the EEG responses; one-sample t-test: $p = 0.387$). More importantly, there was also no significant difference in condition 2 with 0.50 semitone contrast (0.39 for the cochlear channels model and 0.46 ± 0.50 for the EEG responses; $p = 0.621$), suggesting that the emergence of a response to this fine frequency contrast could be explained by peripheral processing of the contrast. However, this was not the case in condition 3 where the contrast/base ratio predicted by the cochlear channels model was much greater than the actual ratio observed in the EEG (4.20 for the cochlear channels model and 1.73 ± 0.76 for the EEG responses; $p < 0.0001$), suggesting the

involvement of central mechanisms in processing the sequence with larger frequency contrast.

These ratios were also significantly different across the three conditions ($F(20.40, 1.26)= 16.66$, $\eta^2 = 0.60$, $p = 0.001$), with greater values in condition 3 as compared to condition 1 (post hoc t-test, $p = 0.001$) and 2 ($p = 0.002$) and also greater values in condition 2 as compared to condition 1 ($p = 0.020$).

Time course of the EEG responses to the AAAAB pattern. As depicted in Figure 5, the EEG waveforms in response to the different sequences were characterized in all conditions by a positive deflection peaking ~50 ms relative to the onset of each tone, thus at a latency compatible with middle latency auditory potentials (Grimm et al. 2016), followed by a negative deflection at ~100 ms.

In Experiment 1, the point-by-point comparison of the waveforms across conditions revealed three significant clusters extending between 91-170 ms, 216-422 ms and 504-601 ms (Fig. 5). Post hoc tests between conditions 1 and 2 revealed one significant cluster (227-278 ms) characterized by more negative amplitude values. Comparison between conditions 2 and 3 revealed three significant clusters exhibiting more negative amplitude values coinciding with the occurrence of the A tones (91-134 ms, 266-371 ms and 382-428 ms) and one significant cluster with more positive amplitude values (504-594 ms) coinciding with the occurrence of the B tone in the AAAAB pattern. In Experiment 2, the point-by-point comparison of the two conditions revealed two significant clusters between 157-241 ms and 537-606 ms (Fig. 5).

Time course of the EEG response across successive repetitions of the AAAAB pattern. In Experiment 1, the magnitude of the base response was stable over time in conditions 1 and 2, but not in condition 3, where a progressive decrease over time was found (Fig. 6). These

observations were confirmed by the ANOVA showing no significant difference in amplitude across the time points in conditions 1 and 2 ($F(53, 11)=1.105$, $\eta^2 = 0.091$, $p = 0.289$ and $F(53, 11)=1.080$, $\eta^2 = 0.089$, $p = 0.329$ respectively), but a significant decrease in condition 3 ($F(53, 11)=1.632$, $\eta^2 = 0.129$, $p = 0.004$; significant linear trend at $p<0.0001$, -6% slope). A similar profile was observed in Experiment 2. Specifically, a stable amplitude of the base response was observed over time in condition 1 (0 μs inter-aural time contrast: $F(53, 11)= 0.601$, $\eta^2 = 0.051$, $p = 0.988$), but not in condition 2 (600 μs inter-aural time contrast: $F(53, 11)= 1.897$, $\eta^2 = 0.147$, $p = 0.0003$; significant linear trend at $p<0.0001$, -10%).

For the contrast response, the ANOVA comparing the amplitude over time yielded overall similar results than for the base response. In Experiment 1, no significant differences across time points were obtained in condition 1 (0.02 semitone contrast: $F(53, 11)=1.335$, $\eta^2 = 0.108$, $p = 0.062$) and in condition 2 (0.50 semitone contrast: $F(53, 11)=1.398$, $\eta^2 = 0.034$, $p = 0.999$), but a significant decrease over time in amplitude was observed in condition 3 (4 semitones contrast: $F(53, 11)=1.874$, $\eta^2 = 0.145$, $p = 0.0003$; significant linear trend at $p<0.0001$, -8%) (Fig. 6). In Experiment 2, a significant difference across time points was found in condition 1 (0 μs inter-aural time contrast: $F(53, 11)=1.622$, $\eta^2 = 0.128$, $p = 0.005$), but this fluctuation in amplitude did not follow a linear tendency over the sequence ($p = 0.778$). In condition 2, the significant difference across time points (600 μs inter-aural time contrast: $F(53, 11)=2.091$, $\eta^2 = 0.159$, $p < 0.0001$) was characterized by a significant linear decrease over the sequence ($p<0.0001$, -15%).

Visual inspection of the time course of the contrast response revealed a similar profile across the two experiments (Fig. 6, red points). In both experiments, at large contrast, the time course of the contrast response was characterized by a gradual increase at the beginning of the sequence, peaking over the fifth time point (i.e. between ~3 and 10 s). This

transient buildup was observed only for perceptually salient contrasts. Moreover, it was observed only in the contrast responses but not in the base responses, thus suggesting that the two responses might be supported by distinct neural mechanisms.

DISCUSSION

In two experiments, contrast-related responses were successfully identified for different types of auditory contrasts embedded in a fast continuous sound sequence, even when the contrast was very fine. Using a frequency-tagging approach, these contrast responses were identified at the exact frequency at which the auditory contrast was introduced (according to an AAAAB pattern). These responses thus emerged in the EEG frequency spectrum at frequencies distinct from those of the EEG response to the envelope of acoustic energy of the sequence. Comparison with a cochlear channels model revealed that peripheral cochlear channeling could not fully account for the EEG responses to the sequence with large tone contrast, thus suggesting an involvement of additional central mechanisms in processing the contrast. This was also suggested by the fact that contrast responses were obtained for a contrast consisting in a difference in inter-aural timing, as such a contrast cannot be resolved at the level of cochlear channels.

Time course analysis of the EEG responses to the AAAAB patterns revealed that these responses occurred at latencies corresponding to that of middle-latency auditory potentials (~50 ms). Finally, analysis of the evolution of the responses *across* repetitions of the AAAAB patterns revealed a stable amplitude over time, except for the responses to the most perceptually salient contrasts which showed dynamics of buildup and decay across repetitions of the pattern.

Undetectable contrast

No contrast response was identified for the tone contrast of 0.02 semitone. This result corroborates previous evidence that the cochlear channels cannot resolve such a small tone contrast (Moore 2003), as also reflected in the subjective report of all the participants who could not discriminate between A and B tones in this condition. Instead, only the base

response to the envelope of acoustic energy of the sequence was observed in the EEG spectrum. This response displayed a fronto-central distribution as typically found for responses to amplitude-modulated tones, compatible with bilateral activation of the Heschl's gyrus in response to fast sequences of tones (Snyder et al., 2006). Time course analysis within the AAAAB pattern revealed that this base response corresponded to a positive deflection peaking ~50 ms relative to the onset of each tone in the pattern, which could be related to auditory middle latency responses originating from the Heschl's gyrus (see e.g. Gutschalk et al. 2005; Snyder et al. 2006). This positive deflection was followed by a negative peak at ~100 ms, which could correspond to higher-level late-latency cortical activity (see e.g. Cornella et al. 2015). Future research using the same design but coupled with recording methods providing a greater spatial resolution (e.g. intracerebral EEG; see e.g. Jonas et al. 2016) or with analysis methods such as those that have been used to investigate the neural generators of evoked potentials such as the MMN (e.g. dynamic causal modelling; see e.g. Garrido et al. 2008) should address the question of the generators of these components more appropriately.

Interestingly, there was no significant decrease in amplitude of these responses over repetitions of the tones in the 40-s sequence. The fact that this response was not subject to adaptation even after dozens of seconds of repetitive stimulation at relatively fast rate is in discrepancy with the hypothesis of a reduction of the neural response to a stimulus by a similar preceding stimulus according to a mechanism of adaptation (Hartmann and Johnson 1991; Bregman et al. 2000; Gutschalk et al. 2005). Alternatively, the adaptation could have been missed by our analysis, if this is a quick process already appearing after one or two repetitions of the sound in the sequence.

Small detectable contrast

As opposed to the 0.02 semitone condition, the 0.50 semitone condition elicited a small but significant contrast response, in accordance with the subjective report of the participants who could all discriminate between A and B tones in this condition. This result highlights the sensitivity of the frequency-tagging method to capture the neural processing of contrast inserted in fast continuous streams and suggests that the method could be possibly used to probe contrast response at threshold in individuals with specific auditory expertise or impairment. Indeed, the semitone is considered as an experience-induced perceptual boundary in Western listeners, as this frequency interval is usually the least common interval found in most Western speech and music (Zarate et al. 2012).

Comparison with a cochlear channels model of response indicated that the channeling processes occurring at cochlear stage could account for the EEG responses to this fine tone contrast. In other words, the contrast-specific response observed here could be explained by the activity of neuronal populations responding specifically to the frequency band corresponding to B tones, distinct from the tonotopic channel responding to A tones (Micheyl et al. 2007). Interestingly, there was no significant decrease in amplitude of these responses over the 40-s sequence, thus revealing again a relatively stable process with no adaptation, even after dozens of repetitions of the tones at fast rate.

Large contrast

With 4 semitones contrast, the contrast-related response was even sharper, thus increasing in amplitude when increasing the contrast. However, these responses could no longer be explained by the cochlear channels model, thus suggesting an involvement of additional central mechanisms. For instance, these neural responses might comprise an effect of bottom-up attention captured by the AAAAB pattern made more perceptually salient due to the sharper contrast, as opposed to sequences conveying a smoother pattern surface

through finer contrast between tones. The involvement of central mechanisms, possibly occurring already at brainstem level, was also suggested by the results of the second experiment using contrast of inter-aural timing instead of tone contrast. Indeed, the contrast responses observed in this second experiment cannot be explained by peripheral channeling as, in this experiment, the A and B tones activated the same tonotopic channels.

Time course of the responses

Time domain analysis revealed a significant modulation of the response to the AAAAB pattern by the contrast, thus explaining the emergence of contrast responses in the frequency domain analysis. For the fine contrast, these modulations appeared within time windows compatible with late responses triggered by the B tone (i.e. during the presentation of the A tones). However, for larger contrasts, these modulations also appeared in a time window compatible with middle latency responses to the B tone (i.e. during the actual presentation of the B tone). To date, deviance-related modulations of middle-latency auditory evoked potentials have been reported to physical feature changes occurring with a low probability, but have not been reported in sequences following a fast repeating pattern such as the one used here (see e.g. Grimm et al. 2016 for a review).

We also analyzed the evolution of the responses across repetitions of the AAAAB pattern. For large contrasts, this analysis revealed a transient increase of the contrast response, peaking 3-10 s after the onset of the sequence. This transient increase could be attributed to the processing of regularity requiring a few instances of the stimulus to lead to the formation of complex perceptual patterns (see e.g. Barascud et al. 2016). Alternatively, this buildup could be interpreted in light of previous work having observed a progressive increase in amplitude of the event-related potentials over a few presentations of ABA- patterns used to explore the perceptual phenomenon of stream segregation between A and

B tones (Snyder et al. 2006). In this work, the buildup was interpreted as a temporal window in which successive tones were analyzed and integrated within separate streams, thus underlying the buildup of stream segregation (Snyder et al. 2006). This interpretation fits well with the results of the current study showing a similar buildup period of the contrast response but not the base response and, most importantly, only in the conditions that generated a stream segregation effect as reported by all participants (i.e. the large frequency contrast of Experiment 1, and the separate spatial sources condition of Experiment 2). Future research using the same frequency-tagging method but manipulating the perceptual segregation into separate streams independently of the contrast, for example, by exploiting the multistability of sequences inducing either stream segregation or integration (see e.g. Bregman 1990), could address this issue more directly.

This transient increase of the contrast response was followed by a progressive decay over the 40-s sequence, possibly reflecting a mechanism of adaptation due to the repetition of the contrast over time. Such adaptation has already been observed in human and non-human animals, by showing modulations of the sensitivity to the contrast due to the increased predictability of the occurrence of the contrast, leading to the formation of complex perceptual patterns (Sussman et al. 1998; Chait et al. 2008; Yaron et al. 2012; Bendixen 2014; Schröger et al. 2014; Simpson et al. 2014; Barascud et al. 2016). Most importantly, the fact that this progressive decay was observed only for large contrasts but not for fine contrasts in the current study suggests that the perceptual saliency of the contrast is critical to this adaptation, in addition to parameters such as contrast predictability.

Overall, the two experiments demonstrate that this new combination of frequency-tagging with an oddball design can be used to characterize the neural responses specific to contrasts in fast continuous acoustic sequences. Specifically, the approach opens to further research studying contrast-specific responses across a large range of presentation rates to investigate the processing of auditory information that arrives rapidly and sequentially, and also investigating the modulation of this processing by endogenous factors such as attention or exogenous factors such as acoustic spectral properties of the stimuli for instance. Moreover, because it does not require an explicit response from the participant, this method could be particularly valuable to probe these processes in individuals with hearing impairment or neurodevelopmental disorders such as dyslexia showing difficulties in processing brief, rapidly changing, acoustic information (Oxenham 2008; Albouy et al. 2016).

Acknowledgements: S.N. is supported by an Australian Research Council (ARC) DECRA DE160101064 and FRSM 3.4558.12 Convention Grant from the Belgian National Fund for Scientific Research (F.R.S.-FNRS) (to Pr. A. Mouraux).

Disclosures: No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

Albouy P, Cousineau M, Caclin A, Tillmann B, Peretz I. Impaired encoding of rapid pitch information underlies perception and memory deficits in congenital amusia. *Sci Rep* 6:18861 doi: 10.1038/srep18861, 2016.

Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM. Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26(45): 11695-708, 2006.

Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M. Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc Natl Acad Sci U S A* 113(5):E616-25, 2016.

Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* 7:1129-59, 1995.

Bregman A. Auditory scene analysis: The perceptual organization of sound. *MIT Press*, Cambridge, Mass., USA, 1990.

Bendixen A. Predictability effects in auditory scene analysis: a review. *Front Neurosci* 8:60, 2014.

Chait M, Poeppel D, Simon JZ. Auditory temporal edge detection in human auditory cortex. *Brain Res* 1213:78-90, 2008.

Chemin B, Mouraux A, Nozaradan S. Body movement selectively shapes the neural representation of musical rhythm. *Psychol Sc* 25(12):2147-59, 2014.

Cirelli LK, Spinelli C, Nozaradan S, Trainor LJ. Measuring Neural Entrainment to Beat and Meter in Infants: Effects of Music Background. *Front Neurosci* 10:229 doi: 10.3389/fnins.2016.00229, 2016.

Cornella M, Bendixen A, Grimm S, Leung S, Schröger E, Escera C. Spatial auditory regularity encoding and prediction: Human middle-latency and long-latency auditory evoked potentials. *Brain Res* 1626:21-30.

de Heering A, Rossion B. Rapid categorization of natural face images in the infant right hemisphere. *Elife* doi:10.7554/eLife.06564, 2015.

Demany L, Semal C. The slow formation of a pitch percept beyond the ending time of a short tone burst. *Percept Psychophys* 67:1376–1383, 2005.

Demany L, Semal C. In Auditory Perception of Sound Sources Vol. 29 *Springer Handbook of Auditory Research* eds R. R. Fay & A. N. Popper 77–113, 2007.

Escera C, Malmierca MS. The auditory novelty system : an attempt to integrate human and animal research. *Psychophysiol* 51:111-23, 2014.

Freyman RL, and Nelson DA. Frequency discrimination as a function of tonal duration and excitation-pattern slopes in normal and hearing-impaired listeners. *J Acoust Soc Am* 79:1034, 1986.

Garrido MI, Friston KJ, Kiebel SJ, Stephan KE, Baldeweg T, Kilner JM. The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* 42(2):936-44.

Grimm S, Escera C. Auditory deviance detection revisited: evidence for a hierarchical novelty system. *Int J Psychophysiol* 85(1):88-92, 2012.

Grimm S, Escera C, Nelken I. Early indices of deviance detection in humans and animal models. *Biol Psychol* 116:23-7, 2016.

Goswami U. Sensory theories of developmental dyslexia: three challenges for research. *Nat Rev Neurosci* 16:43–54, 2015.

Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48(12):1711-25, 2011.

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ. Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci* 25(22):5382-8, 2005.

Hartmann WM, Johnson D. Stream segregation and peripheral channeling. *Music Perception Journal* 9:155–183, 1991.

Heinrich SP, Mell D, Bach M. Frequency-domain analysis of fast oddball responses to visual stimuli: a feasibility study. *Int J Psychophysiol* 73(3):287-93, 2009.

Jonas J, Jacques C, Liu-Shuang J, Brissart H, Colnat-Coulbois S, Maillard L, Rossion B. A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proc Natl Acad Sci U S A* 113(28):E4088-97 doi: 10.1073/pnas.1522033113, 2016.

Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol* 111:1745-58, 2000.

Kraus N, McGee T, Carrell T, King C, Littman T, Nicol T. Discrimination of speech-like contrasts in the auditory thalamus and cortex. *J Acoust Soc Am* 96(5 Pt 1):2758-68, 1994.

Liu-Shuang J, Norcia AM, Rossion B. An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia* 52:57-72, 2014.

Lochy A, Van Reybroeck M, Rossion B. Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proc Natl Acad Sci U S A* 113(30):8544-8549.

Massaro DW, Idson WL. Backward recognition masking in relative pitch judgments. *Percept Mot Skills* 45:87–97, 1977.

May PJ, Tiitinen H. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiol* 47:66-122, 2010.

Makeig S. Response: event-related brain dynamics - unifying brain electrophysiology. *Trends Neurosci* 25:390, 2002.

Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177-90, 2007.

Micheyl C, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Courtenay Wilson E. The role of auditory cortex in the formation of auditory streams. *Hear Res* 229(1-2):116-31, 2007.

Moore BCJ. An Introduction to the Psychology of Hearing. Fifth Edition, *Academic Press*, 2003.

Mouraux A, Iannetti GD, Colon E, Nozaradan S, Legrain V, Plagkhi L. Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. *J Neurosci* 31:6079-87, 2011.

Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118:2544-90, 2007.

Nelken I. Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol Cybern* 108(5):655-63, 2014.

Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B. The steady-state visual evoked potential in vision research: A review. *J Vis* 15(6):4 doi: 10.1167/15.6.4, 2015.

Nozaradan S, Peretz I, Mouraux A. Steady-state evoked potentials as an index of multisensory temporal binding. *Neuroimage* 60(1):21-8, 2012.

Nozaradan S, Zerouali Y, Peretz I, Mouraux A. Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cereb Cortex* doi: 10.1093/cercor/bht261, 2015.

Nozaradan S, Peretz I, Keller PE. Individual Differences in Rhythmic Cortical Entrainment Correlate with Predictive Behavior in Sensorimotor Synchronization. *Sci Rep* 5;6:20612 doi: 10.1038/srep20612, 2016.

Nozaradan S, Schönwiesner M, Caron-Desrochers L, Lehmann A. Enhanced brainstem and cortical encoding of sound during synchronized movement. *Neuroimage* pii: S1053-8119(16)30322-6 doi: 10.1016/j.neuroimage.2016.07.015, 2016.

Nozaradan S, Mouraux A, Jonas J, Colnat-Coulbois S, Rossion B, Maillard L. Intracerebral evidence of rhythm transform in the human auditory cortex. *Brain Structure and Function*. In Press.

Oxenham AJ. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif* 12(4):316-31, 2008.

Retter TL, Rossion B. Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia* 91:9-28, 2016.

Rossion B. Understanding individual face discrimination by means of fast periodic visual stimulation. *Exp Brain Res* 232(6):1599-621, 2014.

Schröger E, Bendixen A, Denham SL, Mill RW, Bőhm TM, Winkler I. Predictive regularity representations in violation detection and auditory stream segregation: from conceptual to computational models. *Brain Topogr* 27(4):565-77, 2014.

Simpson AJ, Harper NS, Reiss JD, McAlpine D. Selective adaptation to "oddball" sounds by the human auditory system. *J Neurosci* 34(5):1963-9, 2014.

Slaney M. Auditory Toolbox Version 2, Technical Report. *Interval Research Corporation* 1998-010, 1998.

Snyder JS, Alain C, Picton TW. Effects of attention on neuroelectric correlates of auditory stream segregation. *J Cogn Neurosci* 18:1-13, 2006.

Sussman E, Ritter W, Vaughan HG Jr. Predictability of stimulus deviance and the mismatch negativity. *Neuroreport* 9(18):4167-70, 1998.

Tallal P. Improving language and literacy is a matter of time. *Nat Rev Neurosci* 5:721–728, 2004.

Tallal P, Miller S, Fitch RH. Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann N Y Acad Sci* 682:27–47, 1993.

van den Broeke EN, Mouraux A, Groneberg AH, Pfau DB, Treede RD, Klein T. Characterizing pinprick-evoked brain potentials before and after experimentally induced secondary hyperalgesia. *J Neurophysiol* 114(5):2672-81, 2015.

van Noorden L. Temporal Coherence in the Perception of Tone Sequences. Unpublished thesis, *Technical University Eindhoven*, 1975.

Yaron A, Hershenhoren I, Nelken I. Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76:603-15, 2012.

Yost WA. Fundamentals of hearing: An Introduction (4th ed.). *New York: Academic Press*, 2000.

Zarate JM, Ritson CR, Poeppel D. Pitch-interval discrimination and musical expertise: Is the semitone a perceptual boundary? *The Journal of the Acoustical Society of America* 132(2):984-99, 2012.

		base response (8Hz f1-f2)	contrast response (1.6Hz f1-f4)
Experiment 1: frequency contrast	**condition 1: 0.02 semitone**	0.094±0.058 µV p=0.0006	-0.001±0.007 µV p=1
	condition 2: 0.50 semitone	0.106±0.054 µV p<0.0001	0.038±0.044 µV p=0.036
	condition 3: 4 semitone	0.143±0.106 µV p=0.0021	0.223±0.132 µV p=0.0003
Experiment 2: inter-aural time contrast	**condition 1: 0 µs**	0.103±0.055 µV p<0.0001	0.001±0.017 µV p=1
	condition 2: 600 µs	0.130±0.074 µV p<0.0001	0.162±0.075 µV p<0.0001

Table 1. T-tests against zero of the noise-subtracted amplitudes for the base and contrast responses. These values were obtained for each participant (mean ± standard deviation, n=12 for each Experiment) from a pool of fronto-central electrodes and averaged across harmonics within contrast (f1-f2) and base (f1-f4) responses.

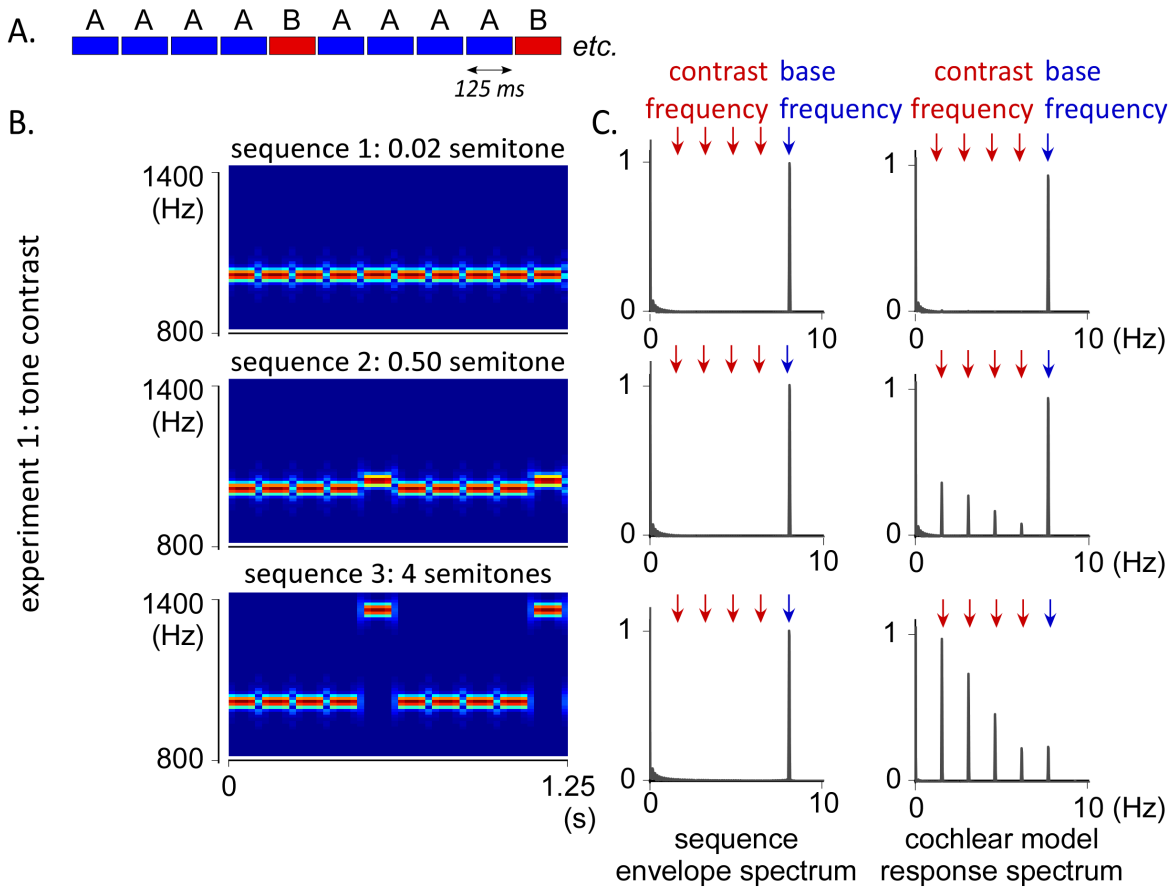

Figure 1. AAAAB pattern and stimuli of Experiment 1 with tone contrast. A. Structure of the AAAAB pattern repeated over the 40-s sequences. B. Time-frequency maps of the AAAAB pattern in Experiment 1. The tone contrast between A and B tones was either 0.02 semitone (undetectable contrast; sequence 1), 0.50 semitone (small detectable contrast; sequence 2) or 4 semitones (large detectable contrast; sequence 3). C. Left column: frequency spectrum of the sound envelope of the three sequences, with a peak at 8 Hz corresponding to the 125-ms repetition rate of the tones (i.e. base frequency). Right column: frequency spectrum of the cochlear response predicted by a cochlear channels model (displayed as normalized relative to the maximum magnitude within each sequence). Note that the cochlear channels model predicts an additional peak at 8 Hz/5 and harmonics, corresponding to the contrast occurrence. Also note that the predicted amplitude of the response at base frequency was much higher than the response at contrast frequency in sequence 2 (small contrast). Conversely, the predicted amplitude of the base response was much smaller than the contrast response in sequence 3 (large contrast).

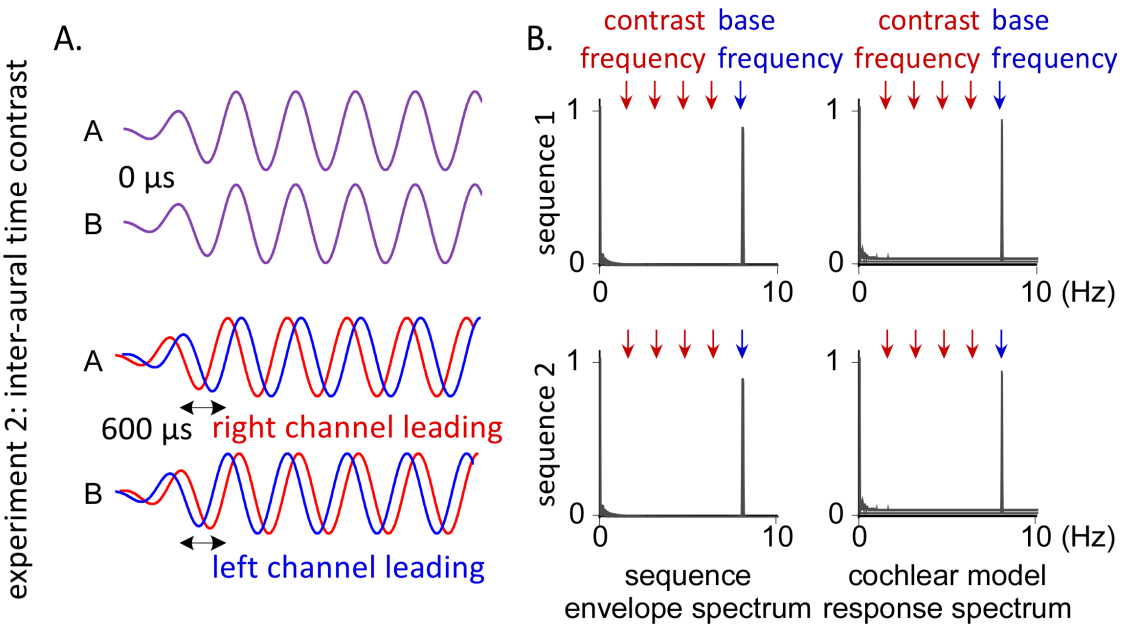

Figure 2. Stimuli of Experiment 2 with inter-aural time contrast. A. In condition 1, there was no inter-aural time contrast between A and B tones (upper graph). In condition 2, the A and B tones were presented with an inter-aural time contrast of 600 µs (bottom graph), inducing an illusion of 180° azimuth contrast between A and B tones and thus an effect of stream segregation due to the perceived separate spatial sources. B. Frequency spectrum of the sound envelope and of the cochlear signals obtained using a cochlear channels model of response (displayed as normalized relative to the maximum magnitude within each sequence). Note that, in both conditions, no peaks emerge at contrast frequencies, because inter-aural time contrasts cannot be processed by cochlear channels.

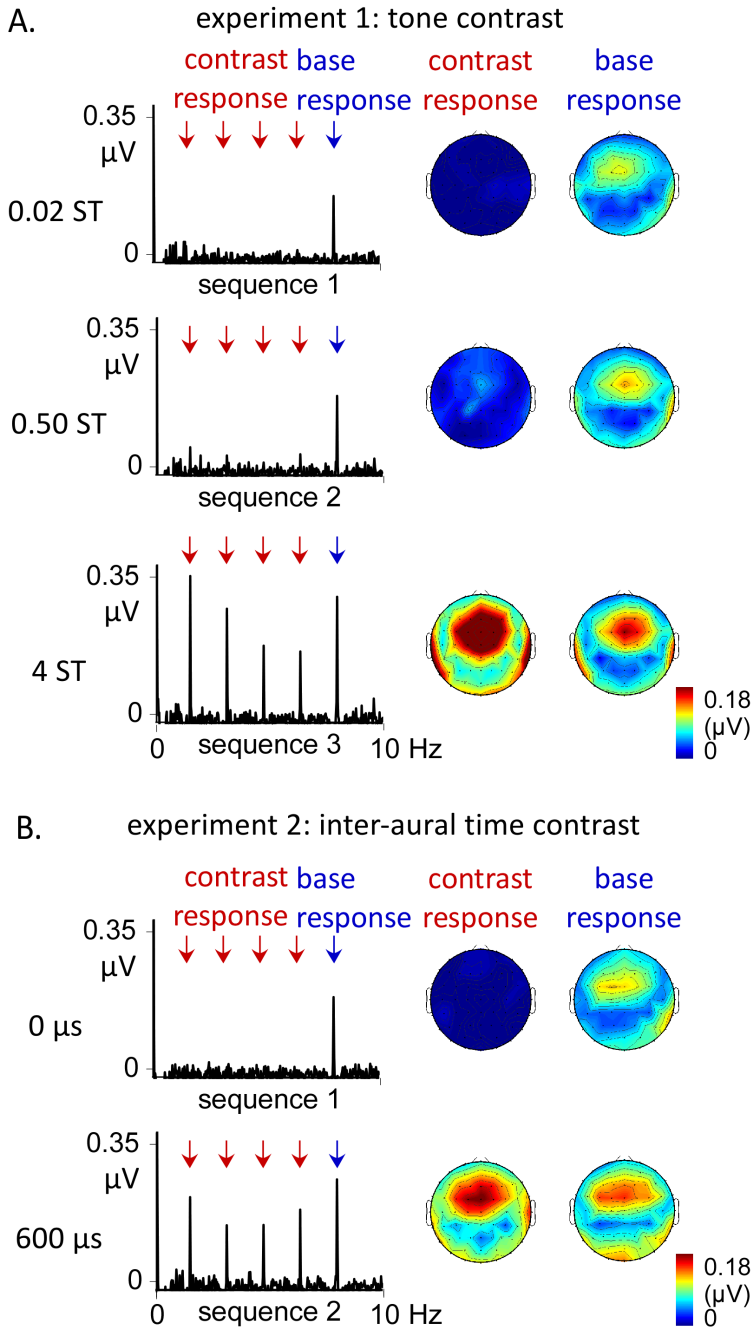

Figure 3. EEG spectra of Experiments 1 and 2 and corresponding topographies. A. In Experiment 1, a base response was observed with similar amplitude across conditions. Most importantly, a contrast response emerged in condition 2, as predicted by the cochlear channels model. This contrast response was markedly increased in condition 3, although it was not associated with a relative reduction of the base response as predicted by the cochlear channels model. B. A similar profile of response was obtained in Experiment 2, with a stable base response across conditions and a contrast response emerging in condition 2. Note that the topographical distribution of these responses was fronto-central and overall similar across conditions and experiments.

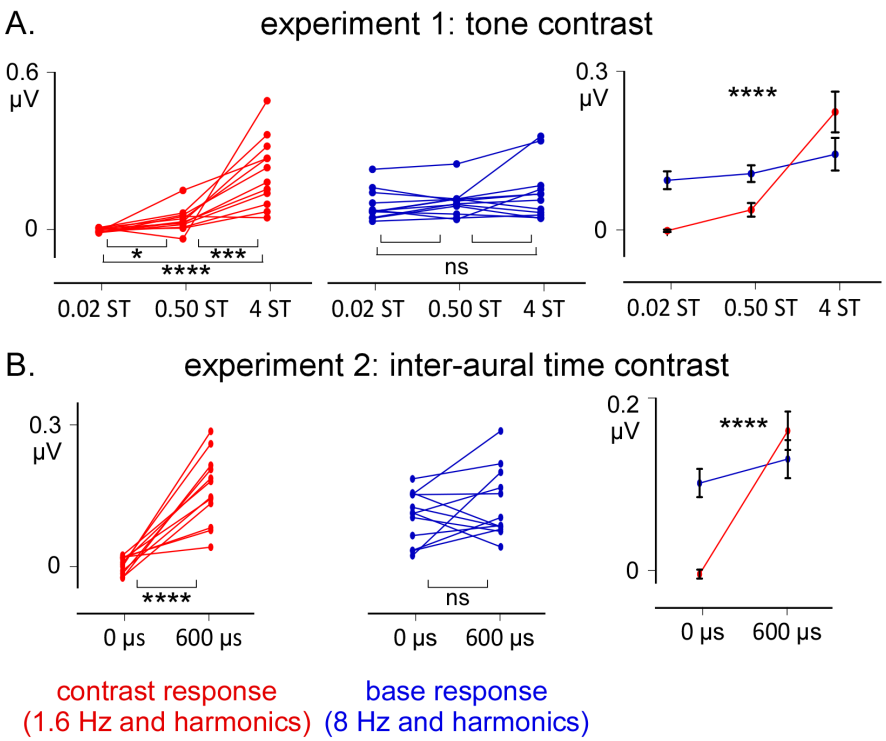

Figure 4. Noise-subtracted amplitudes across participants for Experiments 1 and 2 (fronto-central electrodes). The first two graphs on the left correspond to the amplitude measured in each participant (asterisks = pairwise comparisons). The right graphs correspond to the group-level average ± standard deviation (asterisks = ANOVAs). ns: non significant; *: p≤.05; **: p≤.01; ***: ≤.001; ****: p≤.0001.

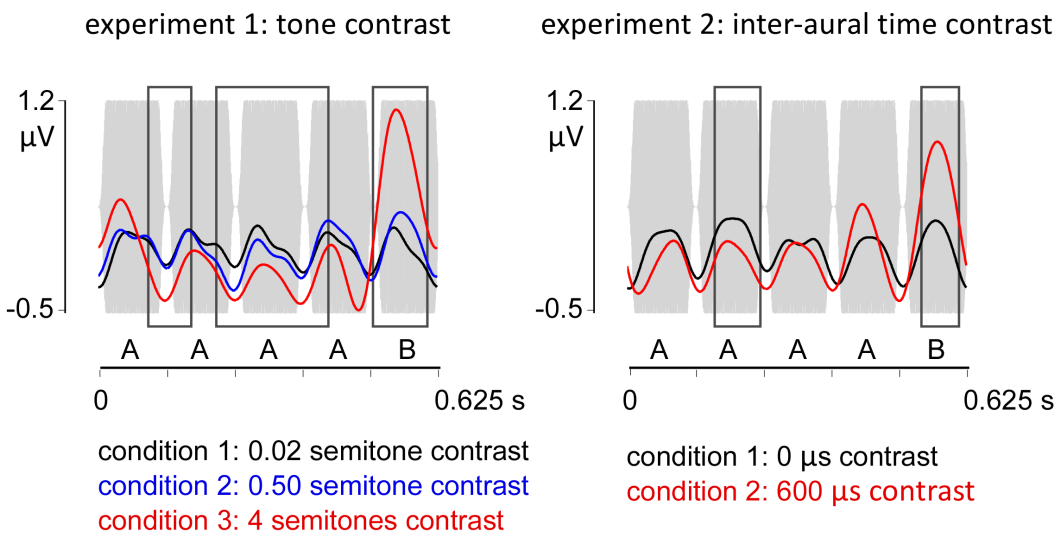

Figure 5. Time course of the EEG responses to the AAAAB pattern. The grey bars represent the pattern envelope. In the two Experiments, responses to each tone were observed as positive deflections peaking at ~50 ms followed by a negative deflection at ~100 ms. A point-by-point comparison (cluster statistics) across conditions revealed significant differences across conditions during and also after the presentation of the B tone (black boxes), thus compatible with middle and late latency auditory potentials triggered by the contrast, respectively.

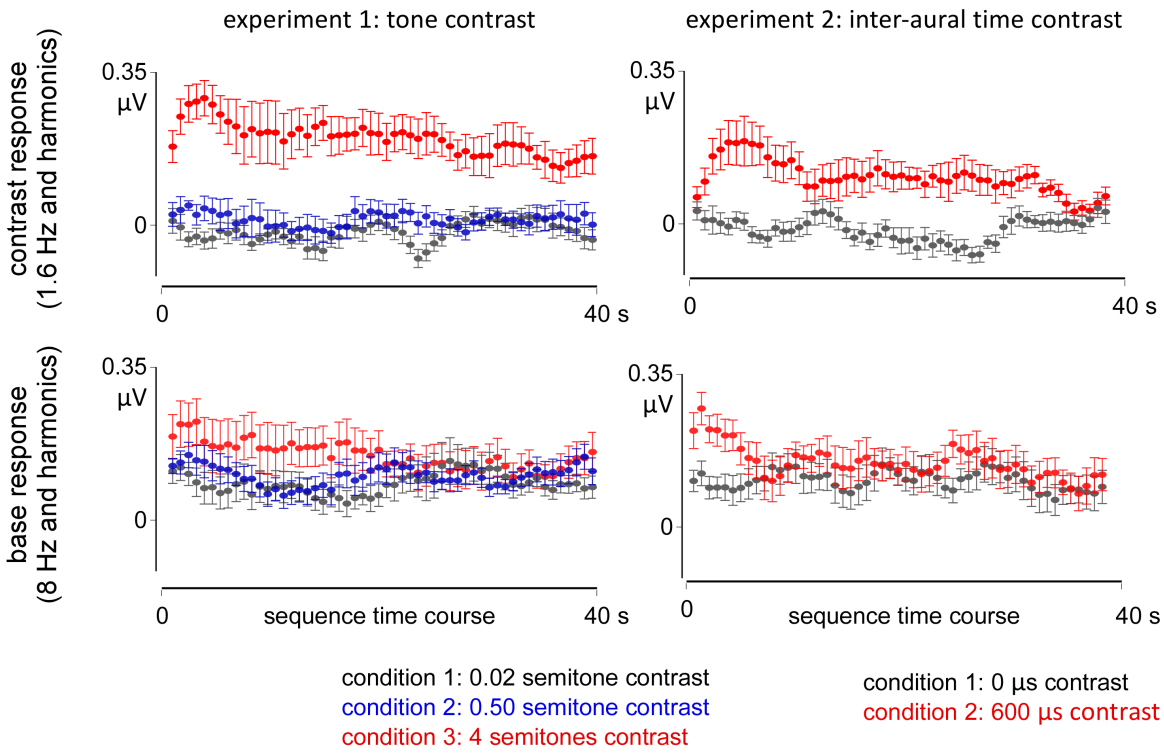

Figure 6. Evolution of the EEG responses across successive AAAAB patterns. The time points were obtained using a sliding FFT (6.25-s time window with 0.625 s steps). For both the contrast and base responses (upper and lower graphs), the amplitude remained stable over time, except in the conditions with large contrast where a transient increase in amplitude was also observed, spanning 3 to 10 s after the onset of the sequence, followed by a significant decrease (in red, upper graphs, Experiments 1 and 2).