

# “Change Deafness” Arising from Inter-feature Masking within a Single Auditory Object

Nicolas Barascud<sup>1</sup>, Timothy D. Griffiths<sup>1,2</sup>, David McAlpine<sup>1</sup>,  
and Maria Chait<sup>1</sup>

## Abstract

■ Our ability to detect prominent changes in complex acoustic scenes depends not only on the ear’s sensitivity but also on the capacity of the brain to process competing incoming information. Here, employing a combination of psychophysics and magnetoencephalography (MEG), we investigate listeners’ sensitivity in situations when two features belonging to the same auditory object change in close succession. The auditory object under investigation is a sequence of tone pips characterized by a regularly repeating frequency pattern. Signals consisted of an initial, regularly alternating sequence of three short (60 msec) pure tone pips (in the form ABCABC...) followed by a long pure tone with a frequency that is either expected based on the on-going regular pattern (“LONG expected”—i.e., “LONG-expected”) or constitutes a pattern violation (“LONG-unexpected”). The change in LONG-expected is manifest as a

change in duration (when the long pure tone exceeds the established duration of a tone pip), whereas the change in LONG-unexpected is manifest as a change in both the frequency pattern and a change in the duration. Our results reveal a form of “change deafness,” in that although changes in both the frequency pattern and the expected duration appear to be processed effectively by the auditory system—cortical signatures of both changes are evident in the MEG data—listeners often fail to detect changes in the frequency pattern when that change is closely followed by a change in duration. By systematically manipulating the properties of the changing features and measuring behavioral and MEG responses, we demonstrate that feature changes within the same auditory object, which occur close together in time, appear to compete for perceptual resources. ■

## INTRODUCTION

Survival often depends on the ability to respond promptly and effectively to new events in the environment. In many cases (e.g., in busy, dynamic surroundings or beyond the field of vision) such events are primarily detected as changes in acoustic input, and the auditory system is commonly thought to possess specialized, highly tuned mechanisms for detecting these changes. Delineating the operational limits of these mechanisms—that is, identifying the events that are easily detectable compared with those that tend to be missed by listeners—is essential to uncovering the neural computations underlying auditory change detection and, more broadly, critical to understanding how the dynamics of the acoustic environment are coded by the brain in the course of scene analysis.

Failure to detect prominent changes in acoustic scenes has been examined in two major contexts: (a) energetic masking—when a new event is physically obscured by other scene components (Moore, 2003; Delgutte, 1990) and (b) informational/perceptual masking—when a new event is clearly resolved by the peripheral auditory system, but the presence of other objects in the scene is

nevertheless sufficient to impair detection (Elhilali, Ma, Micheyl, Oxenham, & Shamma, 2009; Gutschalk, Micheyl, & Oxenham, 2008). This latter case—impaired detection despite a positive signal-to-noise ratio—presumably reflects some limited capacity in information processing. For example, even when resolvability is controlled, it is generally harder to detect a new event in scenes that contain many (as opposed to only a few) concurrent elements, an effect that has been termed “change deafness” (Cervantes Constantino, Pinggera, Paranamana, Kashino, & Chait, 2012; Gregg & Snyder, 2012; Eramudugolla, Irvine, McAnally, Martin, & Mattingley, 2005).

Here, we investigate a different form of change deafness, one that arises when two features of a single auditory object change in close temporal proximity. This manifests as changes in one feature of the sound being perceptually masked by changes in the other. In other words, despite the fact that two features of the same auditory object might be altered, listeners often only become aware of a change in one of these features.

The auditory object under investigation constitutes a rapid sequence of tone pips characterized by a regularly repeating frequency pattern. Frequency modulation is arguably one of the most commonly encountered patterns in natural acoustic scenes, and the perception of many signals, including speech and music, is dependent on the ability to

<sup>1</sup>University College London, <sup>2</sup>Newcastle University Medical School

process such sequences (e.g., Overath et al., 2007). Indeed, a considerable body of evidence now indicates that the auditory cortex is sensitive to the frequency structure of on-going sounds, responding reliably to violations of this pattern even when listeners are not actively engaged in tasks that require them to detect such changes (Sculthorpe, Collin, & Campbell, 2008; Wolff & Schröger, 2001; Alain, Woods, & Ogawa, 1994; Tervaniemi, Maury, & Näätänen, 1994).

In a series of experiments, listeners were presented with tone pip sequences such as those depicted in Figure 1. Signals consisted of an initial, regularly alternating sequence of three short pure tone pips (in the form ABCABC; see Figure 1) followed by a long pure tone with a frequency that is either expected based on the on-going regular pattern, for example, ...ABCABClong(A) (“LONG-expected”; Figure 1A) or constitutes a pattern violation (e.g., ...ABCABClong(B), where the frequency of the long tone is unexpected based on the preceding pattern; “LONG-unexpected”; Figure 1B). Therefore, the change in the LONG-expected pattern is manifest as a change in duration (when the long tone exceeds the established duration of a tone pip), whereas the change in LONG-unexpected is manifest as both a change in the frequency pattern and a change in the duration. The design of this stimulus was motivated by the observation that listeners generally fail to detect changes in the frequency pattern when the change in frequency is accompanied by a change in duration. That is, listeners are often “deaf” to the violation of the frequency pattern in the LONG-unexpected signals. The basic behavioral effect is demonstrated in Experiment 1. Subsequent magnetoencephalography (MEG) and psychophysics experiments were designed to elucidate the properties of this form of “change deafness.”

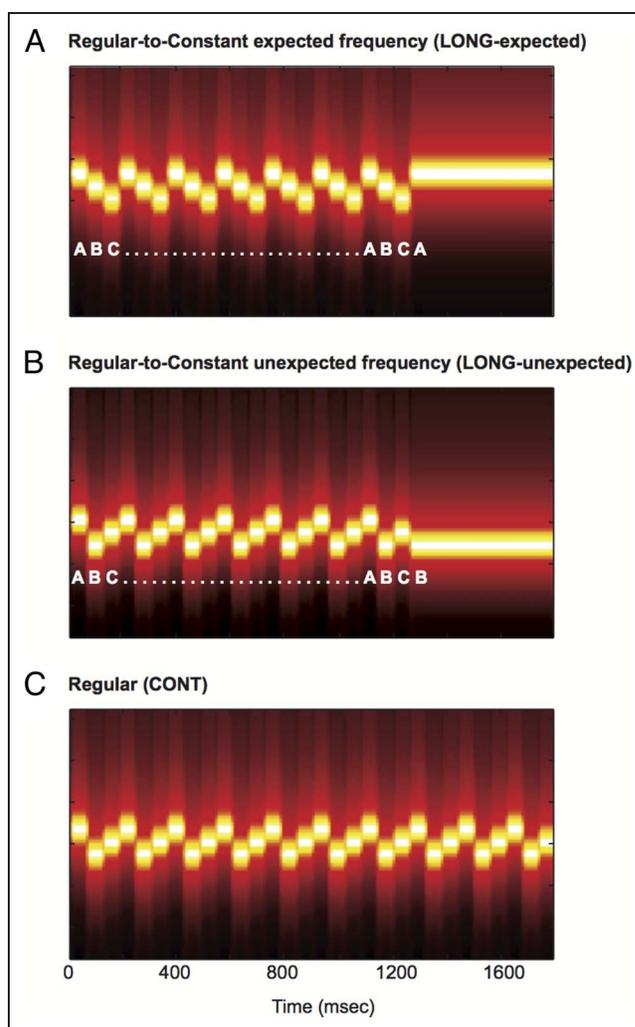
Our data indicate that changes in both the frequency pattern and the expected duration appear to be processed effectively by the auditory system, in that cortical signatures of both changes are evident in the MEG signal. Specifically, activity in the auditory cortex occurring approximately 200 msec after the transition from a regular pattern of ABC tone pips to a long duration tone clearly distinguishes -unexpected from -expected signals. However, the two feature changes appear to compete for perceptual resources, resulting in the suppression of information concerning the change in frequency by information concerning the change in duration, such that it often does not (despite an active effort by the listener) reach conscious awareness.

## EXPERIMENT 1

### Methods

#### Participants

Ten paid participants (4 women; mean age = 28.9 years) participated in the experiment. None were musically trained. All reported normal hearing and no history of neurological disorder. Experimental procedures were approved by the research ethics committee of University



**Figure 1.** Examples of the stimuli in Experiments 1 and 2. (A) The LONG-expected (LONG expected) stimulus consisted of an initial regular pattern, followed by a long pure tone with a frequency that is expected based on the on-going regular frequency pattern. The transition in this case is detectable only from one pip duration after the nominal transition time. (B) The LONG-unexpected (LONG unexpected) stimulus consisted of an initial regular pattern, followed by a long pure tone with a frequency that violated the on-going pattern. Specifically, it was always the frequency that followed the expected frequency in the pattern. (C) The control (no change condition) consisted of a continuous (CONT) regular pattern of three tones A-B-C (this stimulus was used in Experiment 2 but not in Experiment 1). The plots represent “auditory” spectrograms, generated with a filterbank of 1/ERB wide channels (equivalent rectangular bandwidth; Moore & Glasberg, 1983) equally spaced on a scale of ERB rate. Channels are smoothed to obtain a temporal resolution similar to the equivalent rectangular duration (Plack & Moore, 1990).

College London, and written informed consent was obtained from each participant.

#### Stimuli

Stimuli consisted of a pure tone modulated in frequency according to two patterns (LONG-expected and LONG-

unexpected; see Figure 1). Tone pip durations of 200, 100, and 60 msec were employed in different blocks. All tonal transitions were ramped on and off with 15-msec, raised cosine ramps. Frequencies were drawn randomly from a frequency pool of 20 values equally spaced on a logarithmic scale (12% steps) between 222 and 2000 Hz. Both stimulus types consisted of an initial, regularly alternating sequence of three tone pips (A, B, and C; 0-msec intertone interval; duration of 21–24 pips randomly determined for each trial). The frequencies of A, B, and C were drawn randomly (anew for each trial) from the above frequency pool as three consecutive steps, and their order was permuted before assignment to A, B, and C. The spacing between tones (two semitones) assured that frequencies were easily distinguishable and yet close enough so that the sounds were perceived as an integrated sequence of three tones and did not stream (Carlyon, 2004; Bregman, 1990). In the LONG-expected stimulus (Figure 1B), this initial, regular pattern was followed by a 540-msec pure tone with a frequency that was expected in the sequence, that is, based on the ongoing regular pattern (e.g., ...ABCABClong(A); as in Figure 1A). Because the sound frequency after the transition is to a long-duration tone that is consistent with the established pattern, the change is only detectable following a time corresponding to the duration of one tone pip in the pretransition sequence. In the LONG-unexpected stimulus (Figure 1B), the posttransition segment consisted of a pure tone with a frequency that violates the on-going pattern. Specifically, it was always the frequency that followed the expected frequency in the pattern (e.g., ABCABClong(B); as in Figure 1B). Importantly, in both cases, the transition was to a frequency that appeared equally often in the preceding regular pattern, such that the detection of pattern violation could not be based on detecting a novel tone.

In each trial, listeners were presented with a LONG-expected or LONG-unexpected signal and were required to indicate (by pressing one of two keyboard keys—two alternative forced-choice) whether the final (long) tone was expected based on the on-going regular sequence (LONG-expected) or violated the regularity pattern (LONG-unexpected). Visual feedback was provided after each trial.

The stimuli (in this and subsequent experiments) were rendered off-line and stored as 16-bit WAV files at a sampling frequency of 44.1 kHz. The stimulus set was generated anew for each participant.

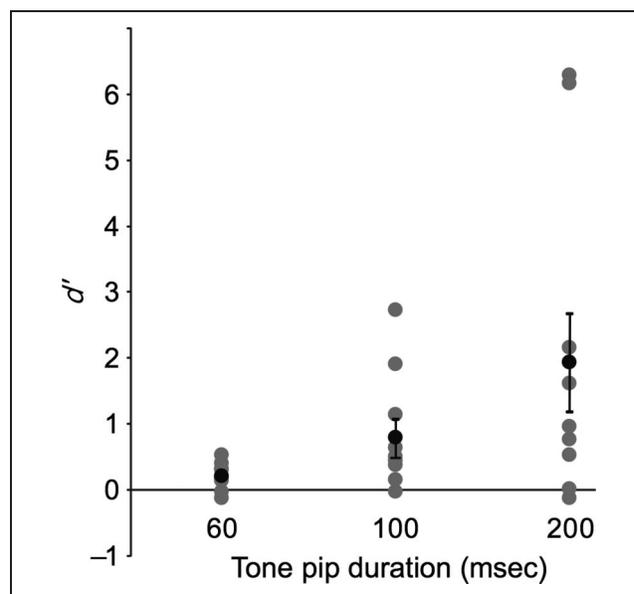
### Procedure

The experimental session began with a practice block, where listeners trained to perform the task with tone pips of 500-msec duration. Practice continued until each listener reached ceiling performance on that condition. In the main experiment, blocks were arranged by decreasing pip duration (200 msec, then 100 msec, then

60 msec) so as to provide ongoing practice, building to the most difficult condition (60-msec pip duration). It was made clear to the listeners that the strategy they employed with the longer tones might not necessarily work for shorter durations (e.g., Warren & Byrnes, 1975), and they were encouraged to use the feedback to evaluate continuously their performance and to refine their strategies for detecting the changes. Each block contained 100 stimuli, and two consecutive blocks were presented for each tone pip duration. Participants were permitted a short rest between blocks.

### Results

Identification performance was assessed for stimuli with tone pip durations of 200, 100, and 60 msec. Figure 2 plots behavioral performance data expressed as  $d'$  sensitivity scores. Plotted are individual participant performance (gray circles) and the mean across participants (black circles). The data demonstrate that, although some listeners could perform the task reasonably well for tone pip durations of 200 and 100 msec, all  $d'$  scores for the 60-msec condition were very low. Notably, even the very best participants (those who obtained exceptionally high  $d'$  scores in the 200- and 100-msec conditions) effectively could not differentiate LONG-unexpected and LONG-expected stimuli (see also results for the same condition in Experiment 3, below).



**Figure 2.** Results of Experiment 1. Participants ( $n = 10$ ) were essentially unable to discriminate between LONG-expected and LONG-unexpected stimuli when pip durations were set to 60 msec. Error bars are 1 SE. The error bars in the 60-msec condition are smaller than the black dot representing the mean and therefore invisible in the figure.

## EXPERIMENT 2

The purpose of this MEG experiment was to investigate the source of the behavioral difficulties exhibited by listeners in Experiment 1. Is the auditory system simply unable to detect violations in frequency patterns, or does reduced performance arise from an “information bottleneck” that limits behavioral performance subsequent to detection of the different sensory events? This issue was explored by measuring MEG responses to the same acoustic signals as in Experiment 1 (tone pip duration = 60 msec; i.e., associated with very poor behavioral performance), while naive listeners performed a decoy task (unrelated to the frequency patterns). In this way, predominantly bottom-up responses to the various transitions are assessed, unaffected by the focus of attention or conscious effort. If the auditory system is sensitive to the violation of the frequency pattern, we expected to observe MEG responses to the change in the LONG-unexpected stimuli as soon as the posttransition frequency is identified. Otherwise, responses to LONG-unexpected and LONG-expected should be identical.

### Methods

#### *Participants*

Sixteen paid participants (nine women; mean age = 28 years) participated in the experiment. All but one were right-handed (Oldfield, 1971). Five of the participants also participated in Experiment 1, above, which was conducted following the MEG study.

#### *Stimuli*

The stimulus set included the LONG-expected and LONG-unexpected conditions identical to those in Experiment 1, above, as well as their no-change control—CONT (Figure 1C). Tone pip duration was 60 msec (corresponding to the shortest duration used in Experiment 1, above). One hundred signals were generated for each of the LONG-expected and LONG-unexpected stimulus conditions, and 200 signals were generated for the regular (CONT) pattern. In this way, the probability of stimulus change was maintained at 0.5, and the occurrence of change within any specific stimulus was unpredictable.

The stimulus set also included a proportion (16%) of 200-msec wide-band noise bursts. These signals were interspersed between the tonal stimuli, and participants were instructed to respond to them as fast as possible by pressing a button. The stimuli were presented to the listeners in a random order with an ISI randomized between 700 and 2000 msec.

#### *Procedure*

Experimental sessions consisted of two phases: recording began with a preliminary functional source-localizer

recording, followed by the main experiment (the main experiment also included another block, with different stimuli, not discussed here). In the functional source-localizer recording, participants listened to 200 repetitions of a 1-kHz, 50-msec sinusoidal tone (ISI randomized between 750 and 1550 msec). These responses were used to verify that the participant was positioned properly in the machine, to verify that signals from auditory cortex showed a satisfactory signal-to-noise ratio, and to determine which MEG channels best reflected activity within auditory cortex. In the main experiment (about 30 min in duration), participants listened to stimuli while performing the noise burst detection task as described above. All listeners were completely naive as to the different stimulus conditions. They were instructed to respond by pressing a button, held in the right hand, as soon as a noise burst appeared. The instructions encouraged speed and accuracy. When questioned after the experiment, none of the participants reported noticing the difference between LONG-expected and LONG-unexpected signals. The presentation was divided into runs of about 5 min (five runs overall). Between runs, participants were permitted a short rest but were required to remain still.

#### *Neuromagnetic Recording and Data Analysis*

Magnetic signals were recorded using a CTF-275 MEG system (axial gradiometers, 274 channels, 30 reference channels; VSM MedTech, Canada). Data were acquired continuously with a sampling rate of 300 Hz and a 100-Hz hardware low pass filter. There was no additional off-line filtering.

Functional localizer data were divided into 700-msec epochs, including 200-msec pre-onset and baseline-corrected to the pre-onset interval. The M100 onset response (Roberts, Ferrari, Stufflebeam, & Poeppel, 2000) was identified for each participant as a source/sink pair in the magnetic field contour plots distributed over the temporal region of each hemisphere. The M100 current source is generally robustly localized to the upper banks of the superior temporal gyrus in both hemispheres (Lütkenhöner & Steinsträter, 1998). For each participant, the 40 strongest channels at the peak of the M100 (20 in each hemisphere) were considered to best reflect activity in the auditory cortex, and these were selected for the root mean square (RMS) analysis below. This procedure serves the dual purpose of enhancing the auditory response components over other response components and compensating for any channel misalignment between participants.

For the main experiment data, 1500-msec epochs (from 500 msec before the transition to 1000 msec after the transition) were created for each of the stimulus conditions (about 100 epochs per condition), averaged, and baseline-corrected to the pretransition interval from 300 msec before the transition. This was done to avoid spurious differences between conditions, introduced by

any drifting in the baseline level of activity (i.e., DC shifts). In each hemisphere, the RMS of the field strength across the 20 channels, selected in the functional source-localizer run, was calculated for each sample point.

The time course of the RMS, reflecting the instantaneous amplitude of neural responses, is employed as a measure of the dynamics of brain response. The congruity of the time course of activation across participants was evaluated using the bootstrap method (Efron & Tibshirani, 1993; 1000 iterations, balanced), based on the individual RMS time series. For illustration purposes, group-RMS (RMS of individual participant RMSs) is plotted, but statistical analysis was always performed across participants, independently for each hemisphere.

To compare the activation between conditions (e.g., Figures 3 and 6), a repeated-measures analysis was used in which, for each participant, the squared RMS value of one condition was subtracted from the squared RMS value of the other condition, and time series of individual differences were subjected to bootstrap analysis (100 iterations, balanced; Efron & Tibshirani, 1993). At each time point, the proportion of iterations below the zero line was counted (see bottom panels in Figures 3 and 5). If the proportion was less than 0.01% or more than 99.99% for 10 adjacent samples (15 msec), the difference was judged to be significant.

Equivalent Current Dipoles (ECDs) were fitted to the single time point corresponding to the maximum of the LONG-expected, SHORT-expected, and CONT-unexpected RMS responses using a variational Bayes ECD algorithm (VB-ECD; Kiebel, Daunizeau, Phillips, & Friston, 2008) based on all available channels (274), in two steps. Initially, two independent ECDs were specified for the whole brain with uninformative prior and moment expectations. In nearly all cases, a clear best ECD, reproducible across 100 repetitions of the VB-ECD procedure, was evident in each hemisphere. To improve further the quality of the fit, the VB-ECD procedure was reiterated using the mean dipole locations as “soft” Bayesian priors (Kiebel et al., 2008). Most of the dipoles obtained that way explained over 80% of the sensor space variance.

## Results

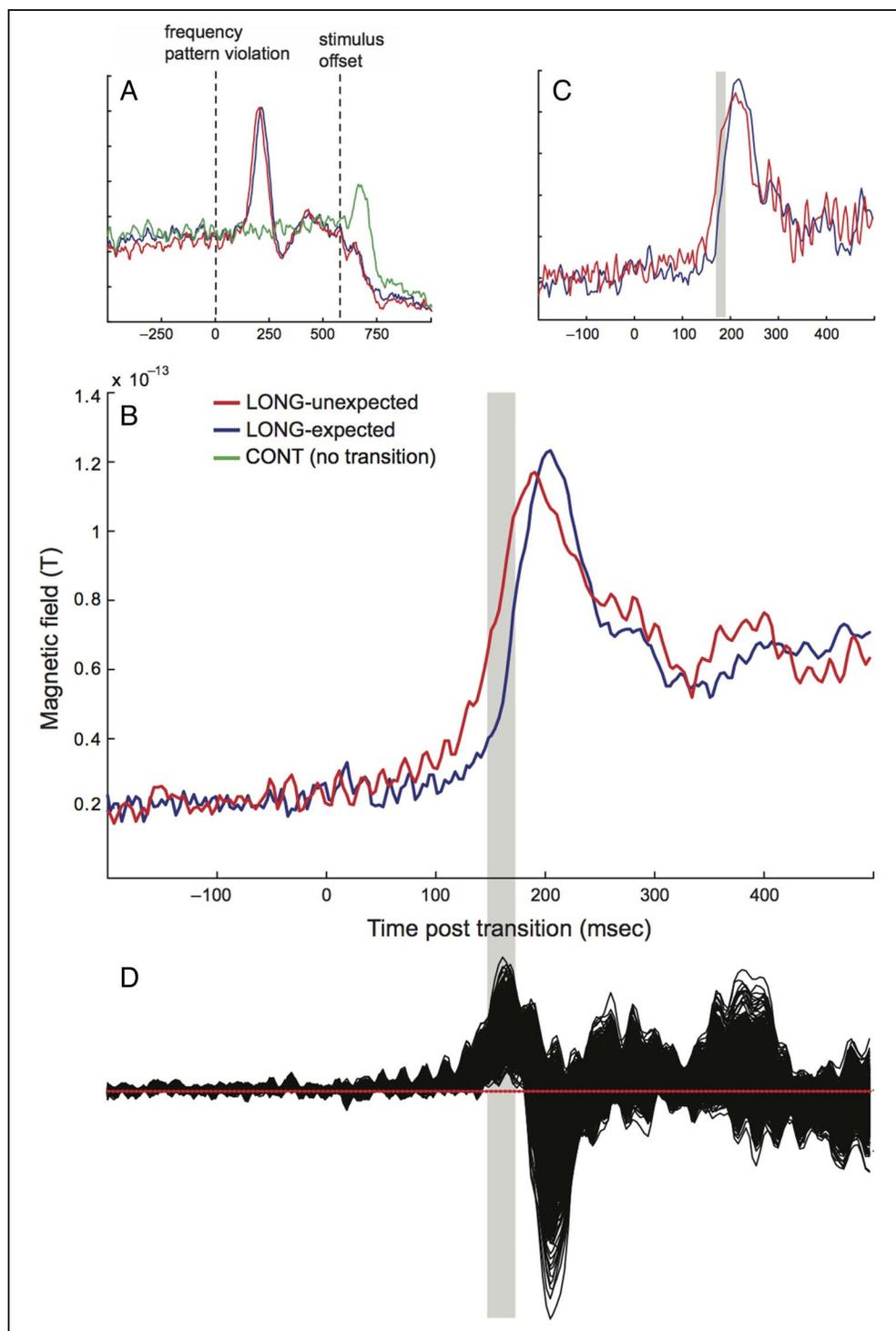
Figure 3A plots the group-RMS (RMS of individual participant RMSs) of auditory-evoked responses to the LONG-expected and LONG-unexpected stimuli. The general response pattern is similar to that observed in previous experiments, which involved the introduction of a pattern change in a continuous sound sequence (e.g., Chait, Poeppel, & Simon, 2008; Chait, Poeppel, de Cheveigné, & Simon, 2007; Jones, 2002; Vaz Pato & Jones, 1999). The auditory-cortical response to the transition from a regularly alternating sequence of tone pips to a constant tone is manifest as an increase in the strength of the magnetic field, peaking some 200 msec following the

transition. The response to the LONG-expected stimuli peaks at 204 msec following the transition, consistent with Chait et al. (2007) in which MEG responses to transitions between a sequence of random frequency tone pips and a long, constant frequency tone (random-to-constant) were assessed. For an ideal observer, this transition is detectable after the duration of exactly one tone pip (i.e., when the tone continues beyond the time when a new tone pip is expected to occur). Indeed, using stimuli with different pip durations (15, 30, 60, and 120 msec), these authors demonstrated that the latency of the first auditory cortical responses to the transition reflected very precise adjustment to this stimulus statistic, such that the latency could be explained by a constant “processing time” of 150 msec + 1 pip duration. This adjustment occurred even when the parameter was not directly behaviorally relevant (i.e., when participants were performing a task that did not require the processing of transitions in the stimuli or adjustment to the duration of the tone pip).

Transition responses to LONG-expected and LONG-unexpected exhibit a very similar profile (including the shape and width of the transition response), which may be taken to indicate that the measured responses do not reflect the detection of changes in frequency or duration but, rather, reflect a subsequent nonspecific change response (see also Oceák, Winkler, & Sussman, 2008; Sussman, Winkler, Ritter, Alho, & Näätänen, 1999; Czigler & Winkler, 1996). However, the two peaks differ in their latency (peak latency of 204 and 191 msec, respectively). This is highlighted in Figure 3B, which magnifies the transition responses. The shaded area indicates the temporal interval over which a repeated-measures bootstrap procedure (see Methods) revealed a significant difference between responses to the two stimuli. Significant differences were observed between the two conditions in the right hemisphere (RH) only, between 140 and 180 msec posttransition. This analysis was conducted on the entire stimulus epoch from 500 msec before the transition to 1000 msec after the transition, and the shaded interval around the transition was the only interval identified in this way across both hemispheres. This pattern demonstrates that the horizontal displacement of the two curves is significant, indicating that the neural responses to LONG-unexpected stimuli preceded neural responses to LONG-expected stimuli by about 10 msec. Source analysis (see “Methods”) both at the response peak and at the center of the time interval, which showed significant differences between the two conditions, revealed no difference between underlying sources.

The difference in latency between LONG-expected and LONG-unexpected conditions is small but robust, as indicated by the fact that all bootstrap iterations are located above the zero line (Figure 3D, bottom). Figure 3C plots the data based on only those five participants who also participated in the behavioral experiment (Experiment 1) and confirms the existence of a robust

**Figure 3.** MEG data (RH) from Experiment 2. (A) Responses to the three stimulus conditions (LONG-unexpected, LONG-expected, CONT) during the full epoch duration from 500 msec before the transition to 1000 msec posttransition. Please note that, for better visualization of the sustained response pretransition and the return to baseline after stimulus offset, these data are baseline-corrected relative to the pre-onset interval. However, the statistical analysis (and the figures in B, C, D) are based on data that are baseline-corrected based on a 300-msec pretransition interval. (B) A closer look at LONG-unexpected and LONG-expected responses around the transition time. Gray shading indicates the temporal intervals where a repeated measures Bootstrap analysis indicated significant differences between the two conditions. (C) Same as B, above, except that the figure is based on data from the five participants who also participated in Experiment 1. Therefore, despite the fact that those participants were unable to distinguish LONG-unexpected and LONG-expected signals, their brain responses demonstrate clearly different response signatures to the two conditions. (D) To demonstrate the robustness of the statistical effect, we show the results of the bootstrap repeated-measures procedure. The black curves are mean differences between conditions ( $x$  axis: time;  $y$  axis: difference (squared) between the -unexpected and -expected stimuli) computed over 1000 bootstrap resamplings; their scatter provides a measure of variability across participants. At each time point, the difference is deemed to be significant if the proportion of iterations above/below the zero line is less than 0.01 for 10 adjacent samples (15 msec).



difference between cortical responses to LONG-expected and LONG-unexpected, despite listeners' inability to distinguish those signals behaviorally. Together, these data suggest that the auditory system reliably acquires the regularity of frequency patterns and rapidly detects violations of this regularity even

when there is no relevant task. Information concerning the frequency change, reflected as an earlier response to LONG-unexpected, is clearly present at a relatively late stage of auditory processing. Why, then are listeners unable to access this information when making behavioral decisions?

### EXPERIMENT 3

Experiments 1 and 2 demonstrate that auditory cortex clearly detects the frequency pattern violation in LONG-unexpected signals but listeners are unable to accomplish this behaviorally. We hypothesized that this “change deafness” results from perceptual suppression by the subsequently occurring change in duration. Namely, despite the fact that information about the frequency change is reliably extracted in the ascending pathways up to auditory cortex, it fails to reach conscious perception because listeners’ attention is captured by the duration change event, which occurs 60 msec later.

To investigate this hypothesis, variants of the basic “LONG” signal employed in Experiment 1 were generated, in which such “informational masking” was explicitly manipulated. All conditions contained a frequency pattern change, as in Experiments 1 and 2, but the properties of the “masking” event, occurring 60 msec subsequent to the frequency change, were varied to produce a range of behavioral effects. To enable such a comparison, it was important that the pattern up to the point of the frequency violation is identical, with the difference between conditions emerging at the same time after that event. Figure 4 illustrates the three stimulus conditions used. The SHORT condition is identical to LONG, except that the last tone in the sequence does not change in duration. As such, the SHORT stimulus also contains two violations (1) and frequency (2) offset. However, based on previous findings indicating that onsets are more perceptually dominant than offsets (Cervantes Constantino et al., 2012; Cole & Kuhn, 2010; Phillips & Hall, 2001), we hypothesized that the offset event will be less effective at masking the frequency violation, therefore resulting in better performance in detecting the frequency change. The CONT (continuous) stimulus consists of an on-going pattern of ABC triplets. The CONT-unexpected condition is generated by skipping a tone in the repeating pattern and then continuing with the same sequence presented before the frequency violation (...ABCABCBCABC...). This sequence, therefore, contains only a violation of the spectral regularity. Performance (i.e., detection of frequency pattern violation) is therefore expected to be highest in CONT, intermediate for SHORT, and lowest for LONG. Note that the hypothesis that performance on SHORT should be better than LONG is, in some sense, counterintuitive because participants have less time in the SHORT condition than in the LONG condition to perceptually resolve the frequency of the last tone in the sequence.

### Methods

#### Participants

Twenty-two paid participants (nine women; mean age = 25.0 years) participated in the experiment. Four participants were musically trained.

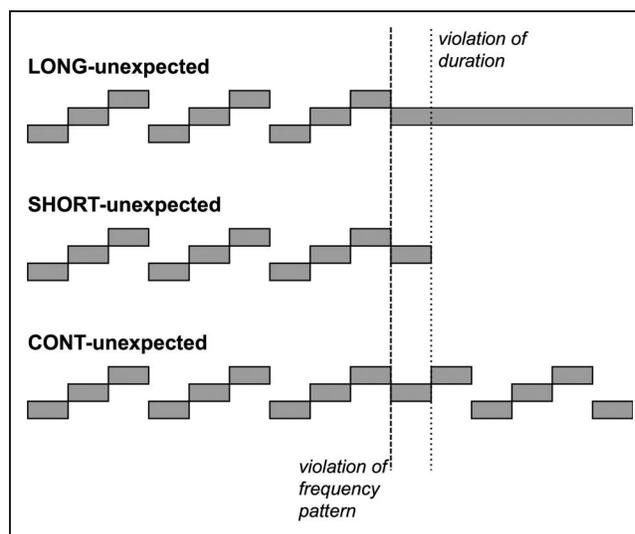
### Stimuli

Three stimulus conditions were tested (Figure 4): (1) LONG stimuli—identical to the stimuli used in Experiment 1; (2) SHORT stimuli—created in the same fashion as LONG signals except that, rather than a long tone, the tone at the transition was of the same duration as the tone pips in the pretransition sequence; (3) CONT (continuous) stimuli—a continuous triplet pattern throughout the duration of the stimulus uninterrupted by a transition to a long tone. All signals were presented in two patterns, -expected and -unexpected. In CONT-unexpected, the violation was introduced by effectively “skipping” a tone anywhere between the 22nd and 25th tone pip (...ABCBCABC...; see Figure 4).

The stimuli were blocked by condition and tone pip duration (100 or 60 msec). On each trial (single-interval forced-choice) participants were required to determine whether the last tone was expected or unexpected based on the preceding pattern by pressing one of two keyboard buttons. The task in the CONT block involved determining whether a tone was skipped in the pattern. Feedback was always provided.

### Procedure

As in Experiment 1, initial training was with tone pips of 500-msec duration. Participants then completed five blocks (100 stimuli per block) for each condition (SHORT, LONG, CONT). In the first two blocks for each condition, tone pip



**Figure 4.** Schematic representation of the stimuli used in Experiments 3 and 4. Three stimulus conditions were tested: (1) LONG stimuli—as those in Figure 1, above; (2) SHORT stimuli—created in the same way as the LONG signals except that the tone at the transition was of the same duration as the tone pips in the pre-transition sequences; and (3) CONT (continuous) stimuli—contained a triplet pattern throughout the duration of the stimulus. The figure presents the -unexpected version of each stimulus, where the tone at the transition violates the established frequency pattern.

duration was 100 msec, and this was reduced to 60 msec in the subsequent three blocks. The order of presentation of the three conditions was counterbalanced across listeners. It quickly became clear that all tested participants performed at ceiling in the CONT condition. This condition was therefore not included in the stimulus set for the remaining participants in Experiment 3.

## Results

Performance on the CONT condition was at ceiling. All listeners performed the task perfectly and required little practice. Thus, the removal of one of the tone pips from the sequence is clearly perceived as an irregularity in the on-going pattern. In contrast, performance on SHORT and LONG conditions was significantly poorer and varied considerably across participants. For pip durations of 100 msec, performance indicated a reasonably high ability to differentiate between -expected and -unexpected signals for both LONG and SHORT conditions (mean  $d'$  of 2.04 and 2.35, respectively). However, performance decreased significantly when pip duration was shortened to 60 msec (mean  $d'$  of 0.68 and 0.914, respectively). Figure 5A compares individual participant performance on the LONG condition, for tone pip durations of 100 msec (abscissa) and 60 msec (ordinate). The four musicians in the participant pool are specified by the dashed circles. As a group, participants in this experiment performed better than those in the identical condition in Experiment 1, largely the result of a few highly performing individuals. It is evident that all participants able to perform the task above chance level experienced a drop in performance when pip duration was shortened from 100 to 60 msec (for the better performers,  $d'$  decreased by more than 50%).

Figure 5B plots the performance of individual participants on the LONG (abscissa) and SHORT (ordinate) conditions. With no difference between conditions, data points are expected to cluster around unity (dashed line). Instead, it is clear that, for both pip duration conditions, the majority of the data points are located above this line, suggesting better performance in the SHORT, compared with the LONG, condition. This observation was confirmed statistically using a repeated-measures ANOVA on  $d'$  data with Condition (LONG/SHORT) and Pip Duration as factors. This analysis yielded significant main effects of Condition [ $F(1, 21) = 20.37, p < .001$ ] and Pip Duration [ $F(1, 21) = 10.08, p = .005$ ] with no interaction. As in Experiment 1, it is clear that many participants had difficulty with the task in both LONG and SHORT conditions, but some performed well, and the majority (17/22) showed better performance for the SHORT condition. Thus, while the  $d'$  difference was small and, on average, most participants found the task difficult, participants systematically showed better performance in the SHORT condition, consistent with our hypothesis.

The results of this experiment demonstrate that two independent factors affect “change deafness” for the frequency pattern: (1) The temporal spacing between the violation of the frequency pattern and the masking event—listeners demonstrate a substantial reduction in performance when pip duration changes from 100 to 60 msec. (2) The identity of the masking event—transitions to a long pure tone (i.e., associated with a new pattern of fluctuation in the auditory object) captured attention (and consequently reduced performance) significantly more strongly than transitions to silence (i.e., associated with offset of an object). Specifically, we hypothesize this effect is related to the perceptual salience of the masking event.

## EXPERIMENT 4

The purpose of Experiment 4 was to compare “bottom-up” cortical responses (i.e., in the absence of directed attention to the transitions) to changes in CONT, SHORT, and LONG signals. Naive listeners heard CONT and SHORT stimuli (in separate blocks) while performing an unrelated task. None of these participants had participated in any of the behavioral experiments, and none were musically trained. Because of session time constraints, it was impossible to include the LONG condition in this assessment, and responses were therefore compared with data for the LONG condition in Experiment 2.

### Methods

#### Participants

Thirteen paid participants (six women; mean age = 24.8 years) took part in the MEG experiment. Two additional participants were excluded from analysis because of excessive movement artifacts.

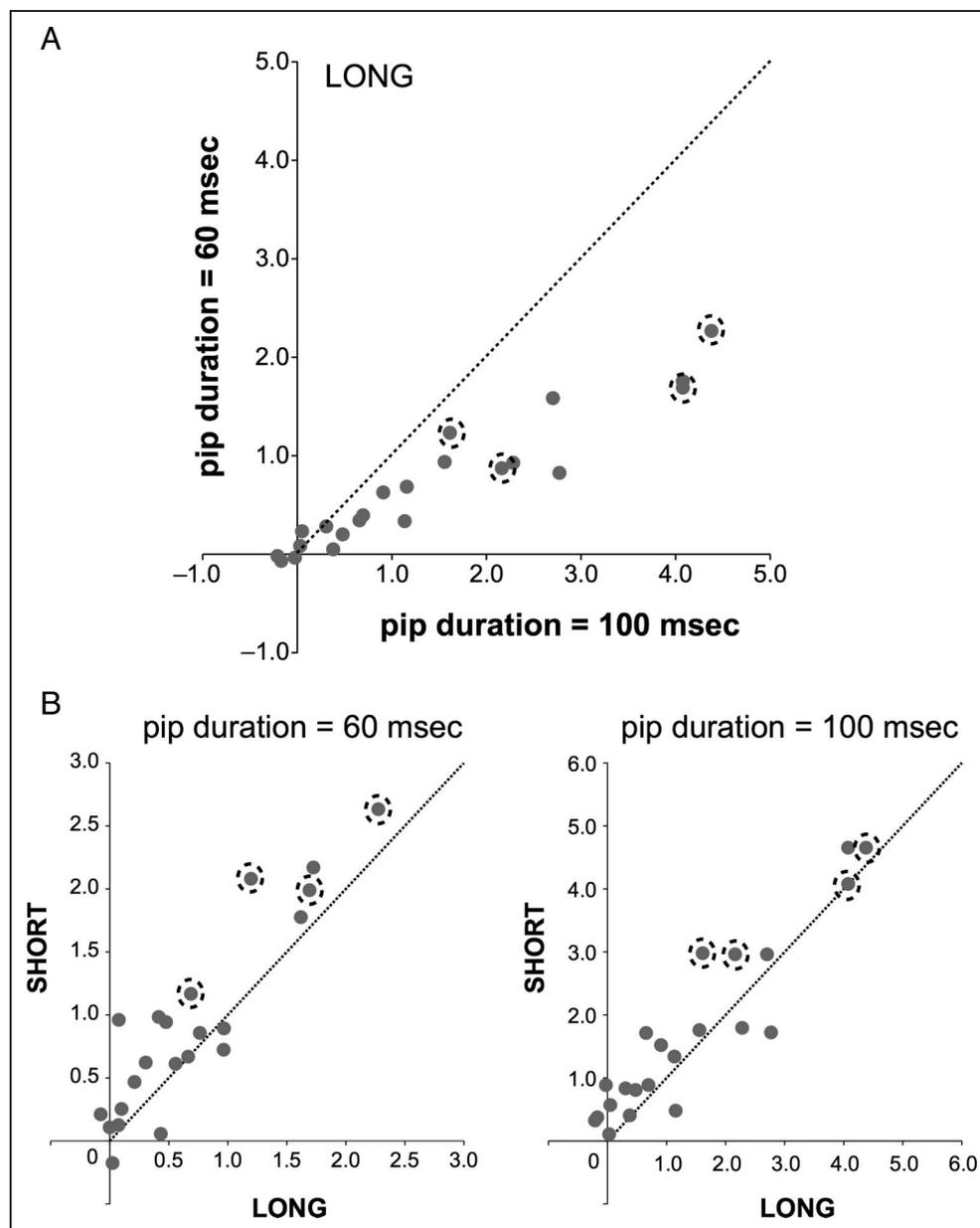
#### Stimuli

The stimulus set included SHORT (SHORT-expected; SHORT-unexpected) and CONT (CONT-expected; CONT-unexpected) stimuli, identical to those in Experiment 3, above. Only signals with tone pip durations of 60 msec were used (total stimulus duration: 1320–1500 msec for SHORT, 2160–2340 msec for CONT; 22–25 pips before transition, randomly determined for each stimulus). Equal numbers of signals (126) for each stimulus type were generated, making the occurrence of change within a stimulus unpredictable. A decoy task, identical to that in Experiment 2, was used.

#### Neuromagnetic Recording and Data Analysis

As in Experiment 2.

**Figure 5.** Results of Experiment 3 ( $n = 22$ ). (A) Comparison between  $d'$  sensitivity scores for pip duration = 100 msec (x axis) and pip duration = 60 msec (y axis). The data for the four participants with musical training are indicated with dashed circles. Overall participants exhibited a substantial decline in performance when pip durations were shortened from 100 to 60 msec. (B) Comparison between  $d'$  sensitivity scores in the LONG and SHORT conditions for pip duration of 60 msec (left) and 100 msec (right). The majority of participants exhibited better performance on SHORT relative to LONG for both durations.

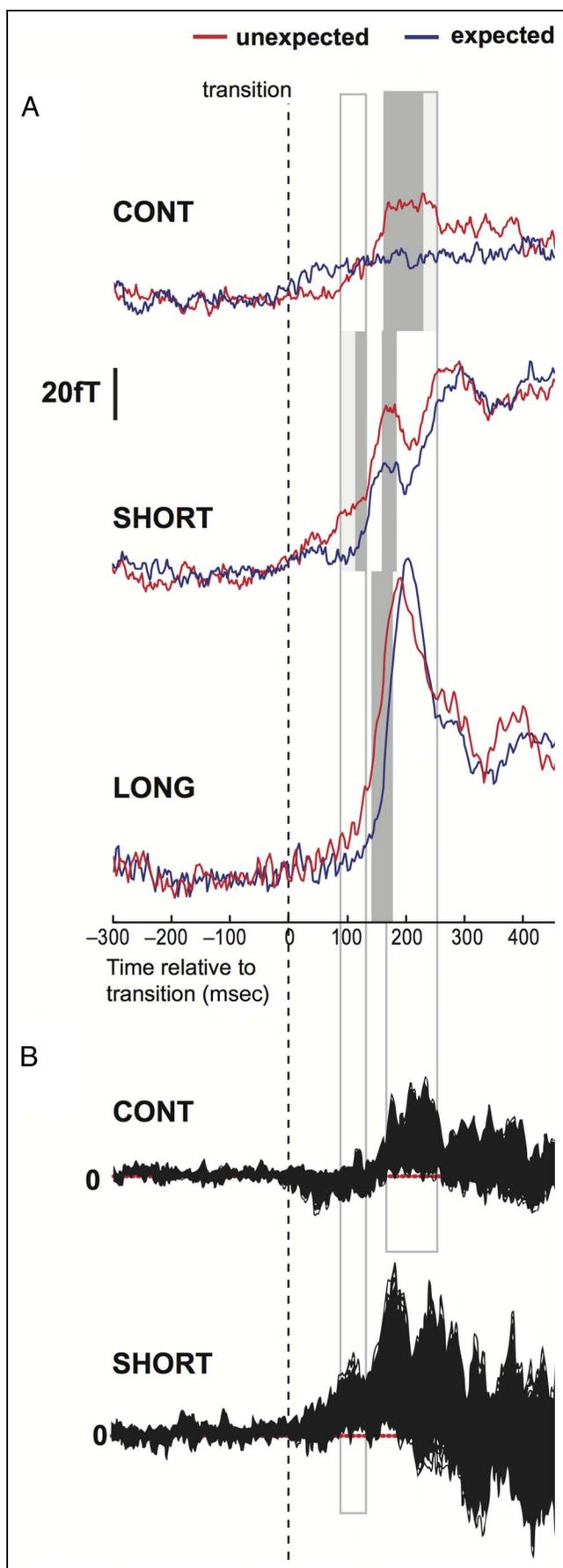


## Results

The group-RMS of auditory-evoked responses for the CONT and SHORT conditions around the transition is shown in Figure 6A. The data from the LONG condition (Experiment 2) are also provided for comparison. The zero on the abscissa corresponds to the onset of the frequency deviant (see Figure 4). The “expected” conditions (CONT-expected; SHORT-expected; LONG-expected) are plotted in blue, and the “unexpected” conditions (CONT-unexpected; SHORT-unexpected; LONG-unexpected) are in red. Gray bands indicate temporal intervals where a bootstrap analysis (see Methods) indicated significant differences between “-expected” and “-unexpected” conditions. This analysis was conducted on the entire stimulus epoch from 500 msec before the onset of the

transition to 1000 msec after the offset (a total duration of 1500 msec), and the intervals shown in Figure 6 were the only significant ones found within the entire range.

For the CONT-unexpected condition, the onset of the deviant tone elicited a prominent response with a first peak at 195 msec in the left hemisphere (LH) and 170 msec in the RH. Bootstrap analysis revealed a significant difference between the expected and unexpected conditions emerging from 160 msec until 250 msec after the onset of the deviant tone for the LH and 163 msec until 220 msec for the right. In the RH, the bootstrap distribution was bimodal, such that, on a certain proportion (less than 50%) of the executions of the bootstrap routine, the significant interval extended to 255 msec. This interval is reported in a



lighter shade of gray in Figure 6. The response returned to baseline 300 msec later.

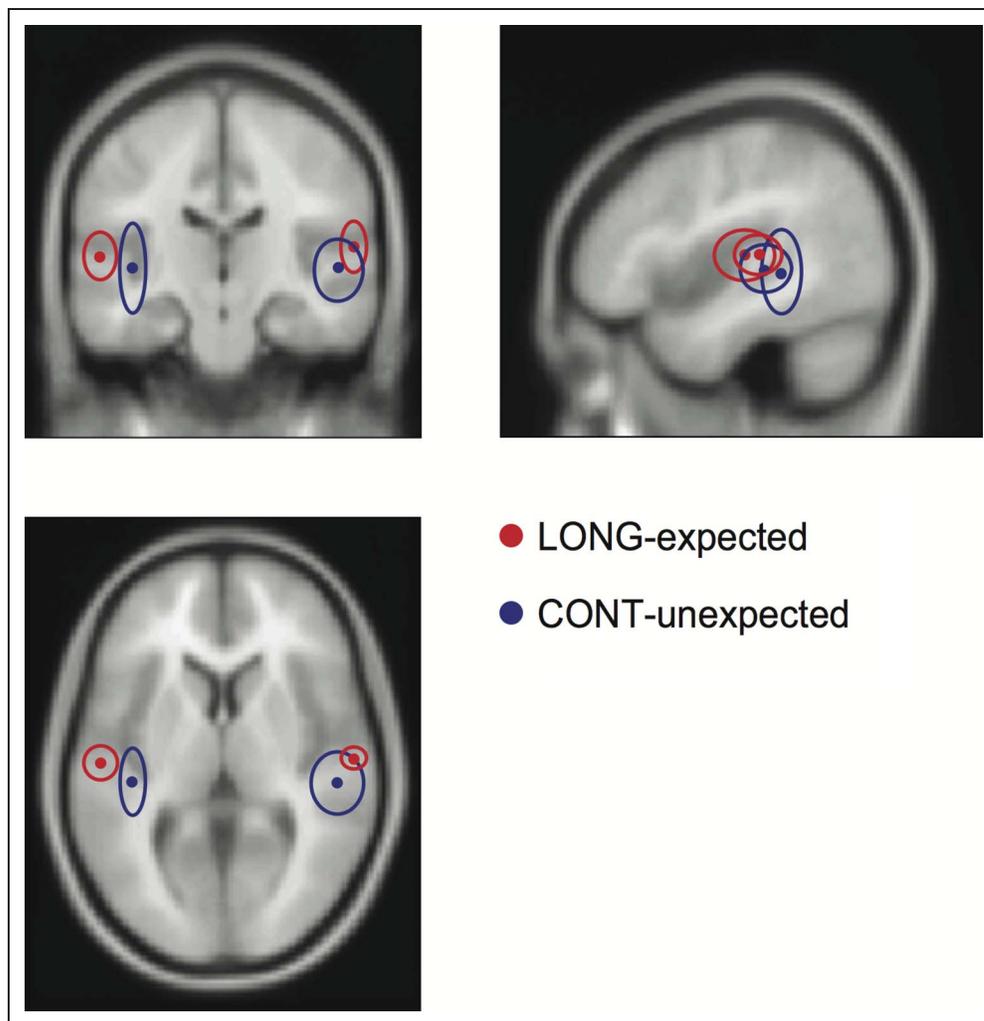
For the SHORT condition, both the expected and unexpected conditions showed a series of peaks at 175 and 290 msec corresponding to the response to the offset of the stimulus sequence (i.e., transition to silence). Similar to data for the LONG condition, the activation data indicate that responses to the -unexpected condition rise slightly before those to the -expected condition. This suggests that the deviation in frequency was detected before the subsequent change in the sequence (offset or violation of duration, respectively). Similar to the data for the LONG condition in Experiment 2, significant differences between responses to SHORT-expected and SHORT-unexpected conditions were observed only in the RH, between 110 and 133 msec and between 160 and 186 msec following the violation of the frequency pattern. The bootstrap distribution here was also bimodal, with about half of the runs yielding a significant interval starting as early as 90 msec (see Figure 6). To demonstrate the robustness of the observed effects, Figure 6B shows the bootstrap difference time series for responses to the CONT and SHORT conditions.

It is important to note that the peak in the response to the SHORT condition is difficult to compare directly to the peak in response to the LONG condition. The response to LONG-expected is a “pure” response to the endogenous process of detecting the violation of duration, whereas the peak in the response to the SHORT-expected condition is a response to an exogenous event (offset of stimulation). Overall, the SHORT condition evokes a pattern of responses very similar to that evoked by the LONG condition; in both conditions, responses to the -unexpected stimulus rise before those to -expected (but with an overall similar shape).

Theoretically, one might expect that the responses to the different -unexpected stimuli would all start at the same time, reflecting the early detection of the frequency deviation. However, the present data (see Figure 6) appear to demonstrate latency differences of up to 30 msec between the time points at which the

**Figure 6.** MEG data (RH) from Experiment 4. (A) Responses to the -unexpected and -expected signals in CONT, SHORT, and LONG conditions. Note that the LONG data are reproduced from Experiment 2. Gray shading indicates the temporal intervals where a repeated-measures Bootstrap analysis indicated significant differences between the -unexpected and -expected conditions. For both SHORT and CONT the bootstrap distribution was bimodal, and time intervals where the difference was not systematically significant are highlighted in a lighter shade of gray. (B) To demonstrate the robustness of the statistical effect, we show the results of the bootstrap repeated-measures procedure. The black curves are mean differences between conditions ( $x$  axis: time;  $y$  axis: difference (squared) between the -unexpected and -expected stimuli) computed over 1000 bootstrap resamplings; their scatter provides a measure of variability across participants. At each time point, the difference is deemed to be significant if the proportion of iterations above/below the zero line is less than 0.01 for 10 adjacent samples (15 msec).

**Figure 7.** Source localizations at the peak of the RMS response in CONT-unexpected (blue) and LONG-expected (red). Dots represent the group mean, and ellipses represent the standard deviation in each direction. These are superimposed in a “glass brain” fashion on a template T1-weighted image in MNI space with the coronal and axial sections at  $y = -20$  and  $z = 5$  mm, respectively.



different “-unexpected” stimuli diverge from their respective control conditions. Although it is difficult to assess these differences formally, because of the physical differences between SHORT and LONG stimuli and that the two responses were recorded from different groups, it appears that the apparent differences might be resulting from noise in the data since all red lines (SHORT, LONG, CONT) appear to show similar trends.

### SOURCE ANALYSIS

All participants showed bilateral dipolar evoked responses to the transition in both LONG-expected, LONG-unexpected (from Experiment 2), and CONT-unexpected, which were well modeled by a single ECD in each hemisphere. Responses to the SHORT-expected and SHORT-unexpected conditions revealed a pattern that is more consistent with two ECDs in each hemisphere, one in A1 (primary auditory cortex) and another in STG. This is perhaps not surprising given that sound offset (unlike the CONT and LONG changes) is a much more distributed process, evoking responses along the ascending auditory pathway.

Here, we focus on comparing responses to LONG-expected—a “pure” change in duration—and CONT-unexpected—a “pure” change in frequency pattern. In 12 of 13 participants in the LONG-expected condition and 10 of 13 in the CONT-unexpected condition, an ECD explaining over 80% of variance was found. The mean dipole location per hemisphere and the standard deviation are shown in Figure 7. For LONG-expected, most ECDs were located in Heschl’s gyrus or planum temporale, the group means occurred at the border of planum temporale and anterior-lateral Heschl’s gyrus (MNI coordinates: RH [57 -15 10], LH [-55 -18 8]). For CONT-unexpected, the group means were located more medially and caudally in superior temporal gyrus (MNI coordinates: RH [53 -28 2], LH [-42 -20 6]). Statistical analysis (a Hotelling T-square test for two multivariate independent samples) revealed significant differences between the two locations ( $p = .02$  in RH and  $p = .0005$  in the LH).

### DISCUSSION

This study demonstrates a form of “change deafness” brought about by inter-feature perceptual masking: We

show that, although information about frequency pattern violation in our stimuli is robustly represented in neural activity at a relatively late stage of auditory processing (secondary auditory cortex), it is rendered perceptually inaccessible by a subsequent change (here: a change in duration or offset) occurring in close temporal proximity. The data have important implications for our understanding of the interaction between “bottom–up” driven feature change extraction and higher level stages of processing that bring such changes to conscious awareness. These are discussed, in turn, below.

### Sensitivity to Frequency Patterns

Our data demonstrate that human auditory cortex is tuned to frequency patterning. In other words, the auditory system is able to, automatically and preattentively, learn an on-going pattern of sound frequencies and respond rapidly when this pattern is violated. Because the deviant tone was of a frequency that was equally present in the preceding sequence, to detect the violation, one must acquire the specific pattern structure, namely, that frequency  $f_1$  is followed by  $f_2$ , then by  $f_3$ , etc., or alternatively learn the specific relation, in terms of frequency shifts, between tones (Demany & Ramos, 2005) and identify a mismatch between the expected frequency and the actual stimulus input. The fact that responses evoked by violations in the frequency pattern were observed even when listeners were unable, at a behavioral level, to detect them, suggests that the MEG responses are tapping into efficient, rapid, mechanisms predominantly based on bottom–up processing.

To date, the most commonly employed tool to study preattentive regularity extraction has been the MMN paradigm (Näätänen, Astikainen, Ruusuvirta, & Huotilainen, 2010). The MMN component is elicited by sounds violating some regular aspect of the preceding sequence of sounds (including abstract rules regarding the succession of elements in a sequence) and is hypothesized to reflect a discrepancy between the memory trace or expectations generated by the standard stimulus and the new, deviant information or processes that update the internal representation when a previously registered regularity is violated. The MMN is elicited even when participants’ attention is diverted away from the stimulus, allowing for the exploration of preattentive, bottom–up mediated, stimulus representations (Näätänen et al., 2010). Using tone pairs (with fixed frequencies) as stimuli, Tervaniemi et al. (1994) demonstrated that an MMN is elicited by reversing the order of the two tones, by repeating the same tone, or by omitting the second tone, suggesting that frequency relations between the members of the pair are extracted by the auditory system. Nordby, Roth, and Pfefferbaum (1988) used a continuous sequence of alternating tones of two frequencies (ABABAB...) and showed that an MMN can be elicited by an occasional repetition

(e.g., ABABAA). Schröger, Paavilainen, and Näätänen (1994) used a continuously repeating sequence of five tones of different frequencies. They demonstrated that the MMN can be elicited when Tone 1 and Tone 3 in the pattern were occasionally interchanged. Tervaniemi et al. (1994) further demonstrated that the auditory system is sensitive to regularities defined in terms of abstract frequency relationships: In their study, an MMN response was elicited for an occasional ascending or repeating tone in a sequence of descending tones. An even more striking demonstration of sensitivity to abstract frequency relations was provided by Wolff and Schröger (2001), who showed that infrequent repetitions of a tone, in a rapidly presented sequence of tones varying in frequency, can elicit an MMN response, suggesting that frequency variation per se can be extracted as a “regularity rule” by the auditory system.

Our paradigm differs “cosmetically” from the classic MMN method in that we record responses to changes in on-going sequences rather than comparing “isolated” responses to standards and deviant tones, a method that is difficult to apply to the very rapid sequences used in this study. It is therefore very likely that the responses we observe are generated by the same neural substrates as those contributing to the MMN (see also Vaz Pato & Jones, 1999). The present results extend previous findings by demonstrating that the auditory system is able to form specific frequency expectations. For instance, the fact that the system can acquire relationships between frequencies (“the next frequency should be same/different/higher/lower”) is not strictly equivalent to expecting a specific frequency (“the next frequency should be  $f_1$ ”). Importantly, unlike many previous studies, the frequency patterns in our stimuli changed from trial to trial (no more than eight cycles were presented before the violation) requiring rapid on-line learning.

### Frequency Violation and Duration Violation

It has been demonstrated previously that frequency and duration violations are computed by distinct systems in auditory cortex (Molholm, Martinez, Ritter, Javitt, & Foxe, 2005; Giard et al., 1995; Levänen, Hari, McEvoy, & Sams, 1993). Molholm et al. (2005) showed that cortical activations related to processing frequency deviance (albeit in a simple oddball design) were situated within primary auditory cortex, whereas those related to detecting duration deviants were localized to nonprimary areas. In our data, the response to a violation in the frequency pattern localizes to secondary auditory areas, probably because that the violation is not an odd ball (a new frequency appearing after a sequence of identical tones) but a more complex pattern violation, likely involving computations in “higher-level” auditory centers. Importantly, however, the source localization data (see Figure 7) clearly indicate a difference between responses to violations of duration

and violations of the frequency pattern, suggesting these are computed by different neural substrates. The observed perceptual interference does not therefore arise at the point of feature change extraction, but rather at a later processing stage where information about both changes is combined.

### Perceptual Masking and “Change Deafness”

The pattern of results in Experiments 1 and 3 suggest that the change in duration captured listeners’ processing/attentional resources before the information about frequency change reached awareness. Namely, despite the fact that our participants were actively trying to listen for frequency pattern violation and ignore the task-irrelevant duration change, the change in duration inadvertently led to an inability to detect the (otherwise easily audible) frequency violation. The form of perceptual masking revealed here depends on the timing of the two events ( $d'$  drops by more than 50% when pip duration is shortened from 100 to 60 msec). Importantly, this “change deafness” can be alleviated by manipulating the properties of the masking event. We show that, when the frequency change is followed by stimulus offset (“SHORT” stimuli), performance is enhanced relative to when it is followed by violation of duration (“LONG” stimuli). Because violations of duration might be indicative of the emergence of a new event, this finding is consistent with accumulating evidence suggesting that onsets are perceptually more salient than offsets (Cervantes Constantino et al., 2012; Cole & Kuhn, 2010; Phillips & Hall, 2001). These data, therefore, demonstrate that change deafness does not arise simply because two events occur close together in time but depends, at least in part, on the perceptual salience of the second event.

Incidentally, this change deafness is not limited to changes that arise in the frequency and duration of tone sequences. Similar effects also occur when a step-wise increase in loudness is introduced with the frequency change. Specifically, in a preliminary study, we compared performance in the CONT condition (-expected and -unexpected; tone pip duration of 60 msec) and a modified condition (CONT\_LC; “CONT loudness change”) where a step change in sound level (+10 dB) was introduced, starting from the tone following the frequency change (in other words, a change in loudness occurred 60 msec after the change in frequency). Both -expected and -unexpected conditions contained a loudness change. In a “pilot” run with five participants, we found that, although the mean performance on CONT (differentiating -expected from -unexpected) was  $d' = 3.14$ , that in CONT\_LC dropped to  $d' = 0.6$ , suggesting that the introduction of a change in loudness 60 msec after the frequency change made the frequency change essentially impossible to detect.

Previous examinations of auditory distraction demonstrate that rare changes in an unattended sound feature

(e.g., a rare change in frequency when listeners are categorizing tones according to duration) impair performance by leading to longer RTs and increased confusion (Schröger & Wolff, 1998; see also Roeber, Widmann, & Schröger, 2003; Schröger, Giard, & Wolff, 2000). The degree to which participants are distracted depends both on the properties of the distracting event and also on the characteristics of the main task performed by listeners (e.g., Lavie, 2005). The present paradigm differs from those described in previous studies in several important respects: first, the changes in the two features (frequency and duration) always occur together, and the distraction therefore does not arise because of a novel event momentarily seizing attention, rather different features of the same object change (almost) simultaneously, potentially competing for processing resources. Second, in previous experiments, the to-be-attended feature was easily discriminable and the negative effects of the deviant were mostly because of momentary confusion. In contrast, the present results demonstrate severe “change deafness.”

Instead, our findings appear to be closely related to the “attentional blink” phenomenon (AB). AB refers to the inability to successfully identify the second of two sensory targets presented in rapid succession, and although it has mostly been observed in vision (for a review, McLaughlin, Shore, & Klein, 2001), increasing evidence suggests the existence of a similar effect in the auditory modality (Horváth & Burgyán, 2011; Shen & Mondor, 2006; Duncan, Martens, & Ward, 1997). A key distinction is that the “change deafness” at the focus of the present work demonstrates temporally opposite effects—a later occurring signal rendering an earlier event, perceptually undetectable. The attentional blink effect is generally thought to reflect restrictions on the deployment of limited capacity attention/processing resources in time, such that, while the first target is being processed, the internal representation of the second target (stored in working memory) decays or is overwritten by subsequently presented input (e.g., Shen & Mondor, 2006). The present results suggest that processing is not necessarily serial and that attentional/computational resources can be captured by a subsequent event before processing of a previous event is complete, resulting in a “backward masking”-induced “deafness.”

In a series of seminal experiments designed to measure the duration of auditory working memory, Massaro (1970, 1971) instructed listeners to identify the frequency of a short (20 msec) test tone. He demonstrated that if this tone is followed by a subsequent, “masking” tone, participants’ performance worsens significantly, with the disruption persisting up to an intertone interval of about 250 msec. This effect was interpreted as suggesting that the masking tone overwrote the representation of the test tone, rendering the pitch information inaccessible, and the critical intertone interval was suggested to reflect the duration of a central auditory image. The

present findings offer an alternative explanation to this effect, not in terms of working memory, but rather as an attentional phenomenon. Both the MEG findings, which suggest that the frequency violation is coded by auditory cortex, and the fact that the degree of backward masking varies as a function of the salience of the second event are more consistent with an attentional account.

Importantly, while the masking in Massaro's work and the attentional blink literature involved externally occurring events (i.e., the onset of a subsequent sensory stimulus), the present example is different in that the violation of duration in the LONG signals is an internally generated event. It is not evoked by an exogenous sensory occurrence (nothing changes at the level of the sensory receptor 60 msec after the change in frequency) but rather involves the output of a mechanism, likely located relatively high up in the processing hierarchy, that signals a violation of expectation—i.e., that the duration of the tone exceeded the expected duration. Consequently, the data demonstrate that competition for attentional/perceptual resources is not reserved for exogenous events; the violation of frequency and duration, reflecting changes in the features of the same on-going object, are computed in parallel, by different neural systems that then compete for perceptual priority.

These results reveal a rather late bottleneck in the route from “bottom-up” change detection to the level of conscious awareness—despite the fact that a violation in the pattern of an evolving sequence of tone pips has been identified by the auditory system, it nevertheless fails to reach conscious awareness because some form of internal attention required to achieve this awareness has been suppressed or diverted away by a subsequently occurring event (here, a change in duration). The consequence is a severe perceptual “deafness” to an otherwise easily detectable feature change.

## Acknowledgments

We are grateful to David Bradbury and the Radiographer team at the University College London Wellcome Trust Centre for Neuroimaging for excellent MEG technical support and to Alain de Cheveigné for comments and discussion. This study was supported by a Deafness Research UK fellowship and a Wellcome Trust project grant to M. C.

Reprint requests should be sent to Dr. Maria Chait, Ear Institute, 332 Gray's Inn Road, London WC1X 8EE, UK, or via e-mail: [m.chait@ucl.ac.uk](mailto:m.chait@ucl.ac.uk).

## REFERENCES

- Alain, C., Woods, D. L., & Ogawa, K. H. (1994). Brain indices of automatic pattern processing. *NeuroReport*, *6*, 140–144.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, *8*, 465–471.
- Cervantes Constantino, F., Pinggera, L., Paranamana, S., Kashino, M., & Chait, M. (2012). Detection of appearing and disappearing objects in complex acoustic scenes. *PLoS One*, *7*, e46167.
- Chait, M., Poeppel, D., de Cheveigné, A., & Simon, J. Z. (2007). Processing asymmetry of transitions between order and disorder in human auditory cortex. *Journal of Neuroscience*, *27*, 5207–5214.
- Chait, M., Poeppel, D., & Simon, J. Z. (2008). Auditory temporal edge detection in human auditory cortex. *Brain Research*, *1213*, 78–90.
- Cole, G. G., & Kuhn, G. (2010). Attentional capture by object appearance and disappearance. *Quarterly Journal of Experimental Psychology* (2006), *63*, 147–159.
- Czigler, I., & Winkler, I. (1996). Preattentive auditory change detection relies on unitary sensory memory representations. *NeuroReport*, *7*, 2413–2417.
- Delgutte, B. (1990). Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers. *The Journal of the Acoustical Society of America*, *87*, 791–809.
- Demany, L., & Ramos, C. (2005). On the binding of successive sounds: Perceiving shifts in nonperceived pitches. *The Journal of the Acoustical Society of America*, *117*, 833–841.
- Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, *387*, 808–810.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. *Monographs on statistics and applied probability*. New York: Chapman & Hall.
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. A. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron*, *61*, 317–329.
- Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates “change deafness” in complex auditory scenes. *Current Biology: CB*, *15*, 1108–1113.
- Giard, M. H., Lavikahen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, *7*, 133–143.
- Gregg, M. K., & Snyder, J. S. (2012). Enhanced sensory processing accompanies successful detection of change for real-world sounds. *Neuroimage*, *62*, 113–119.
- Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, *6*, e138.
- Horváth, J., & Burgyán, A. (2011). Distraction and the auditory attentional blink. *Attention, Perception & Psychophysics*, *73*, 695–701.
- Jones, S. (2002). The internal auditory clock: What can evoked potentials reveal about the analysis of temporal sound patterns, and abnormal states of consciousness? *Neurophysiologie Clinique/Clinical Neurophysiology*, *32*, 241–253.
- Kiebel, S. J., Daunizeau, J., Phillips, C., & Friston, K. J. (2008). Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. *Neuroimage*, *39*, 728–741.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75–82.
- Levänen, S., Hari, R., McEvoy, L., & Sams, M. (1993). Responses of the human auditory cortex to changes in one versus two stimulus features. *Experimental Brain Research*, *97*, 177–183.
- Lütkenhöner, B., & Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neurotology*, *3*, 191–213.

- Massaro, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, *85*, 411–417.
- Massaro, D. W. (1971). Effect of masking tone duration on preperceptual auditory images. *Journal of Experimental Psychology*, *87*, 146–148.
- McLaughlin, E. N., Shore, D. I., & Klein, R. M. (2001). The attentional blink is immune to masking-induced data limits. *The Quarterly Journal of Experimental Psychology: Series A, Human Experimental Psychology*, *54*, 169–196.
- Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*, *15*, 545–551.
- Moore, B. C. (2003). *An introduction to the psychology of hearing* (vol. 4.). San Diego: Academic Press.
- Moore, B. C., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America*, *74*, 750–753.
- Näätänen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory-cognitive core of cognitive processes. *Brain Research Reviews*, *64*, 123–136.
- Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988). Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. *Psychophysiology*, *25*, 262–268.
- Oceák, A., Winkler, I., & Sussman, E. (2008). Units of sound representation and temporal integration: A mismatch negativity study. *Neuroscience Letters*, *436*, 85–89.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Overath, T., Cusack, R., Kumar, S., Von Kriegstein, K., Warren, J. D., Grube, M., et al. (2007). An information theoretic characterisation of auditory encoding. *PLoS Biology*, *5*, e288.
- Phillips, D. P., & Hall, S. E. (2001). Spatial and temporal factors in auditory saltation. *The Journal of the Acoustical Society of America*, *110*, 1539–1547.
- Plack, C. J., & Moore, B. C. (1990). Temporal window shape as a function of frequency and level. *The Journal of the Acoustical Society of America*, *87*, 2178–2187.
- Roberts, T. P. L., Ferrari, P., Stufflebeam, S. M., & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology*, *17*, 114.
- Roeber, U., Widmann, A., & Schröger, E. (2003). Auditory distraction by duration and location deviants: A behavioral and event-related potential study. *Brain Research. Cognitive Brain Research*, *17*, 347–357.
- Schröger, E., Giard, M. H., & Wolff, C. (2000). Auditory distraction: Event-related potential and behavioral indices. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *111*, 1450–1460.
- Schröger, E., Paavilainen, P., & Näätänen, R. (1994). Mismatch negativity to changes in a continuous tone with regularly varying frequencies. *Electroencephalography and Clinical Neurophysiology*, *92*, 140–147.
- Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Brain Research. Cognitive Brain Research*, *7*, 71–87.
- Sculthorpe, L. D., Collin, C. A., & Campbell, K. B. (2008). The influence of strongly focused visual attention on the detection of change in an auditory pattern. *Brain Research*, *1234*, 78–86.
- Shen, D., & Mondor, T. A. (2006). Effect of distractor sounds on the auditory attentional blink. *Perception & Psychophysics*, *68*, 228–243.
- Sussman, E., Winkler, I., Ritter, W., Alho, K., & Näätänen, R. (1999). Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. *Neuroscience Letters*, *264*, 161–164.
- Tervaniemi, M., Maury, S., & Näätänen, R. (1994). Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *NeuroReport*, *5*, 844–846.
- Vaz Pato, M., & Jones, S. J. (1999). Cortical processing of complex tone stimuli: Mismatch negativity at the end of a period of rapid pitch modulation. *Cognitive Brain Research*, *7*, 295–306.
- Warren, R. M., & Byrnes, D. L. (1975). Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners. *Attention, Perception, & Psychophysics*, *18*, 273–280.
- Wolff, C., & Schröger, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neuroscience Letters*, *311*, 37–40.