Maastricht University

Hearing an illusory vowel in noise: suppression of auditory cortical activity

Citation for published version (APA):

Riecke, L., Vanbussel, M., Hausfeld, L., Baskent, D., Formisano, E., & Esposito, F. (2012). Hearing an illusory vowel in noise: suppression of auditory cortical activity. *Journal of Neuroscience*, *32*(23), 8024-8034. https://doi.org/10.1523/JNEUROSCI.0440-12.2012

Document status and date:
Published: 06/06/2012

DOI:
10.1523/JNEUROSCI.0440-12.2012

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 19 Apr. 2024

Behavioral/Systems/Cognitive

Hearing an Illusory Vowel in Noise: Suppression of Auditory Cortical Activity

Lars Riecke,[1] Mieke Vanbussel,[1] Lars Hausfeld,[1] Deniz Başkent,[2] Elia Formisano,[1] and Fabrizio Esposito[1,3]

[1]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, The Netherlands, [2]Department of Otorhinolaryngology/Head and Neck Surgery, Graduate School of Medical Sciences (Research School of Behavioural and Cognitive Neurosciences), University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands, and [3]Department of Neuroscience, University of Naples "Federico II," 80131 Naples, Italy

Human hearing is constructive. For example, when a voice is partially replaced by an extraneous sound (e.g., on the telephone due to a transmission problem), the auditory system may restore the missing portion so that the voice can be perceived as continuous (Miller and Licklider, 1950; for review, see Bregman, 1990; Warren, 1999). The neural mechanisms underlying this continuity illusion have been studied mostly with schematic stimuli (e.g., simple tones) and are still a matter of debate (for review, see Petkov and Sutter, 2011). The goal of the present study was to elucidate how these mechanisms operate under more natural conditions. Using psychophysics and electroencephalography (EEG), we assessed simultaneously the perceived continuity of a human vowel sound through interrupting noise and the concurrent neural activity. We found that vowel continuity illusions were accompanied by a suppression of the 4 Hz EEG power in auditory cortex (AC) that was evoked by the vowel interruption. This suppression was stronger than the suppression accompanying continuity illusions of a simple tone. Finally, continuity perception and 4 Hz power depended on the intactness of the sound that preceded the vowel (i.e., the auditory context). These findings show that a natural sound may be restored during noise due to the suppression of 4 Hz AC activity evoked early during the noise. This mechanism may attenuate sudden pitch changes, adapt the resistance of the auditory system to extraneous sounds across auditory scenes, and provide a useful model for assisted hearing devices.

Introduction

One prerequisite for the restoration of an interrupted sound (target) in noise is that the intensity of the target does not change markedly at the target–noise transitions (Bregman and Dannenbring, 1977), i.e., the offset and reintroduction (reonset) of the target need to be masked (Warren et al., 1972). This "no-discontinuity" rule implies that the auditory system fails to represent the offset/reonset of the target (Houtgast, 1972; Bregman, 1990). For that case, there is no sensory evidence to infer that the target has stopped; thus, a continuous percept may emerge. The sensory evidence may be provided by on/off neurons, which are characterized by transient responses to sound onsets/offsets and exist throughout the auditory pathway (for review, see Phillips et al., 2002; Scholl et al., 2010). For example, in the thalamus and auditory cortex (AC) of animals, neural on/off responses to an interrupted tone may be suppressed, provided that the tone's onset/offset is masked (Schreiner, 1980; Petkov et al., 2007). At a larger spatial scale, 4 Hz activity in human AC evoked by a gap in an interrupted tone may be suppressed, provided that the tone elicits a continuity illusion (Riecke et al., 2009b). Thus, suppression of onset/offset-evoked central responses may be a prerequisite for auditory restoration. This putative mechanism could also be involved in the restoration of ecologically valid sounds: the no-discontinuity rule has been validated behaviorally for speech (Başkent et al., 2009), and continuity illusions of vowels and speech have been associated with changes in hemodynamic activity in temporal and parietal cortices (Shahin et al., 2009; Riecke et al., 2011a), although neural on/off responses could not be resolved in these studies. Therefore, evidence for a role of onset/offset suppression for the restoration of ecologically valid sounds is sparse (Shahin et al., 2012).

Another important factor whose underlying mechanisms are poorly understood is context. The restoration of a tone, for example, may be hampered by a preceding tone that is clearly interrupted, but facilitated by a preceding tone that is clearly uninterrupted (Riecke et al., 2011b). This context-specific perceptual adaptation could serve to accommodate listeners' resistance to extraneous sounds across different scenes. While context-specific adaptation has been demonstrated for AC neurons in animals, comparable evidence for on/off neurons involved in auditory restoration is missing.

Here, we investigated three critical issues. (1) Does restoration of a natural sound involve suppression of neural on/off responses? (2) Does this putative mechanism contribute differently to restorations of differently complex sounds? (3) Do the on/off

Received Jan. 31, 2012; revised April 5, 2012; accepted April 26, 2012.

Author contributions: L.R., M.V., E.F., and F.E. designed research; L.R. and M.V. performed research; L.R., M.V., L.H., and F.E. analyzed data; L.R., M.V., L.H., D.B., E.F., and F.E. wrote the paper.

This work was supported by the Netherlands Organization for Scientific Research Veni Grant 451-11-014 (to L.R.). We thank Niels Disbergen for help with data acquisition and three anonymous reviewers for useful comments.

The authors declare no conflict of interest.

Correspondence should be addressed to Lars Riecke, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: L.Riecke@MaastrichtUniversity.nl.

DOI:10.1523/JNEUROSCI.0440-12.2012

Copyright © 2012 the authors 0270-6474/12/328024-11$15.00/0

responses adapt restoration to different auditory contexts? Human listeners judged the continuity of a vowel that was partially replaced by noise (Fig. 1). Using EEG, we identified on/off responses and tested whether they are suppressed during continuity illusions. Further, we tested whether the strength, latency, and cortical origin of the suppression differed for restorations of differently complex sounds, exploiting an illusory vowel and an illusory tone. Finally, we tested whether restoration-related on/off responses depended on the intactness of auditory context.

Materials and Methods

Stimuli. Auditory stimuli were comprised of a vowel and a noise burst (Fig. 1*A*). The vowel was recorded from a 24-year-old male speaker (fundamental frequency f_0: 100 Hz) uttering a sustained [a:]. The utterance contained no linguistic information and had strong pitch; the latter feature is critical for speaker identification (van Dommelen, 1990) and concurrent vowel identification (Assmann and Summerfield, 1990). After digitization (sampling rate: 44.1 kHz), the recording was bandpass-filtered into an audio frequency range similar to that of telephone lines [cutoffs: 680 and 3900 Hz, finite impulse response (FIR) filter] to make the conditions more relevant to communication devices. The noise burst was obtained by bandpass-filtering broadband Gaussian noise into a frequency range that completely overlapped the vowel spectrum [cutoffs: same as for vowel filtering ±⅓ octave (oct.), FIR filter].

The first 2800 ms of the filtered vowel recording, starting from the voice onset, was used for creating the stimuli. An interruption of 600 ms was introduced by muting the portion between 800 and 1400 ms. The resulting gap was filled with the filtered noise. All onsets and offsets were ramped in such a way that the midpoints of the vowel off-ramps coincided with the midpoints of the noise on-ramps and vice versa (linear ramp durations: 75 ms for the vowel, 25 ms for the gap and the noise).

Contextual stimuli were comprised of a longer version of the filtered vowel (duration: 3800 ms) without noise. These contextual stimuli were either intact (containing no gap) or fragmented (containing several short gaps, introduced by modulating the amplitude of the vowel with a periodic square wave with 200 ms period and 50% duty cycle). The resulting onsets and offsets were ramped (linear ramps, duration: 3 ms). The auditory stimuli were processed using Matlab (MathWorks) and presented diotically through loudspeakers (JBL Professional) using Presentation software (Neurobehavioral Systems).

Participants and task. Fourteen volunteers (eight females) between 20 and 39 years old (mean: 25 years) with no reported hearing or motor problems participated. Their perception of the continuity of the vowel through the noise was assessed using a continuity rating task (Riecke et al., 2008). Each trial was comprised of a stimulus interval and a delayed response interval (Fig. 1*B*). These intervals were separated by a silent rest period (duration: 1 s) and indicated by color changes of a central fixation cross. On adaptation trials, the stimulus interval was preceded by an additional interval during which the contextual stimulus was presented, followed by a 180 ms silence (see Fig. 8) (Riecke et al., 2009a). Participants performed ratings during brief response intervals by pressing one of four buttons with the middle or index finger of their left hand (corresponding to buttons labeled "most likely discontinuous" and "likely discontinuous") or right hand (corresponding to buttons labeled with "likely continuous" and "most likely continuous"). The button labels were displayed together with the fixation cross on a LCD screen. Trials were preceded and followed by fixed rest periods (duration: 1 s each).

Design. The design included four stimulus conditions and two adaptation conditions. First, an interrupted vowel and an intact vowel were used to extract neural on/off responses (see Fig. 4*A*). To ensure clear audibility of the gap or the intact vowel during the noise, the noise spectrum level was attenuated in and around the strongest frequency component of the filtered vowel (the 12th harmonic of the voice fundamental). This was done equally for the interrupted vowel and the intact vowel by inserting a broad spectral notch in the noise (width: 1.2 oct., centered logarithmically on 1200 Hz, FIR filter). Thus, these two stimulus conditions differed only in the clear absence/presence of the vowel during the noise.

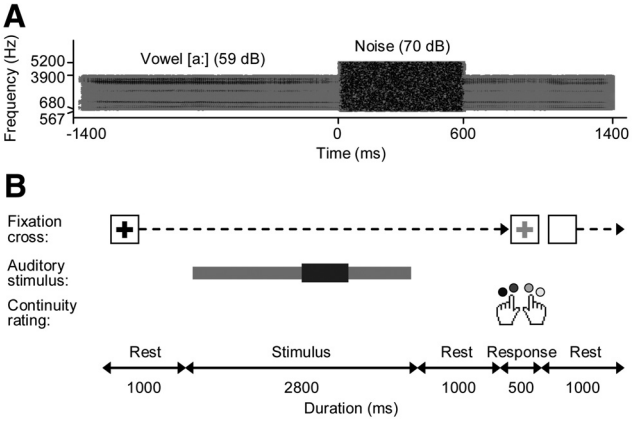

Figure 1. Vowel stimulus and continuity rating task. ***A***, A vowel recording from a male speaker was filtered and a portion of it was replaced by noise (stimulus spectrogram). ***B***, Participants were instructed to attend to the vowel and rate its continuity through the noise on a four-point scale. Stimulus and response intervals were visually cued and separated by silent rest periods.

Second, an ambiguous version of the interrupted vowel was used to investigate auditory restoration. This vowel could be perceived either veridically as discontinuous or illusorily as continuous (see Fig. 4*B*, left). To ensure perceptual ambiguity, the sound level of the vowel was set to a threshold that was estimated individually before the EEG experiment (see Procedure, below). To increase the sample size, a second ambiguous stimulus was included that was physically identical, except that a narrow notch was inserted in the noise (instead of no notch; notch width: 0.1 oct., centered logarithmically on 1200 Hz, FIR filter; see Fig. 4*B*, right). For each of these two ambiguous stimulus conditions, two rating conditions were defined *post hoc* based on the participant's dichotomized rating on each trial. These physically identical rating conditions were labeled as "No restoration" or "Restoration."

Finally, a clearly interrupted vowel (probe; same as Gap condition; see Fig. 4*A*, top) and a contextual stimulus that was either uninterrupted or repeatedly interrupted (see Stimuli, above; see Fig. 8) were used to investigate context-specific adaptation. The contextual stimuli served to adapt listeners to context that was either intact or fragmented. The subsequently presented probe served to evoke neural on/off responses and assess the strength of the putative adaptation under physically identical conditions (Fig. 4*C*).

The conditions described above were presented during 10 experimental runs (duration: 12 min each), each containing 100 randomly presented trials (20 trials of each of the four stimulus conditions and 10 trials of each of the two adaptation conditions). Inspection of participants' rating data revealed *post hoc* that reports of continuity and discontinuity occurred approximately equally often (probability of discontinuity reports: 51 ± 10%, mean ± SD across participants).

Procedure. Participants were seated in a comfortable chair in an electrically shielded and sound-attenuating booth. They were given written instructions to attend to the vowel, ignore the noise, and rate the continuity of the vowel through the noise. Participants first practiced this task in short runs of the experiment (run duration: 8 min; 3–7 runs depending on the participant). After each practice run, the presentation level of the vowel was adjusted based on the participant's ratings in the no-notch and narrow-notch conditions, while the overall noise level was held constant at 70 dB SPL. This was done repeatedly for maximally 1 h, until vowel continuity in the aforementioned conditions was reported reliably as ambiguous (i.e., each condition was rated approximately equally often as continuous and discontinuous for two consecutive runs). This adaptive procedure yielded an individualized vowel level, equaling, on average, 59 ± 3 dB SPL (mean ± SD across participants). This level was then used in the subsequent EEG experiments, where it remained constant across all conditions.

After completing the training, participants were admitted to the EEG experiments, which were conducted in the same booth on a consecutive day. For these experiments, participants were further instructed to keep their gaze at the fixation cross and delay any movement (e.g., button presses or eye blinks) to the response intervals. After practicing the EEG task for another 5–12 min (depending on the participant), each participant completed 10 experimental runs (including short breaks in between) while EEG was recorded simultaneously (see EEG recordings, below). All participants provided informed consent and the local ethics committee (Ethische Commissie Psychologie at Maastricht University) approved the study procedures.

EEG recordings. EEG was recorded from 64 scalp locations and referenced to the left mastoid, using Ag/AgCl electrodes (mounted in Easycaps, 10/20 system) and Neuroscan amplifiers. Electrooculography was recorded below the left eye using an additional electrode. Individual electrode locations were digitized in reference to a coordinate system defined by the nasion and the two preauricular points, using a Polhemus system. Interelectrode impedances were kept <5 kΩ by abrading the skin. EEG recordings were bandpass-filtered (cutoffs: 0.05 and 100 Hz, analog filter) and then digitized (sampling rate: 250 Hz).

EEG data preprocessing. The EEG data were analyzed using the EEGLAB toolbox (Delorme and Makeig, 2004) and custom Matlab scripts. A series of preprocessing and analysis steps were applied to the data, as summarized in Figure 2. The purpose of the preprocessing was, first, to reduce artifacts while preserving a reasonable amount of brain-related EEG data (epochs) and second, to extract auditory cortical activities from the artifact-reduced brain-related activities. Initial preprocessing steps involved bandpass filtering (cutoffs: 0.5 and 50 Hz, FIR filter), resampling (rate: 125 Hz), rereferencing to an average reference (based on the mean activity of all channels), and extraction of epochs. Trials contaminated by movements and nonrepetitive artifacts were identified semiautomatically using a ± 75 μV criterion and visual inspection. Trials during which participants were not engaged actively in the task were identified based on missing button responses (derived from a dataset containing longer epochs). Both types of trials were discarded, resulting in individually single-trial artifact-cleaned datasets (950 ± 13 trials; mean $\pm$ SD across participants). Trials affected by eye movements, blinks, or other repetitive artifacts were not discarded as these artifacts could be reduced, as follows.

Independent component analysis. To further facilitate the extraction of EEG activity resembling brain activity, the precleaned datasets were decomposed into linear sums of 65 spatially fixed and maximally temporally independent components (ICs) using the extended Infomax ICA algorithm (Bell and Sejnowski, 1995; Lee et al., 2000). These ICs can be considered as separate spatially filtered portions of the scalp EEG activities. Brain-related ICs were separated from artifact-related ICs using standard criteria (Jung et al., 2000; Delorme et al., 2007). ICs primarily accounting for eye movements or blinks were identified based on their far-frontal scalp distributions and irregular occurrence/timing across trials. Remaining artifacts were identified based on their nondipolar scalp maps, flat activity spectra, and irregular occurrence/timing across trials. Removal of these ICs resulted in individually pruned artifact-reduced datasets containing 23 ± 4 ICs (mean $\pm$ SD across participants).

To estimate auditory cortical activities and enable group analyses, the 318 artifact-reduced ICs from all participants were clustered based on similarities in their scalp topographies, event-related activity waveforms, and power spectra (for similar applications, see Jung et al., 2001; Makeig et al., 2002, 2004; Contreras-Vidal and Kerick, 2004; Onton et al., 2006). This was done automatically using principal component analysis and the *k*-means algorithm (Matlab Signal Processing Toolbox). The clustering was further refined based on visual inspection to improve uniformity among the clustered ICs that showed strong auditory event-locking. The resulting cluster contained 13 ± 3 ICs (mean $\pm$ SD across participants) and revealed clear auditory event-locking (Fig. 3), which was deemed most indicative of auditory cortical activities.

Time–frequency analysis. Short-term power spectra were computed for each trial by sliding a Gaussian-tapered window (duration: 1100 ms) in 20 ms steps across the IC activity waveforms. The number of cycles per wavelet was increased linearly from 3 to 17.5 across the investigated frequency range (3–35 Hz). For exploring lower frequencies (1–7 Hz; see Fig. 7*B*), 1 to 3.5 cycles per wavelet were used. Event-related spectra were computed by trial-averaging the short-term spectra and log-subtracting the average spectrum within the 500 ms interval that immediately preceded the noise burst (baseline). These computations were done separately for each participant, IC, and condition. The trial averages included 190 ± 4, 41 ± 28, 52 ± 31, and 92 ± 5 trials (mean $\pm$ SD across participants) for each of the conditions illustrated in Figure 4 (*A*; *B*, left; *B*, right; and *C*, respectively). For the ambiguous stimulus conditions (Fig. 4*B*), the rating conditions were first balanced separately for each participant by selecting a random subset of trials from the rating condition that occurred more frequently.

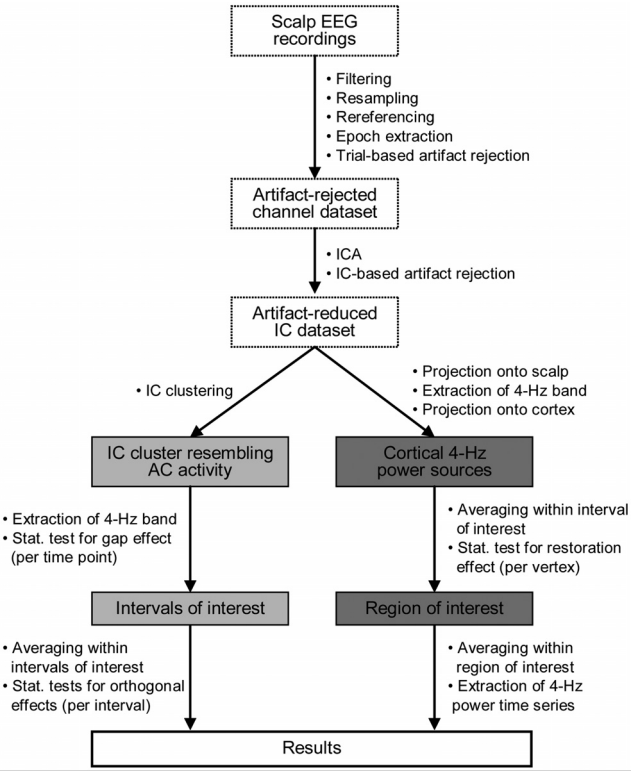

Figure 2. Summary of EEG data preprocessing and analysis steps. After EEG data preprocessing (dotted rectangles), a cluster of maximally temporally independent EEG components deemed to resemble AC activity was extracted (light gray rectangles). Intervals showing gap effects were identified and deemed to resemble neural on/off responses. These responses were further investigated for orthogonal effects related to restoration, restored sound, and context. To verify whether restoration influenced activity in AC, distributed source analyses were conducted (dark gray rectangles). These cortically constrained analyses were based on the whole set of artifact-reduced components and served to localize the restoration-related changes that were identified in the cluster-based analyses (see Results).

Analysis of EEG power and activity. Event-related cluster features (power spectra and activity waveforms) were computed by averaging across the ICs belonging to the cluster of interest. Slow phase-locked activity (Figs. 3, 5) was defined by low-pass filtering the event-related cluster activity waveforms (cutoff: 10 Hz, FIR filter). The cluster features were analyzed on the basis of individual time points or within specific intervals of interest (time averages), and at specific frequencies (for power measures). All measures were extracted separately for each participant and for each condition. Time–frequency windows of interest remained fixed throughout. The extracted measures were submitted to repeated-measures analyses.

Comparison of vowel illusions and tone illusions. To compare continuity illusions of different sounds, data from a previous tone continuity illusion study (Riecke et al., 2009b) that used essentially the same methods as here were analyzed together with the present data. First, the two groups were balanced ($N = 10$ each) by excluding data from those participants in the present study whose dichotomized ratings showed the least inter-

trial variability in the ambiguous conditions (participants 4, 5, 13, and 14; Fig. 4*B*). Second, each rating condition was pooled across the two ambiguous stimulus conditions. Third, effects of auditory restoration on event-related power were defined by subtracting the Restoration condition from the No-restoration condition. Finally, the resulting event-related power differences were standardized using *z*-transforms. These computations were done separately for each participant and each group. The resulting individual *z*-scores were submitted to independent-measures analyses.

Analysis of EEG sources. Two limitations of the present IC-based approach are its dependency on partially subjective clustering criteria and its low spatial specificity. To test more objectively whether restoration-related effects occurred in AC, the nonclustered IC data were submitted to a distributed EEG source analysis. This analysis was based on a cortically constrained inverse model using the minimum L2-norm approach (Dale and Sereno, 1993; Lin et al., 2004, 2006a,b; Esposito et al., 2009a,b) as implemented in BrainVoyager QX (Brain Innovation). For each participant, the nonclustered IC activity waveforms were projected collectively back onto the scalp electrode locations. The resulting channel activities were filtered using a continuous wavelet transform (CWT) such that the frequency of interest (4 Hz; see Results, below) was passed (center frequency: 4 Hz; wavelet ratio: 5). The individually digitized electrode locations were fitted to a common scalp model reconstructed from T1-weighted magnetic resonance images of the Montreal Neurological Institute 152-template head. The CWT-filtered channel activities were then projected onto a cortical surface mesh that was extracted from the aforementioned template brain images. The projection involved placing a single-current dipole at each of the 2500 vertices of the mesh and estimating a distributed solution for the EEG inverse problem using a standard linear model (Mosher et al., 1999; Baillet et al., 2001) at each vertex.

The lead fields for the EEG channel configuration were computed assuming a four-layer sphere model (Berg and Scherg, 1994). The inverse solution was estimated individually based on a depth-weighted minimum-norm solution (Dale and Sereno, 1993; Dale et al., 2000; Lin et al., 2004, 2006a,b). For regularization, the covariance matrix of the channel noise was estimated from channel activities (duration: 50 s) recorded from a polystyrene head model using the same EEG channel configuration as in the subsequent EEG experiment, and the λ-parameter was set according to Lin et al. (2004; 2006a,b) assuming a source power signal-to-noise ratio of 5. The depth-weighting parameter was set to a value of 2 and the final inverse filter weights were noise-normalized (Dale et al., 2000; Lin et al., 2006a,b). Using this inverse solution, EEG source activity was estimated for each of the three dipole components (*x*, *y*, *z*). Finally, from these dipole component activities, a single source 4 Hz power (root-mean-square) time series was obtained. This was done for each vertex and each trial. Note that this distributed source modeling approach may not allow differentiating between sources within small cortical patches because of the low spatial resolution of EEG, yet it is suitable for resolving correlated and uncorrelated sources in left and right AC, as suggested by results from prior auditory studies using similar approaches (Sorokin et al., 2010; Hämäläinen et al., 2011).

Statistical analyses in source space were focused on the interval that showed a main effect of auditory restoration in the IC cluster analysis (−40 to 330 ms relative to vowel offset; $F_{(1,13)} > 4.7, p < 0.05$; see Results, below). Differences in average source power during this interval versus the baseline interval were assessed at each vertex using *t* statistics. The resulting surface *t* maps were then standardized using *z*-transforms. This was done separately for each participant and each condition. The resulting individual *z*-maps were then submitted to repeated-measures statistical analyses. Regions of interest (ROIs) were defined based on the clusters of contiguous sources that revealed the highest statistical significances. For ROI analyses, regional power time series were extracted and averaged (first across trials and then across the sources within the region), each rating condition was pooled across the two ambiguous stimulus conditions, and the baseline was log-subtracted (see Time–frequency analysis, above). This was done separately for each participant and each rating condition. The resulting individual ROI power estimates were submitted to repeated-measures analyses.

Analysis of continuity ratings. Behavioral data analyses were focused on two measures: trial-averaged continuity ratings and variability of single-trial continuity ratings. Both measures were computed separately for each participant and each condition, and then submitted to repeated-measures analyses. The measure of intertrial variability was taken to quantify changes in continuity perception that were independent of changes in acoustic input (Fig. 4, error bars).

Statistical analyses. None of the distributions of the relevant measures (participants' ratings and time-averaged power) diverged significantly from normality as assessed by Kolmogorov–Smirnov tests. Thus, these measures could be analyzed using parametric statistical tests including *t* tests and two-way ANOVAs for repeated-measures, and *t* tests for independent-measures (for vowel vs tone group comparisons). When noted in the text, probability values were corrected for multiple comparisons across time–frequency points or time points (for time–frequency data or time series data, respectively) such that the false-discovery rate (FDR) was controlled at a level of 0.05. The random-effect *t* map resulting from source analysis was corrected using a spatial cluster criterion (minimum vertex cluster size: 75 mm^2; Fig. 6*B*).

Results

General effects of auditory stimulation and task

Based on known correlates of tone continuity illusions in AC (for review, see Petkov and Sutter, 2011), we focused our initial analyses on auditory cortical processes. To that end, we filtered the measured EEG data in the spatial domain (see Materials and Methods, Independent component analysis, above) and extracted the cortical activities that were evoked by the acoustic input and the behavioral task (Fig. 3). These activities clearly deviated after the onset of the noise compared with the average baseline interval (i.e., the vowel portion before the noise). First, EEG power in the delta and theta range (1–2 and 3–7 Hz, respectively) increased significantly after noise onset and offset (Fig. 3*A*, red). Second, power in the alpha and beta/gamma range (8–12 and 15–35 Hz) decreased significantly ~300 ms after the noise onset and remained low thereafter (Fig. 3*A*, blue). The enhancements in the lower frequency range were most compatible with auditory-evoked cortical processes (Cacace and McFarland, 2003), whereas the latter attenuations in the higher frequency range were more compatible with task-related arousal shifts (Klimesch et al., 1998) and motor response planning (Pfurtscheller, 1981). We focused on the putative auditory cortical processes indexed by the lower-frequency power changes and investigated them under the acoustic/perceptual conditions depicted in Figure 4 (top). We hypothesized that neural activity specifically in the 4 Hz band is critical for vowel restoration because EEG power changes in this band have been associated with auditory restoration (Riecke et al., 2009b; Shahin et al., 2012), sound onsets (Makeig, 1993), and auditory change detection (Cacace and McFarland, 2003). These changes could resemble the major spectral component of auditory-evoked on/off responses, which also have been associated with tone restoration (Petkov et al., 2007).

Neural on/off responses to the missing vowel portion

To identify neural on/off responses to the missing vowel portion, we compared conditions that contained an interrupted vowel versus an intact vowel, both combined with noise (i.e., Gap vs No Gap; Fig. 4*A*, top). Participants' continuity ratings of the vowel revealed that these stimuli were perceived differently, i.e., as clearly discontinuous or clearly continuous ($d' = 2.8 \pm 0.4 > 0$, mean ± SEM across participants; $t_{13} = 7.9$, $p < 10^{-6}$; Fig. 4*A*, bottom).

Inspection of the EEG spectrum revealed that the onset and offset of the noise evoked an increase of power in the lower theta range (Fig. 5*A*, event-related power spectrum in the No-gap condition). More importantly, this noise onset/offset-evoked increase revealed

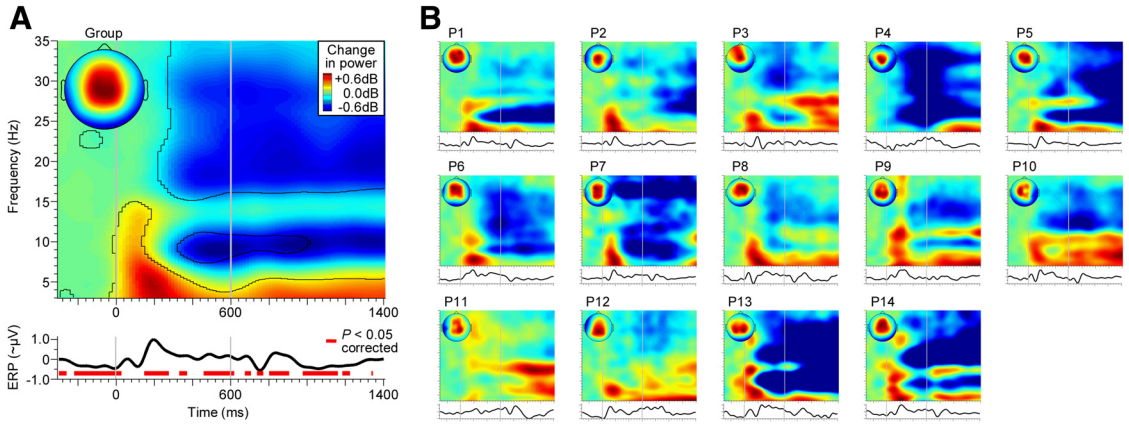

Figure 3. General effects of auditory stimulation and task. ***A***, Auditory cortical processes were defined by extracting activities that showed a centrally distributed topography and time-locking to the acoustic input and behavioral task [scalp map insert, event-related power spectrum, and event-related activity waveform (ERP), averaged first across the stimulus conditions in Fig. 4 and then across participants; for details, see Materials and Methods]. The time–frequency contours and red horizontal bars delineate significant deviations in EEG power and significant deviations in slow phase-locked activity, respectively, relative to the 500 ms interval before the noise (baseline; $p < 0.05$, FDR-corrected). Red and blue time-frequency regions indicate power increases and power decreases, respectively. Vertical lines delineate the noise interval (0 – 600 ms; Fig. 1*A*). ***B***, Same as ***A***, but for individual participants (P) (without statistics; plots normalized with respect to their maximum).

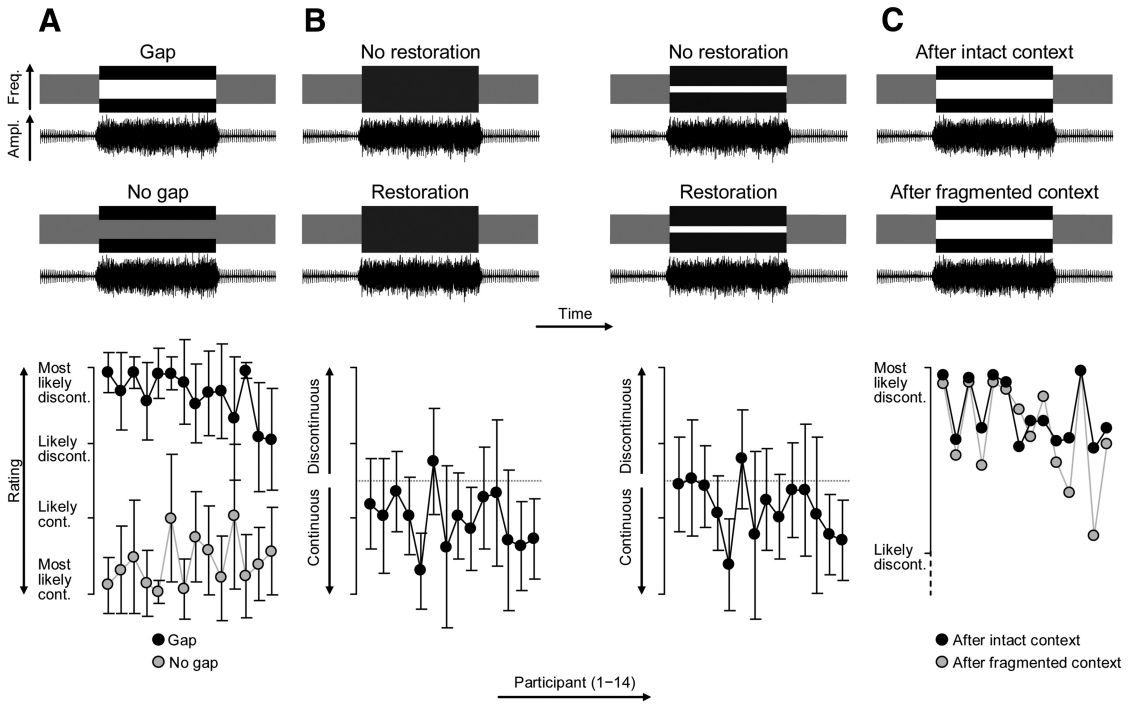

Figure 4. Experimental conditions and behavioral results. ***A***, Top, Schematic stimulus spectrograms and stimulus waveforms; truncated for visibility. Neural on/off responses to the missing vowel portion were identified by comparing a clearly interrupted vowel (Gap condition) and a clearly uninterrupted vowel (No-gap condition). Bottom, Individual continuity ratings (mean ± SD across trials). All participants perceived the vowel as strongly discontinuous or continuous in these conditions. ***B***, Neural suppression related to the restoration of the vowel was investigated using two pairs of physically matched stimulus conditions. For each pair, participants had rated the interrupted vowel as either discontinuous (No-restoration condition) or continuous (Restoration condition) through noise. The perceptual ambiguity of these conditions is illustrated by both the length of the error bars (intertrial variability) and the overlap of these bars with the midpoint of the rating scale (dotted line). ***C***, Context-specific adaptation was investigated by comparing the clearly interrupted vowel (same as in Gap condition, see ***A***) across conditions in which the preceding auditory context was either intact or fragmented (Fig. 8). Participants' ratings were typically correct (i.e., strongly discontinuous), but following fragmented context, these ratings were significantly less decisive (i.e., less likely discontinuous, see ratings; error bars omitted for visibility).

condition-specific differences that paralleled the aforementioned acoustic/perceptual differences due to the vowel gap. In particular, 4 Hz power was significantly higher when the vowel contained a gap versus no gap. This gap-specific enhancement emerged 40 ms after the vowel offset and lasted for 240 ms. It emerged again 80 ms after the vowel reonset, where it lasted for another 330 ms ($t_{13} = 2.2$, $p < 0.05$, FDR-corrected; Fig. 5*A*). The latter enhancement was further significantly stronger than the former ($t_{13} = 2.9$, $p < 0.02$). The timing of these enhancements was consistent with that reported by EEG studies on gap detection in noise (Michalewski et al., 2005; Pratt et al., 2005; Lister et al., 2007), tones (Heinrich et al., 2004), or noise-interrupted tones (Riecke et al., 2009b). They further matched previously observed increases in 4 Hz power to sound onsets and offsets (Makeig, 1993; Riecke et al., 2009b). Therefore, these enhancements

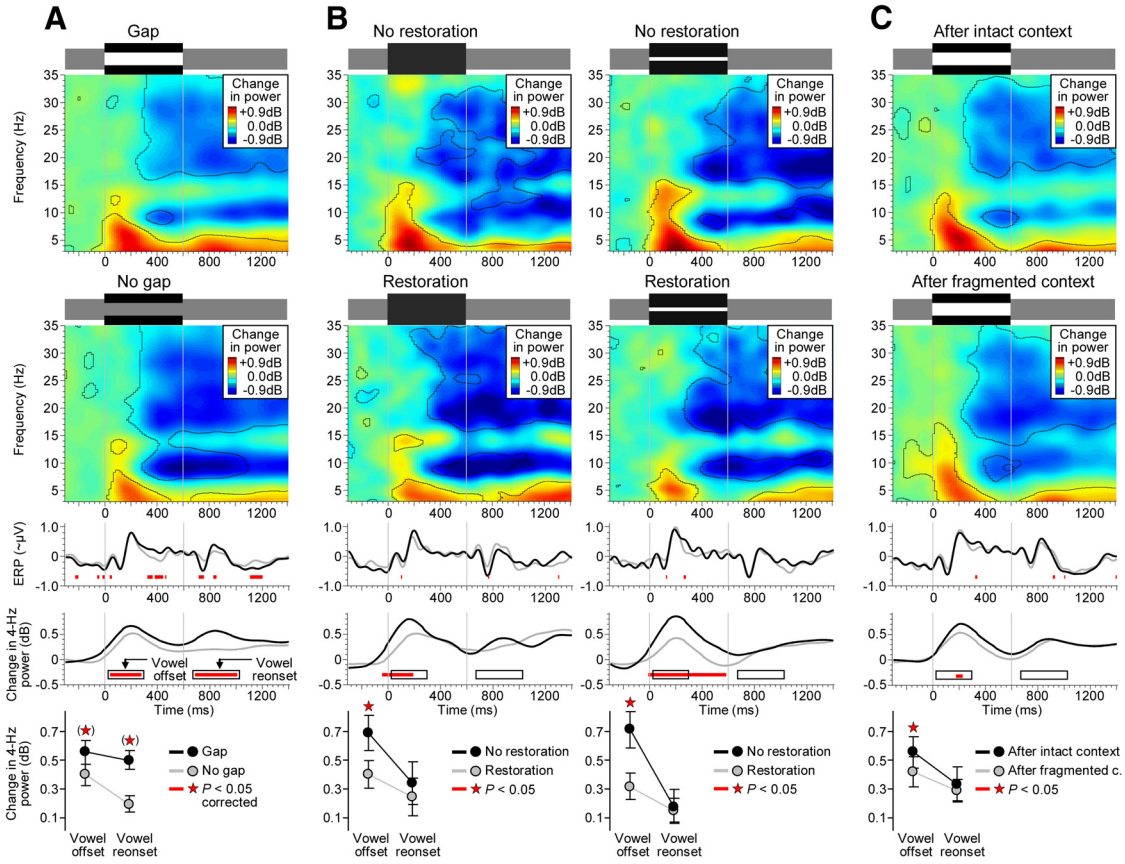

Figure 5. Effects of the missing vowel portion, auditory restoration, and auditory context. ***A***, Both the onset and offset of the noise evoked power increases in the lower theta range (second row: event-related power spectrum in No-gap condition). The vowel offset and reonset in the Gap condition further evoked increases in 4 Hz power (fourth row: event-related 4 Hz power series). The intervals during which these on/off responses reached significance (40–280 ms and 80–410 ms postevent; denoted by the rectangular outlines) thus provided the auditory system with sensory evidence for the offset and reonset of the vowel. ***B***, Auditory restoration of the vowel was accompanied by partial suppression of 4 Hz power. This suppression was observed consistently for each pair of physically matched conditions. It reached significance during the previously identified off-response interval (fifth row: time-averaged 4 Hz power, mean ± SEM across 14 participants). ***C***, After fragmented auditory context, 4 Hz power was partially suppressed compared with intact context. As for restoration, this context-specific suppression reached significance during the off-response interval (fifth row: time-averaged 4 Hz power). For details, see Figure 8. The time–frequency contours delineate significant changes in EEG power relative to baseline, and the red horizontal bars and red asterisks denote significant differences between conditions ($p < 0.05$, FDR-corrected in ***A***). Third row is analogous to the fourth row, except that slow phase-locked activity is plotted and no FDR correction was applied.

in 4 Hz power were most compatible with evoked on/off responses or, more generally, with sensory evidence of the offset/reonset of the vowel.

Neural suppression related to auditory restoration

Is suppression of neural on/off responses, as identified in the previous section, a prerequisite for the restoration of an interrupted vowel? To address this question, we compared an individually defined stimulus condition under which participants had rated the same interrupted vowel as discontinuous versus continuous (i.e., No-restoration vs Restoration). We did this comparison twice, for two slightly different stimuli (see Materials and Methods, above; Fig. 4*B*, top). Participants' continuity ratings revealed significantly more intertrial variability and also significantly more switches (continuity ↔ discontinuity) for these two stimuli than for all the other stimuli (all $t_{13} > 1.9$, $p < 0.05$; Fig. 4*B*, bottom). Therefore, the continuity of these two stimuli was perceived as most ambiguous.

Inspection of the EEG spectrum (Fig. 5*B*) revealed that restorations of the vowel (inferred from participants' reports of continuity) were accompanied by partial suppression of the on/off responses identified in the previous section. In particular, 4 Hz power appeared lower when the vowel portion was restored versus not restored. This attenuation was observed consistently under both ambiguous stimulus conditions (in the following, these conditions were pooled in order to increase sample size). The attenuation emerged shortly before the vowel offset and was most evident during the previously identified interval of the vowel offset response (40–280 ms; $t_{13} = 3.7$, $p < 0.002$). It did not reach significance during the interval of the vowel reonset response (80–410 ms). It was further significantly stronger during the former interval than the latter interval ($t_{13} = 2.5$, $p < 0.03$). Therefore, restoration of the vowel was accompanied by partial suppression of the neural response evoked by the vowel offset.

Source analysis based on the spatially unfiltered EEG data (see Materials and Methods, above) confirmed that the offset suppression was most significant in the right temporal cortex and the parietal cortices ($F_{(1,13)} = 11.0$, $p < 0.005$; Fig. 6*A*). The peak in temporal cortex was situated in a posterior portion of the superior temporal gyrus (STG; Talairach coordinates in mm: $x = 57$, $y = -37$, $z = 16$; Fig. 6*B*, right). This region was in the vicinity of the region in right middle STG (interpeak distance in Talairach space: 13 mm) that showed a corresponding effect in a tone continuity illusion study using essentially the same methods as here (Riecke et al., 2009b). Therefore, the sound offset suppression accompanying the restored sounds was located

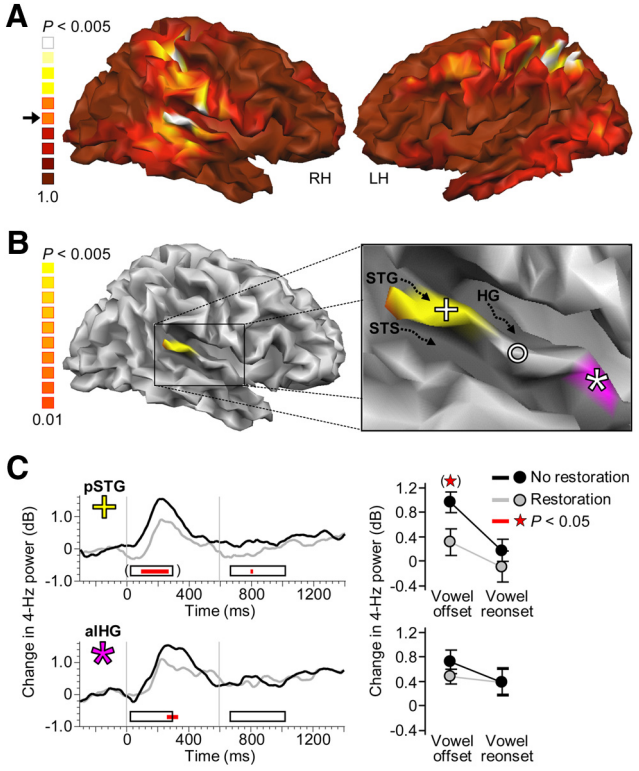

Figure 6. Cortical origin of the restoration-related suppression. ***A***, Auditory restoration was accompanied by partial suppression of 4 Hz power, specifically in the right temporal cortex and the parietal cortices. Brighter spots in the nonthresholded maps indicate EEG sources exhibiting stronger suppression. Arrow, $p < 0.05$; LH, left hemisphere; RH, right hemisphere. ***B***, The most significant suppression in temporal cortex was localized in a posterior portion of STG (cross; $p < 0.005$, excluding vertex clusters <75 mm^2). This region was in the vicinity of a region that showed a corresponding effect for the restoration of a simple tone ($p < 0.001$; circle) in an accompanying EEG study (Riecke et al., 2009b). The red patch delineates the alHG that showed a corresponding effect for restored vowels ($p < 0.05$) in a similar fMRI study (Riecke et al., 2011a). STS, Superior temporal sulcus; HG, Heschl's gyrus. ***C***, Suppression in posterior portion of STG (pSTG) affected the vowel offset response and also the vowel reonset response as similar trends could be observed during the intervals of these responses (top, event-related 4 Hz power series and time-averaged 4 Hz power). Suppression in alHG revealed a similar and less prominent pattern (bottom). Error bars indicate SEM across participants.

consistently in the superior temporal gyrus of the right AC. The offset suppression could also be observed in the anterolateral portion of Heschl's gyrus (alHG) (Fig. 6*B*,*C*, bottom), a region that has been associated with vowel continuity illusions in an fMRI study using similar methods as here (Riecke et al., 2011a).

Neural suppression related to restorations of vowel versus tone

Does neural suppression that is associated with auditory restoration depend on which sound is restored? To address this question, we compared restorations of the vowel with restorations of a tone. To that end, data from the accompanying tone continuity illusion study (Riecke et al., 2009b) were matched and analyzed together with the present data (see Materials and Methods, above). We compared the strength of neural suppression, defined as the average difference in 4 Hz power between the Restoration and No-restoration conditions (Fig. 4*B*), across the two types of restored sounds.

The time course of suppression was overall similar for the restored vowel and the restored tone, including prominent peaks that were time-locked to the offsets and reonsets of these sounds

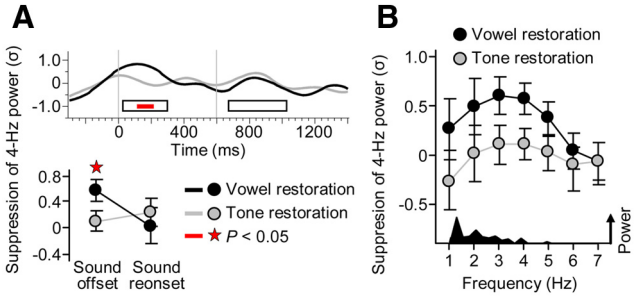

Figure 7. Effects of the restored sound on neural suppression. ***A***, The waveforms show the evolution of restoration-related suppression, defined as the difference in 4 Hz power between the Restoration and No-restoration conditions. The suppression that accompanied the restored vowel revealed peaks during the intervals of the offset/onset responses in Figure 5*A* (black waveform). A similar pattern was observed for the restoration of a simple tone (gray waveform). Offset suppression was significantly stronger for the restored vowel than the restored tone ($p < 0.05$). ***B***, The vowel offset suppression (black circles) was reflected most strongly by activities within the lower theta range (3–4 Hz). A similar pattern was observed for the restored tone (gray circles). The offset suppression was reflected less clearly at the peak frequency of the vowel envelope spectrum (1.3 Hz; see amplitude modulation spectrum, bottom) indicating little suppression of sound envelope-locked neural activity. Error bars indicate SEM across 10 participants.

(Fig. 7*A*, top). However, for the restored vowel, the offset-locked peak was larger and occurred slightly later. The strengthening of suppression reached significance during the previously identified interval of the offset response ($t_{18} = 2.1$, $p < 0.05$; Fig. 7*A*, bottom), whereas no effect was observed during the interval of the reonset response or for the peak latencies. Therefore, the restoration of the more complex sound (i.e., the vowel) was associated with stronger offset suppression.

Exploration of further EEG frequencies within the delta/theta range revealed that the offset suppression affected mostly frequencies in the lower theta range (3–4 Hz) for both restored sounds (Fig. 7*B*). Supplemental analyses for effects of vowel restoration on other EEG frequency bands (alpha, beta, and gamma range) revealed no significant result.

Context-specific adaptation

Do neural on/off responses associated with auditory restoration, as identified in the previous sections, adapt to auditory context? To address this question, we compared conditions comprising the same probe stimulus (same as Gap condition in Fig. 4*A*, top) that was preceded either by intact or fragmented auditory context (Fig. 8). Participants generally deemed the probe correctly as discontinuous (Fig. 4*C*). Their continuity ratings of the probe further depended on the context: following fragmented context versus intact context, participants were significantly less confident about the discontinuity of the interrupted vowel ($t_{13} = 1.8$, $p < 0.05$), in line with previous results based on tones (Riecke et al., 2009a, 2011b).

Inspection of the EEG spectrum revealed that the contrastive aftereffect on participants' judgment was paralleled by a corresponding aftereffect on the neural activity that was evoked by the probe (Fig. 5*C*). In particular, 4 Hz power was lower following fragmented versus intact context. This suppressive aftereffect was evident during the interval of the vowel offset response (40–280 ms; $t_{13} = 1.8$, $p < 0.05$) and a similar, nonsignificant trend could be observed during the interval of the vowel reonset response (80–410 ms). This pattern was similar to that of the restoration-related suppression (compare to Fig. 5*B*). Exploration of the whole trial interval revealed that the suppression following frag-

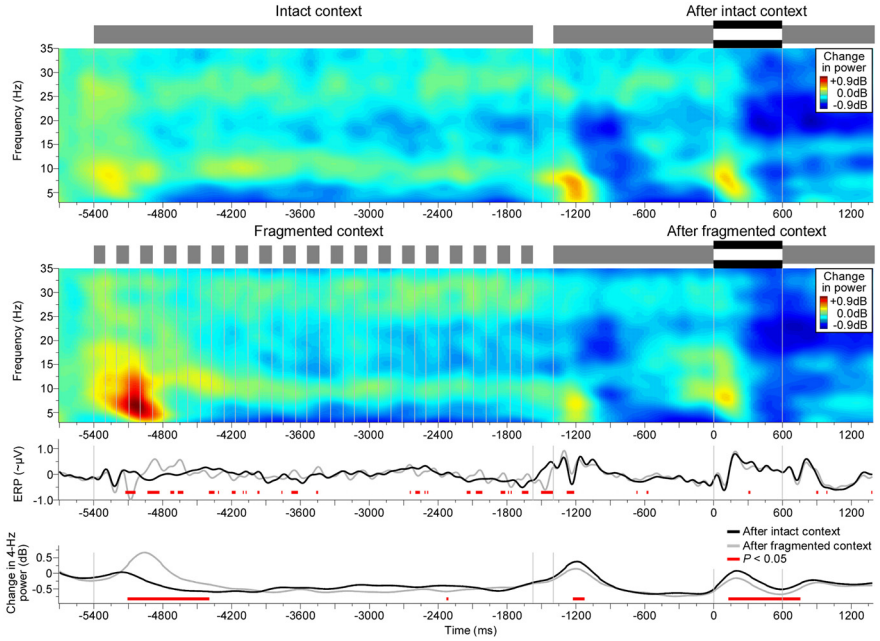

Figure 8. Effects of auditory context. Same as first through fourth rows in Figure 5*C*, but plotted across a longer (whole-trial) interval and using the 500 ms interval preceding the contextual stimulus as baseline. Four Hertz power showed context-specific differences, i.e., differences that reflected the gaps in the fragmented context (fourth row). An initial increase was observed shortly after the onset of the fragmented context; this increase decayed over the remainder of the contextual stimulus. Then, during the subsequent probe, a decrease was observed shortly after the probe onset and also shortly after the probe offset at the interruption ($p < 0.05$; red horizontal bars). Thus the fragmented context exhibited suppressive aftereffects that were time-locked to sound onsets/offsets. Vertical lines denote sound onsets and offsets.

mented context was time-locked exclusively to sound onsets/offsets (Fig. 8). Therefore, the fragmented auditory context induced a temporary suppression of specifically on/off responses.

Evoked and induced effects

The observed effects on 4 Hz power reflected changes in brain activity that were evoked by sensory input and/or induced by ongoing brain processes. To disentangle these two possible sources, we visually compared the results based on 4 Hz power (Fig. 5, fourth row) to results from corresponding analyses of slow phase-locked activity (ERP; Fig. 5, third row); the latter measure does not reflect induced activity (Tallon-Baudry and Bertrand, 1999).

The missing vowel portion had a similar effect on both measures, mostly around the vowel reonset. As shown in Figure 5*A*, the gap-related effects on 4 Hz power co-occurred with corresponding phase-locked effects, which confirms that these effects were evoked by sensory input, i.e., the offset/reonset of the vowel. As described in the previous sections, these evoked on/off responses were influenced by restoration and context. Since these influences concerned mostly 4 Hz power rather than phase-locked activity (Fig. 5*B*,*C*), they were induced by nonsensory factors. In summary, the gap-related increases in 4 Hz power likely resembled auditory-evoked on/off responses (or, more exactly, spectral "signatures" thereof), whereas the restoration- and context-related effects resembled modulations of these on/off responses induced by ongoing brain processes.

Discussion

Our findings provide neurobehavioral evidence for three statements: (1) restoration of an interrupted vowel is accompanied by suppression of off-responses in right AC (indexed by sound offset-evoked 4 Hz power); (2) this suppression is stronger for an illusory vowel than an illusory tone; (3) auditory context influences the off-responses and thereby continuity perception. These statements are elaborated below.

Auditory restoration is accompanied by offset suppression in right AC

Our finding that off-responses in right AC were partially suppressed when listeners experienced continuity illusions of the vowel (as opposed to veridical discontinuity percepts) is consistent with previous results on continuity illusions (Petkov et al., 2007; Riecke et al., 2009b; Shahin et al., 2012). Importantly, our findings extend the no-discontinuity rule for auditory restoration to vowels and reveal the underlying substrate (right AC) and interval (40–280 ms after vowel offset). Since the offset suppression occurred variably across physically identical conditions, it was induced by nonacoustic factors that cannot be resolved here. A possible factor was top-down processing, which has been proposed to initiate the sensory repair of ambiguities in speech signals (Samuel, 1981; Shahin et al., 2009). Perhaps ongoing feedback signals counteracted the synchronization of auditory neurons that were excited by the interruption (Pfurtscheller and Lopes da Silva, 1999). The restoration-related differences shortly before the interruption suggest that such processes could have operated in anticipatory fashion (Winkler et al., 2012), but this idea remains uncertain because of the limited temporal resolution of our 4 Hz power measure. In summary, suppression of evoked activity in right AC may impair the neural evidence for sound offsets and thereby bias the auditory system toward interpreting interrupted sounds as continuous during noise.

Auditory restoration of complex sound features

Our observation that vowel restoration and tone restoration are accompanied by similar neural patterns may hint to a common mechanism. However, we also observed differences associated with the different illusory sounds. First, offset suppression was stronger for the restored vowel. This indicates that the auditory system may put additional efforts in overcoming the breakdown of complex vocal pitch. Our interpretation is supported by psychophysical results (Darwin, 2005) showing that restoration of harmonic sounds occurs on the basis of complex sound features (high-order acoustic properties including pitch) rather than individual harmonics. Second, for the restored vowel, offset suppression occurred slightly later and farther apart from Heschl's gyrus, the most likely location of human primary AC (Rademacher et al., 2001). These differences could hint to vowel-specific processes in nonprimary AC. Our methods cannot disentangle primary and nonprimary AC regions, but several fMRI studies have associated the latter regions with the restoration of complex sounds (Heinrich et al., 2008, 2011; Shahin et al., 2009). For example, a study using the same illusory vocal sound as here reported suppression of hemodynamic activity in a right-

lateralized, presumably nonprimary, AC region on alHG (Riecke et al., 2011a) that may be sensitive to vocal pitch (Lattner et al., 2005). Consistently, this region revealed offset suppression associated with vowel restoration in our study.

The observed right-lateralization can be reconciled with the left-lateralized results of several imaging studies on speech perception (cited in Leff et al., 2009) based on methodological differences. For example, our listeners judged the continuity of a noise-interrupted vowel that conveyed strong pitch, whereas listeners in most of the previous studies judged the intelligibility of uninterrupted speech signals conveying relatively rich semantic/syntactic information. Such stimuli and tasks may activate differently lateralized mechanisms; specifically, pitch processing has been associated with the right AC, whereas the processing of meaning and rapid spectral changes, which are inherent in speech, has been associated more with the left AC (for review, see Zatorre and Gandour, 2008).

Following these considerations, the offset suppression in our study possibly enabled the preservation of vocal pitch through the noise, as suggested by previous psychophysical work (Bregman and Dannenbring, 1977). Vocal pitch may improve the intelligibility of fragmented speech (Başkent and Chatterjee, 2010); thus, restorations of fragmented speech or other complex sounds (e.g., melodies) may also benefit from suppression of offset-evoked activity (Shahin et al., 2012). In this view, the extended no-discontinuity rule may further include that there exists no neural evidence of the offset of the complex sound features that belong to the target sound (Chait et al., 2008; Shinn-Cunningham and Wang, 2008).

Onset/offset asymmetry of suppression

Our finding that the neural response to the vowel reonset was stronger than that to the vowel offset is consistent with neurobehavioral studies showing that sound onsets have a more elaborate neural representation and receive a greater perceptual weighting than sound offsets (for review, see Phillips et al., 2002). Interestingly, the restoration-related suppression found here affected the vowel offset response rather than the vowel reonset response. This opposite asymmetry suggests that the suppression of sound offsets is most vital for auditory continuity perception. Indeed, psychophysical studies have emphasized that the auditory system may use sound offsets for resetting itself at points of sudden intensity change to carry out a new pitch analysis (Bregman et al., 1994; White and Plack, 1998; Plack and White, 2000). Our findings suggest that the auditory system may suppress these resets to smooth ongoing vocal pitch contours.

Another possibility is that the suppression affected the noise onset response. If that were the case, it seems likely that the noise onset would have been attenuated, leading to a comparatively salient vowel offset and thus a rather discontinuous vowel percept (see Introduction, above), and an inhibition of an attention shift to the noise. However, these predictions are incompatible with our psychophysical results showing a link between suppression and continuous vowel percepts, and psychophysical and neuroimaging studies reporting no link between attention and continuity illusions (Drake and McAdams, 1999; Micheyl et al., 2003; Heinrich et al., 2011). Therefore, suppression more likely affected the vowel offset response (for a similar discussion, see Lütkenhöner et al., 2011). Overall, our findings indicate that the auditory system may engage in restoring an interrupted vocal sound early during the interruption. This implicates that restoration does not need to be determined retrospectively after the interruption (Bregman, 1990; Bendixen et al., 2009).

Neural off-responses adapt continuity perception to auditory context

We found that the neural response to a vowel interruption may be suppressed after fragmented context; a result that is indicative of context-specific adaptation. The neuronal circuitry underlying this adaptation cannot be resolved here, but similarly long-lasting ($>$1 s) adaptation to simple tone properties has been demonstrated for AC neurons in animals (Ulanovsky et al., 2004; Bartlett and Wang, 2005). Perhaps sound offset-sensitive AC neurons reduced their gain during their repetitive excitation by fragmented input in our study. This putative gain reduction (Rabinowitz et al., 2011) could have impaired the neural evidence of the vowel offset and biased the auditory system toward interpreting the vowel as continuous. These ideas are supported by our psychophysical results showing that listeners were less decisive about the discontinuity of the interrupted sound when that sound followed fragmented context.

The observed context-specific adaptation complements EEG results linking tone continuity illusions to mismatch detection (Micheyl et al., 2003). Auditory mismatch detection is thought to arise from preattentive mechanisms in AC that compare current sounds to repetitive aspects of auditory context (for review, see Näätänen et al., 2001). Thus, adaptation of the vowel offset response to fragmented context, as observed here, could reflect absence of mismatch, i.e., a highly expected vowel offset (Nelken and Ulanovsky, 2007). Interestingly, context-specific adaptation and auditory restoration both involved suppression of vowel offset-evoked activity, which may hint to a common mechanism. In this view, the auditory system may continuously track the ongoing acoustic input to adapt accordingly the encoding of pitch boundaries and thereby the restoration of vocal sounds in noise.

Adaptive encoding of spectrotemporal boundaries

Our findings on vowel gap detection and auditory restoration suggest that auditory continuity perception depends on the salience of auditory boundaries. In a computational model for auditory scene analysis (Wang and Brown, 2006), this parameter could be manipulated, for example, by sliding a spectrotemporal integration window across the time–frequency representation of an input signal, with larger windows producing lower contrast, i.e., smoother and less salient transitions (Husain et al., 2005). Following our discussion of context-specific adaptation, the size of such integration window may be scaled flexibly based on the spectrotemporal contrast of the ongoing signal. This proposed input-driven adaptation can be incorporated in existing computational models for the continuity illusion (Cooke and Brown, 1993; Masuda-Katsuse and Kawahara, 1999; Grossberg et al., 2004; Husain et al., 2005; Vinnik et al., 2010) or acoustic edge detection (Fishbach et al., 2001). These augmented models may then be applied, for example, in hearing aids to adapt optimally the time constants of front-end dynamic-range compression. This could be useful as these time constants may determine envelope discontinuities in noise-interrupted speech signals, a factor that is critical for the restoration of these signals in hearing aid users (Başkent et al., 2009).

Notes

Supplemental material for this article is available at http://www.maastrichtuniversity.nl/web/Faculties/PsychologyAndNeuroscience/LR_SM_JN2012.htm. The Supplemental Information provides details on the ICA-based approach and results from analyses of delta, theta, alpha, beta, and gamma band power. This material has not been peer reviewed.

References

Assmann PF, Summerfield Q (1990) Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. J Acoust Soc Am 88:680–697.

Baillet S, Riera JJ, Marin G, Mangin JF, Aubert J, Garnero L (2001) Evaluation of inverse methods and head models for EEG source localization using a human skull phantom. Phys Med Biol 46:77–96.

Bartlett EL, Wang X (2005) Long-lasting modulation by stimulus context in primate auditory cortex. J Neurophysiol 94:83–104.

Başkent D, Chatterjee M (2010) Recognition of temporally interrupted and spectrally degraded sentences with additional unprocessed low-frequency speech. Hear Res 270:127–133.

Başkent D, Eiler C, Edwards B (2009) Effects of envelope discontinuities on perceptual restoration of amplitude-compressed speech. J Acoust Soc Am 125:3995–4005.

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1129–1159.

Bendixen A, Schröger E, Winkler I (2009) I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. J Neurosci 29:8447–8451.

Berg P, Scherg M (1994) A fast method for forward computation of multiple-shell spherical head models. Electroencephalogr Clin Neurophysiol 90:58–64.

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge: MIT.

Bregman AS, Dannenbring GL (1977) Auditory continuity and amplitude edges. Can J Psychol 31:151–159.

Bregman AS, Ahad PA, Kim J (1994) Resetting the pitch-analysis system. 2. Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones. J Acoust Soc Am 96:2694–2703.

Cacace AT, McFarland DJ (2003) Spectral dynamics of electroencephalographic activity during auditory information processing. Hear Res 176:25–41.

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

Contreras-Vidal JL, Kerick SE (2004) Independent component analysis of dynamic brain responses during visuomotor adaptation. Neuroimage 21:936–945.

Cooke MP, Brown GJ (1993) Computational auditory scene analysis: exploiting principles of perceived continuity. Speech Commun 13:391–399.

Dale AM, Sereno MI (1993) Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci 5:162–176.

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26:55–67.

Darwin CJ (2005) Simultaneous grouping and auditory continuity. Percept Psychophys 67:1384–1390.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Delorme A, Sejnowski T, Makeig S (2007) Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34:1443–1449.

Drake C, McAdams S (1999) The auditory continuity phenomenon: role of temporal sequence structure. J Acoust Soc Am 106:3529–3538.

Esposito F, Mulert C, Goebel R (2009a) Combined distributed source and single-trial EEG-fMRI modeling: application to effortful decision making processes. Neuroimage 47:112–121.

Esposito F, Aragri A, Piccoli T, Tedeschi G, Goebel R, Di Salle F (2009b) Distributed analysis of simultaneous EEG-fMRI time-series: modeling and interpretation issues. Magn Reson Imaging 27:1120–1130.

Fishbach A, Nelken I, Yeshurun Y (2001) Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. J Neurophysiol 85:2303–2323.

Grossberg S, Govindarajan KK, Wyse LL, Cohen MA (2004) ARTSTREAM: a neural network model of auditory scene analysis and source segregation. Neural Netw 17:511–536.

Hämäläinen JA, Ortiz-Mantilla S, Benasich AA (2011) Source localization of event-related potentials to pitch change mapped onto age-appropriate MRIs at 6 months of age. Neuroimage 54:1910–1918.

Heinrich A, Alain C, Schneider BA (2004) Within- and between-channel gap detection in the human auditory cortex. Neuroreport 15:2051–2056.

Heinrich A, Carlyon RP, Davis MH, Johnsrude IS (2008) Illusory vowels resulting from perceptual continuity: a functional magnetic resonance imaging study. J Cogn Neurosci 20:1737–1752.

Heinrich A, Carlyon RP, Davis MH, Johnsrude IS (2011) The continuity illusion does not depend on attentional state: fMRI evidence from illusory vowels. J Cogn Neurosci 23:2675–2689.

Houtgast T (1972) Psychophysical evidence for lateral inhibition in hearing. J Acoust Soc Am 51:1885–1894.

Husain FT, Lozito TP, Ulloa A, Horwitz B (2005) Investigating the neural basis of the auditory continuity illusion. J Cogn Neurosci 17:1275–1292.

Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ (2000) Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37:163–178.

Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ (2001) Analysis and visualization of single-trial event-related potentials. Hum Brain Mapp 14:166–185.

Klimesch W, Doppelmayr M, Russegger H, Pachinger T, Schwaiger J (1998) Induced alpha band power changes in the human EEG and attention. Neurosci Lett 244:73–76.

Lattner S, Meyer ME, Friederici AD (2005) Voice perception: sex, pitch, and the right hemisphere. Hum Brain Mapp 24:11–20.

Lee TW, Girolami M, Bell AJ, Sejnowski TJ (2000) A unifying information-theoretic framework for independent component analysis. Comput Math Appl 39:1–21.

Leff AP, Iverson P, Schofield TM, Kilner JM, Crinion JT, Friston KJ, Price CJ (2009) Vowel-specific mismatch responses in the anterior superior temporal gyrus: an fMRI study. Cortex 45:517–526.

Lin FH, Witzel T, Hämäläinen MS, Dale AM, Belliveau JW, Stufflebeam SM (2004) Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain. Neuroimage 23:582–595.

Lin FH, Belliveau JW, Dale AM, Hamalainen MS (2006a) Distributed current estimates using cortical orientation constraints. Hum Brain Mapp 27:1–13.

Lin FH, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW, Hämäläinen MS (2006b) Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. Neuroimage 31:160–171.

Lister JJ, Maxfield ND, Pitt GJ (2007) Cortical evoked response to gaps in noise: within-channel and across-channel conditions. Ear Hear 28:862–878.

Lütkenhöner B, Seither-Preisler A, Krumbholz K, Patterson RD (2011) Auditory cortex tracks the temporal regularity of sustained noisy sounds. Hear Res 272:85–94.

Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr Clin Neurophysiol 86:283–293.

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ (2002) Dynamic brain sources of visual evoked responses. Science 295:690–694.

Makeig S, Delorme A, Westerfield M, Jung TP, Townsend J, Courchesne E, Sejnowski TJ (2004) Electroencephalographic brain dynamics following manually responded visual targets. PLoS Biol 2:e176.

Masuda-Katsuse I, Kawahara H (1999) Dynamic sound stream formation based on continuity of spectral change. Speech Commun 27:235–259.

Michalewski HJ, Starr A, Nguyen TT, Kong YY, Zeng FG (2005) Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. Clin Neurophysiol 116:669–680.

Micheyl C, Carlyon RP, Shtyrov Y, Hauk O, Dodson T, Pullvermüller F (2003) The neurophysiological basis of the auditory continuity illusion: a mismatch negativity study. J Cogn Neurosci 15:747–758.

Miller GA, Licklider JCR (1950) The intelligibility of interrupted speech. J Acoust Soc Am 22:167–173.

Mosher JC, Leahy RM, Lewis PS (1999) EEG and MEG: forward solutions for inverse methods. IEEE Trans Biomed Eng 46:245–259.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence" in the auditory cortex. Trends Neurosci 24:283–288.

Nelken I, Ulanovsky N (2007) Mismatch negativity and stimulus-specific adaptation in animal models. J Psychophysiol 21:214–223.

Onton J, Westerfield M, Townsend J, Makeig S (2006) Imaging human EEG dynamics using independent component analysis. Neurosci Biobehav Rev 30:808–822.

Petkov CI, Sutter ML (2011) Evolutionary conservation and neuronal mechanisms of auditory perceptual restoration. Hear Res 271:54–65.

Petkov CI, O'Connor KN, Sutter ML (2007) Encoding of illusory continuity in primary auditory cortex. Neuron 54:153–165.

Pfurtscheller G (1981) Central beta rhythm during sensorimotor activities in man. Electroencephalogr Clin Neurophysiol 51:253–264.

Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110:1842–1857.

Phillips DP, Hall SE, Boehnke SE (2002) Central auditory onset responses, and temporal asymmetries in auditory perception. Hear Res 167:192–205.

Plack CJ, White LJ (2000) Perceived continuity and pitch perception. J Acoust Soc Am 108:1162–1169.

Pratt H, Bleich N, Mittelman N (2005) The composite N1 component to gaps in noise. Clin Neurophysiol 116:2648–2663.

Rabinowitz NC, Willmore BD, Schnupp JW, King AJ (2011) Contrast gain control in auditory cortex. Neuron 70:1178–1191.

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage 13:669–683.

Riecke L, Van Orstal AJ, Formisano E (2008) The auditory continuity illusion: a parametric investigation and filter model. Percept Psychophys 70:1–12.

Riecke L, Mendelsohn D, Schreiner C, Formisano E (2009a) The continuity illusion adapts to the auditory scene. Hear Res 247:71–77.

Riecke L, Esposito F, Bonte M, Formisano E (2009b) Hearing illusory sounds in noise: the timing of sensory-perceptual transformations in auditory cortex. Neuron 64:550–561.

Riecke L, Walter A, Sorger B, Formisano E (2011a) Tracking vocal pitch through noise: neural correlates in nonprimary auditory cortex. J Neurosci 31:1479–1488.

Riecke L, Micheyl C, Vanbussel M, Schreiner CS, Mendelsohn D, Formisano E (2011b) Recalibration of the auditory continuity illusion: sensory and decisional effects. Hear Res 277:152–162.

Samuel AG (1981) Phonemic restoration: insights from a new methodology. J Exp Psychol Gen 110:474–494.

Scholl B, Gao X, Wehr M (2010) Nonoverlapping sets of synapses drive on responses and off responses in auditory cortex. Neuron 65:412–421.

Schreiner C (1980) Encoding of alternating acoustical signals in the medial geniculate body of guinea pigs. Hear Res 3:265–278.

Shahin AJ, Bishop CW, Miller LM (2009) Neural mechanisms for illusory filling-in of degraded speech. Neuroimage 44:1133–1143.

Shahin AJ, Kerlin JR, Bhat J, Miller LM (2012) Neural restoration of degraded audiovisual speech. Neuroimage 60:530–538.

Shinn-Cunningham BG, Wang D (2008) Influences of auditory object formation on phonemic restoration. J Acoust Soc Am 123:295–301.

Sorokin A, Alku P, Kujala T (2010) Change and novelty detection in speech and non-speech sound streams. Brain Res 1327:77–90.

Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 3:151–162.

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

van Dommelen WA (1990) Acoustic parameters in human speaker recognition. Lang Speech 33:259–272.

Vinnik E, Itskov P, Balaban E (2010) A proposed neural mechanism underlying auditory continuity illusions. J Acoust Soc Am 128:EL20–EL25.

Wang D, Brown GJ (2006) Computational auditory scene analysis: principles, algorithms, and applications. New York: Wiley-IEEE.

Warren RM (1999) Auditory perception: a new analysis and synthesis. Cambridge, UK: Cambridge UP.

Warren RM, Obusek CJ, Ackroff JM (1972) Auditory induction: perceptual synthesis of absent sounds. Science 176:1149–1151.

White LJ, Plack CJ (1998) Temporal processing of the pitch of complex tones. J Acoust Soc Am 103:2051–2063.

Winkler I, Denham S, Mill R, Bohm TM, Bendixen A (2012) Multistability in auditory stream segregation: a predictive coding view. Philos Trans R Soc Lond B Biol Sci 367:1001–1012.

Zatorre RJ, Gandour JT (2008) Neural specializations for speech and pitch: moving beyond the dichotomies. Philos Trans R Soc Lond B Biol Sci 363:1087–1104.