

Tracking Vocal Pitch through Noise: Neural Correlates in Nonprimary Auditory Cortex

Lars Riecke,^{1,2} Anke Walter,^{1,2,3} Bettina Sorger,^{1,2,4} and Elia Formisano^{1,2}

¹Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience and ²Maastricht Brain Imaging Center, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, The Netherlands, ³Department of Neuropsychology, Faculty of Social and Behavioral Sciences, Tilburg University, 5000 LE Tilburg, The Netherlands, and ⁴Coma Science Group, Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium

In natural environments, a sound can be heard as stable despite the presence of other occasionally louder sounds. For example, when a portion in a voice is replaced by masking noise, the interrupted voice may still appear illusorily continuous. Previous research found that continuity illusions of simple interrupted sounds, such as tones, are accompanied by weaker activity in the primary auditory cortex (PAC) during the interruption than veridical discontinuity percepts of these sounds. Here, we studied whether continuity illusions of more natural and more complex sounds also emerge from this mechanism. We used psychophysics and functional magnetic resonance imaging in humans to measure simultaneously continuity ratings and blood oxygenation level-dependent activity to vowels that were partially replaced by masking noise. Consistent with previous results on tone continuity illusions, we found listeners' reports of more salient vowel continuity illusions associated with weaker activity in auditory cortex (compared with reports of veridical discontinuity percepts of physically identical stimuli). In contrast to the reduced activity to tone continuity illusions in PAC, this reduction was localized in the right anterolateral Heschl's gyrus, a region that corresponds more to the non-PAC. Our findings suggest that the ability to hear differently complex sounds as stable during other louder sounds may be attributable to a common suppressive mechanism that operates at different levels of sound representation in auditory cortex.

Introduction

How does the auditory system enable the tracking of a relevant sound in the presence of other occasionally louder sounds? A striking illustration is the auditory continuity illusion in which an interrupted sound, for example a tone or a voice, can be heard continuing through another sound that masks the interruption (Bregman, 1990; Warren, 1999). Neurophysiologic studies on the mechanisms underlying continuity illusions of tones have stressed a role of the primary auditory cortex (PAC) (Sugita, 1997; Micheyl et al., 2003). More recently, findings from animal electrophysiology and human electroencephalography (EEG) have indicated that continuity illusions of interrupted tones may be accompanied by weaker activity in PAC during the interruption compared with veridical discontinuity percepts of these tones (Petkov et al., 2007; Riecke et al., 2009). However, whether this putative mechanism also operates on sounds more complex than tones could not be resolved. Several non-PAC regions are more sensitive to complex sounds than to simple sounds (Wessinger et al., 2001). Furthermore, voice-sensitive regions in temporal cortex respond more strongly to vocal sounds than to

nonvocal sounds (Belin et al., 2000; Binder et al., 2000). Therefore, continuity illusions of complex sounds, such as voices, might rely on processing stages beyond PAC that cannot be addressed using tones.

Few studies have investigated the neural basis of continuity illusions of complex sounds. Sivonen et al. (2006) used spoken words in which phonemes were replaced by a cough. They varied the likelihood of these words during EEG by manipulating sentence context. Their results supported the view that speech restoration emerges mainly from top-down processes, i.e., listeners' expectations. Heinrich et al. (2008) studied vowel perception using stimuli synthesized of artificial formants. They found that partially masked formants that may elicit continuity illusions evoke more vowel-like percepts and stronger blood oxygenation level-dependent (BOLD) responses in posterior temporal cortex than formants not eliciting such illusions. Shahin et al. (2009) used spoken words in which phonemes were replaced by noise. They found that BOLD responses in posterotemporal and frontoparietal cortices covary with listeners' reports of continuity illusions. Although the latter findings suggest that complex continuity illusions may emerge in temporal regions outside PAC, most studies did not link reports of continuity illusions to concurrent brain activity. Thus, more evidence is needed to confirm the mechanism and the stage of sound representation that enable continuity illusions of complex sounds.

We addressed this issue by combining psychophysics with functional magnetic resonance imaging (fMRI). We filled noise into a silent portion of vowel recordings and manipulated the masking potential of the noise. This enabled us to vary the illu-

Received July 3, 2010; revised Oct. 28, 2010; accepted Nov. 10, 2010.

This work was supported by the Netherlands Organization for Scientific Research Cognitive Programma Grant 05104020. We thank Mieke Vanbussel for support with magnetic resonance image segmentation and Pascal Belin, Daniel Mendelsohn and three anonymous reviewers for useful comments on a previous version of this manuscript.

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.3450-10.2011

Copyright © 2011 the authors 0270-6474/11/311479-10\$15.00/0

sory continuity of the voice, as confirmed by listeners who rated voice continuity during fMRI. We controlled illusion-unrelated aspects of this task using non-illusory stimuli and we factored out acoustic variations during data analysis. We identified a region in auditory cortex (AC) in which BOLD responses covaried with the reported salience of continuity illusions. To infer the stage at which these vocal pitch illusions were represented, we compared this region with a PAC region implicated in tone continuity illusions (Riecke et al., 2007).

Materials and Methods

Participants

Fifteen volunteers (nine females) between 20 and 36 years old (mean, 24 years) with no reported hearing or motor problems participated in the study after providing informed consent. The local ethics committee (Ethische Commissie Psychologie at Maastricht University) approved the study procedures.

Stimuli

Vowel recordings. Voices from two females and two males each uttering a sustained [a:] vowel were recorded and digitized (sampling rate, 44.1 kHz, 16 bit resolution) (Fig. 1A). The fundamental frequency (f_0) of the voices was chosen to play a central role for the continuity illusions in the present study (for details, see section on Design) because this parameter is known to support important aspects of hearing such as vocal pitch, speaker identification (van Dommelen, 1990), and perceptual segregation of concurrent vowels (Culling and Darwin, 1993; Assmann and Paschall, 1998). The f_0 of each voice was defined as the frequency that exhibited the maximum of the long-term power spectrum of the recording (Hess, 1983). The extracted f_0 values clearly differed for the female voices (224 and 231 Hz) and the male voices (100 and 148 Hz), and they were compatible with values from the literature (Hwa Chen, 2007; Schweinberger et al., 2008). Similar f_0 values were obtained when using an approach based on the autocorrelation function (Hess, 1983).

The recordings and digitization were performed in a sound-attenuated chamber using a condenser microphone (ME 64; Sennheiser Electronic), a Sound Blaster sound card (Audigy 2ZS; Creative Technology), and Praat software (Boersma, 2001).

Stimulus generation. The auditory stimuli comprised a voice and a noise burst (Fig. 1A), except for stimuli that comprised only a voice (no-noise conditions; see below, Design). The vowels from the four different speakers were matched for their number of harmonics by bandpass filtering each voice in such a way that the frequencies from four octaves below f_0 to four octaves above f_0 were preserved [finite impulse response (FIR) filter centered on f_0]. The filtered recordings were matched for their duration by removing silent portions before the voice onset and truncating each recording at 2800 ms (Fig. 1A).

The noise bursts had a duration of 600 ms and were obtained by bandpass filtering white noise. This was done separately for the different voices, preserving the frequencies from 4.05 octaves below f_0 to 4.05 octaves above f_0 (FIR filter centered on f_0). The different noise bursts were matched for their overall sound level by adjusting their root mean square (RMS) values. The RMS-matched noise bursts were superimposed on the center of the corresponding voice using a log-frequency scale and a linear timescale.

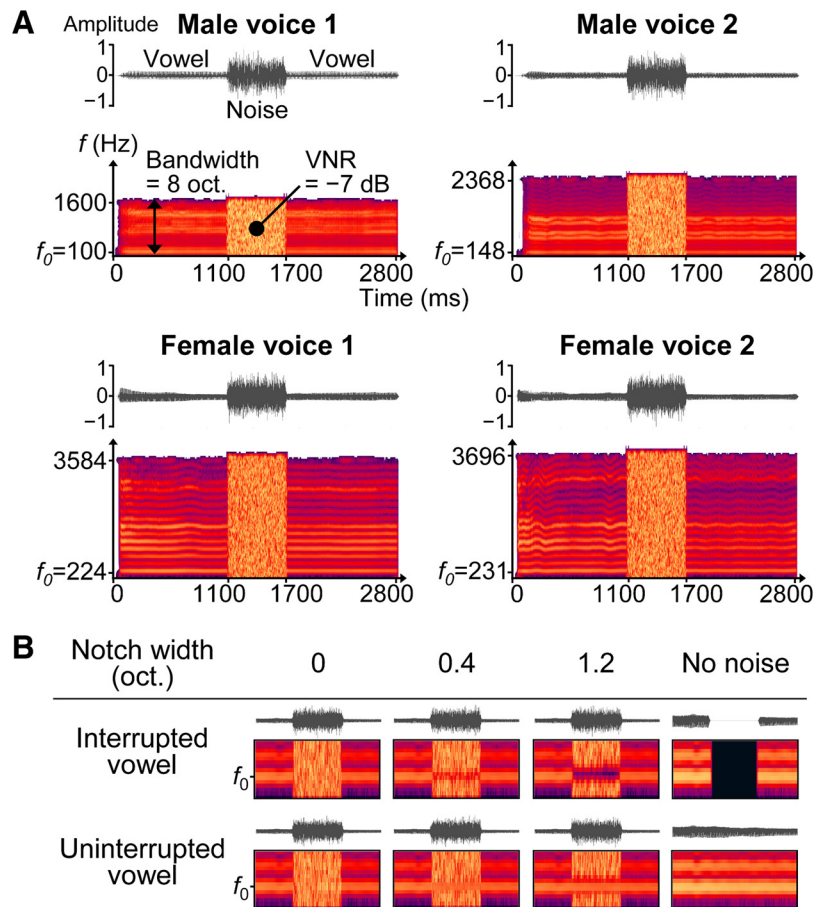


Figure 1. Auditory stimuli and experimental design. **A**, Waveforms and short-term power spectra are shown for four exemplar auditory stimuli. Each stimulus comprised a sustained [a:] vowel and a noise burst that overlapped with the middle portion of the vowel. The vowel was recorded from four speakers with different fundamental frequency f_0 , oct., Octaves. **B**, Waveforms and short-term power spectra (focused on the first 2 harmonics during the middle stimulus portion) are shown for one voice in the different stimulus conditions. The vowel was either interrupted by a silent gap or uninterrupted (control), thus the middle portion was either missing or intact. The amount of acoustic energy in and around f_0 in this portion was manipulated by varying the width of a spectral notch in the overlapping noise (columns 1–3). Additional reference stimuli comprised the vowel without noise (column 4).

For each stimulus, the RMS value of the voice was adjusted so that the voice level was 7 dB below the noise level during the noise interval. This “voice-to-noise ratio” (VNR) ensured that the noise would mask the frequency components of the voice in the middle stimulus portion, which is one determining factor for the continuity illusion (Houtgast, 1972; Warren et al., 1972).

The main stimuli comprised a physically interrupted vowel and were created as described above for the uninterrupted control stimuli, except that a silent gap was inserted in the middle portion (Fig. 1B, top row). These interrupted stimuli served to evoke continuity illusions or veridical discontinuity percepts, depending on the masking of the gap (see below, Design).

The amplitude onsets and offsets of the voices and noise bursts were linearly ramped with 25 ms rise–fall times. For the interrupted vowel, the midpoints of the voice off-ramps coincided with the midpoints of the noise on-ramps and vice versa. All stimuli were sampled at 44.1 kHz using 16 bit resolution and processed in Matlab (MathWorks).

Design

The masking potential of the noise burst in the interrupted vowel was manipulated across three levels so that these main stimuli could evoke continuity illusions and discontinuity percepts of varying salience. More specifically, the amount of noise energy in and around f_0 was manipulated by varying the width of a spectral notch that was centered on f_0 (Fig. 1B, columns 1–3). The first level comprised no notch (notch width, 0

octaves) and was expected to mask the gap, a prerequisite for the continuity illusion (see previous section). The second level comprised a narrow notch (notch width, 0.4 octaves) and was expected to elicit ambiguous percepts of the middle voice portion. This specific notch width was derived from averaging individual continuity illusion thresholds that were obtained from psychophysical pilot experiments ($n = 5$). The third level comprised a broad notch (notch width, 1.2 octaves) and was expected to evoke discontinuity percepts, i.e., a perceptual change of similar magnitude as the zero-notch condition but in the opposite direction (relative to the intermediate notch condition).

The spectral notch was inserted by band-stop filtering the different noise bursts (FIR filter centered on f_0). After filtering, the RMS values of the notched noise bursts were readjusted, so that the overall noise level and the VNR remained constant for all stimuli.

For the stimuli comprising an uninterrupted vowel (Fig. 1*B*, bottom row), the notch parameter was varied as described above for the main stimuli. These uninterrupted control stimuli served to evoke non-illusory continuity percepts, as expected based on previous work using tones (Riecke et al., 2008). Furthermore, inclusion of these non-illusory stimuli enabled controlling influences on neural activity that were not specifically related to continuity illusions (i.e., influences attributable to the notch width or potential response biases that would possibly affect both illusory and non-illusory conditions). To extract influences that were uniquely related to continuity illusions, influences attributable to non-illusory continuity (i.e., as evoked by the uninterrupted vowel) were subtracted from influences attributable to illusory continuity (i.e., as evoked by the interrupted vowel) (for details, see below, Contrast analysis).

The stimuli in the no-noise conditions (Fig. 1*B*, last column) were RMS matched to the stimuli described above and were included to provide listeners with clear references for percepts of veridical continuity and veridical discontinuity, respectively.

Task

Participants were instructed on a screen to attend to the voices and rate the perceived continuity of each voice by pressing a button on an optical response keypad. Listeners performed the ratings on a four-point scale that was labeled with the following: most likely continuous; likely continuous; likely discontinuous; and most likely discontinuous (visible to the listener throughout the experiment). The task was presented in blocks, with each block representing one of the eight conditions shown in Figure 1*B* (for details, see supplemental Fig. S1, available at www.jneurosci.org as supplemental material). Each task block comprised four trials, with each trial comprising a different voice followed by a visual cue to rate the preceding voice. Sixteen such task blocks alternating with rest (stimulus-free baseline) blocks were presented during each experiment (corresponding to one functional run). All blocks had the same duration (20 s). The order of the task blocks and the order of the trials within each block were pseudorandomized. Each listener completed four differently randomized functional runs within a single session; thus, in total, eight blocks (32 stimuli) of each stimulus condition were presented.

The four voices within each task block were from different speakers, which served to induce variability and avoid that listeners would adopt simple response strategies such as maintaining the same rating for all four stimuli. This strategy would be reflected, for example, by rating shifts being larger across the first two stimuli of a block (i.e., when the listener was figuring out which specific rating “should fit” the remaining stimuli in the block) than across the last two stimuli of the same block (i.e., when the listener had already settled on a specific rating). Supplemental analyses based on intrablock variances of listeners’ ratings revealed no evidence for this hypothesis (paired t test, $t_{(12)} = 1.1$, $p = 0.6$), suggesting that listeners unlikely adopted the above strategy.

Procedure

The auditory stimuli were presented during stimulus intervals using MRI-compatible headphones (Commander XG; Resonance Technology). The stimulus intervals were alternated with scanning intervals during which BOLD responses were measured (for details, see supplemental Fig. S1, available at www.jneurosci.org as supplemental material). This

interleaved acquisition design allows presenting auditory stimuli in silence and collecting a relatively large number of samples. Although it induces some overlap of BOLD responses to auditory stimuli and those to intermittent MRI scanner noise, this overlap is significantly smaller than that induced by continuous acquisition designs (Shah et al., 2000; Gaab et al., 2007).

Listeners wore ear plugs and ear muffs that together attenuated MRI scanner noise to a sound pressure level (~ 60 dB) that was likely below that of the interleaved stimuli (~ 70 dB; estimated levels in the ear canal). After listeners had inserted the ear plugs, the loudness of the stimuli was equalized across the two ears. Afterward, listeners practiced the task for 7 min inside the scanner before the first functional run was conducted. The presentation of the auditory stimuli was timed using trigger pulses from the MR scanner and Presentation software (Neurobehavioral Systems).

Imaging

Brain images were collected with a Siemens Allegra 3-Tesla MRI scanner (Siemens Medical Systems). BOLD signal changes were measured with a head coil using a gradient echo planar imaging sequence (time to echo, 30 ms; acquisition time, 2200 ms; repetition time, 5000 ms; field of view, 256×256 mm²; matrix size, 128×128 ; slice thickness, 2 mm; corresponding to a voxel size of $2 \times 2 \times 2$ mm³). During each functional run, 134 volumes were collected, each comprising 27 axial continuous slices (i.e., without interslice gap) centered on the Sylvian fissures and covering the temporal lobes. Structural T1-weighted images optimized for gray-white matter contrast were obtained using a magnetization-prepared rapid-acquisition gradient echo pulse sequence (voxel size, $1 \times 1 \times 1$ mm³).

Preprocessing of functional images included correction for head motion, coregistration with individuals’ structural images, normalization to stereotaxic Talairach space, and spatial smoothing with a 4 mm full-width half-maximum Gaussian kernel. BOLD signal time series were corrected for interslice acquisition time differences, high-pass filtered (cutoff, three cycles per run), and normalized to the mean signal level. Gray-white matter boundaries were extracted from the structural images and used for reconstructing individual cortical surface meshes.

Statistical analyses

Behavioral data analysis. Listeners’ continuity ratings were analyzed using ANOVAs. The notch width, the gap (i.e., the interrupted or uninterrupted vowel), and f_0 were treated as fixed factors, and subjects were included as a random factor. One main aim was to investigate influences of continuity perception that were independent of stimulus changes. To determine the amount of such stimulus-unrelated changes, the amount of interblock variability in listeners’ average ratings was quantified as follows. First, listeners’ average ratings were obtained for each task block by averaging the four ratings within a block. Second, the interblock variability was obtained by computing the variance of listeners’ average ratings across repetitions of the same block; this was done separately for each stimulus condition and for each listener. Finally, the resulting amounts of variance were expressed as percentages of the total variance in all stimulus conditions (see Results) (see Fig. 2*A*).

Functional mapping. To map local differences in activity associated with differences among the experimental conditions, the pre-processed fMRI data were analyzed using a voxel-by-voxel, two-level random-effect analysis. At the first level, a general linear model (GLM) (described in next section) was computed for each subject by fitting the BOLD signal time series with the predicted time series in the different conditions. The sluggishness of the hemodynamic response was taken into account by convolving the predicted time series with a hemodynamic response function (HRF). This HRF was composed of the sum of two gamma functions (Friston et al., 1998) excluding temporal derivatives. At the second level, maps of repeated-measure statistics were obtained by contrasting the fitted model parameters that were derived from the first-level analysis (i.e., individual β values) across specific conditions using paired t tests (for details, see below, Contrast analysis).

Maps of statistically significant differences were obtained by applying a threshold to the repeated-measure contrast maps. This threshold was defined by combining a voxelwise t statistic criterion with a contiguous

voxel (cluster size) criterion, which together yielded a corrected false-positive rate of $p < 0.05$. The cluster size criterion was estimated from the spatial smoothness of a map of sound-related voxels (see below, Contrast analysis) and from 1000 Monte Carlo simulations (Forman et al., 1995; Goebel et al., 2006).

General linear model. The aims of the fMRI analyses were threefold. The first aim was to define AC regions that responded strongly to sound (i.e., to the no-noise stimuli). The main aim was to identify regions within these AC regions that showed effects specifically reflecting continuity illusions in the absence of stimulus changes (as indicated by listeners' continuity reports to the same interrupted stimuli). The final aim was to identify those AC regions that showed effects specifically reflecting acoustic discontinuities (as induced by the notch in the interrupted stimuli).

Therefore, three predictors were included in the GLM, which coded for three different properties of the interrupted stimuli as follows. The first predictor coded for the presentation of the stimuli in the no-noise condition. The second predictor coded for the perceived properties of the three different noise conditions (i.e., the listener's average continuity rating in these conditions). Finally, the third predictor coded for the acoustic properties of these three noise conditions (i.e., the notch width). For the uninterrupted control stimuli, three additional predictors were included that were analogous to the three main predictors described above. All predictors were uncorrelated (for details, see next section).

The predictors that coded for the no-noise conditions were binary and thus they allowed investigating simple "on-off" changes in the BOLD response to the presentation of the auditory stimuli. The predictors that coded for the ratings and the notch widths were both modeled as linear functions, and thus they allowed investigating parametric changes in the BOLD response to different levels of perceived continuity and different notch widths, respectively. For the parametric predictors, smaller values coded reports of more salient continuity percepts and narrower notches, respectively (for a similar application, see Overath et al., 2010). The β value that was derived from fitting a given parametric predictor with the BOLD signal thus provided the slope that described best the changes in the BOLD signal as a function of the changes in that predictor. After orthogonalizing the predictors (described in the next section), for the rating predictors, positive β values indicated decreases in the BOLD signal as a function of more decisive reports of continuity. For the notch width predictors, positive β values indicated increases in the BOLD signal as a function of broader notches. Negative β values indicated the opposites.

Orthogonal factors. To allow investigating influences of continuity perception that were unrelated to stimulus changes, stimulus-related factors and percept-related factors were separated before fMRI data analysis by orthogonalizing the corresponding predictors to each other. This was achieved by extracting the variance that was unique to each predictor, i.e., by removing the covariance of a given predictor and all other predictors using the Gram–Schmidt process (Wilf, 1962). This "decorrelation" procedure thus allowed extracting the interblock variability that was unique to listeners' ratings, i.e., the perceived changes that were linearly unrelated to changes in the notch. A subsequent correlation analysis confirmed that, after orthogonalization, the rating predictors and notch width predictors indeed were unrelated (average Pearson's $r = -0.004\%$). It should be noted that this approach does not take into account statistical dependencies of second or higher order and may depend on the model used for the psychophysical parameters. In the present model, the four possible ratings and three different notch widths were treated as equidistant using simple linear functions (slopes, 1).

Contrast analysis. To localize voxel clusters (cortical regions) that would exhibit the three effects described above (see above, General linear model), three orthogonal contrast analyses were performed as follows. First, sound-sensitive regions (i.e., regions responding significantly to the presentation of the stimuli) were identified by contrasting the averaged no-noise conditions versus baseline.

Second, within these AC regions, effects specifically reflecting continuity illusions were localized by contrasting the rating predictors for the interrupted stimuli versus the rating predictors for the uninterrupted stimuli. This test for rating \times gap interactions was used to extract activity

changes that were attributable to illusory continuity (as inferred from listeners' perceptual reports of the interrupted vowel) but not to non-illusory continuity (as inferred from reports of the uninterrupted vowel). The region that showed such interaction (see Results) was also observed in control analyses testing for simple parametric effects (by contrasting the rating predictors for the interrupted stimuli versus baseline).

Finally, also within the AC regions, effects specifically reflecting the amount of acoustic discontinuities in and around f_0 were localized by contrasting the notch width predictors for the interrupted stimuli versus the notch width predictors for the uninterrupted stimuli. This test for notch width \times gap interactions was used to extract activity changes that were attributable to the acoustic energy during the gap (as determined by the notch width in the interrupted stimuli) but not to the notch in the noise itself (as determined by the notch width in the uninterrupted stimuli). The regions that showed such interactions (see Results) were also observed in control analyses testing for simple parametric effects (by contrasting the notch width predictors for the interrupted stimuli versus baseline).

Region-of-interest analysis. To determine whether continuity illusions and acoustic discontinuities were accompanied by enhanced or reduced neural responses in AC, the activity patterns in regions-of-interest (ROIs) were investigated. Individual ROIs were defined for each of the interaction effects localized in the voxelwise group analyses using the following three steps. First, group ROIs were labeled from the voxels exhibiting the most significant interaction in the multiple comparisons-corrected group map. Second, to allow exploration of potential lateralization, contralateral group ROIs were approximated by mirroring the Talairach x coordinates of the respective unilateral group ROIs (for a similar application, see Overath et al., 2008), gradually lowering the t threshold, and selecting the voxel clusters with the highest t statistic that were situated most proximally to the mirrored locations. Finally, for each of the above group ROIs, an individual ROI was extracted. Specifically, for each listener, the sound-sensitive voxel cluster that exhibited the highest t statistic and the highest proximity to the center of a sphere around the group ROI was selected (sphere radius, 20 mm in Talairach space; approximated by the circles in Fig. 3). The contrast used for this individual ROI definition (averaged no-noise conditions vs baseline) was orthogonal to the interaction tests used for the group analyses.

To investigate the directionality of the presumed interactions, the voxel-averaged BOLD signal time series of each of the individual ROIs was fitted using the GLM described above. Individual β values were extracted and averaged across listeners so that the direction of parametric BOLD signal changes attributable to the notch width or the listener's ratings could be determined for each ROI cluster.

To further validate these parametric changes, simple activity changes associated with individual conditions were examined using a GLM comprising only binary predictors. For the rating conditions, individual β values were extracted from the corresponding group ROI (see Fig. 3B) for each stimulus block. The single-block β values were dichotomized into two categories (continuous or discontinuous, according to the listener's rating) and averaged (first across blocks and then across listeners) so that activity changes associated with each rating category could be determined. These supplemental analyses were done separately for the ambiguous stimulus conditions (no-notch condition and 0.4 octave notch condition).

For the notch conditions, individual β values were extracted from the corresponding individual ROIs (see Fig. 4A). The β values were averaged across listeners so that activity changes associated with each notch width could be determined. For display purposes, β values associated with the individual notch widths were normalized to the no-notch condition. These supplemental analyses were done separately for the interrupted vowel and the uninterrupted vowel.

Vowel versus tone group comparisons. Finally, to infer the stage at which vowel continuity illusions were represented, the region that showed the most significant rating \times gap interaction was compared with a PAC region that showed corresponding effects in a previous tone continuity illusion study (Riecke et al., 2007). That study used essentially the same methods as here; thus, data from that study could be analyzed in the context of the present data as follows. First, the group sizes were matched

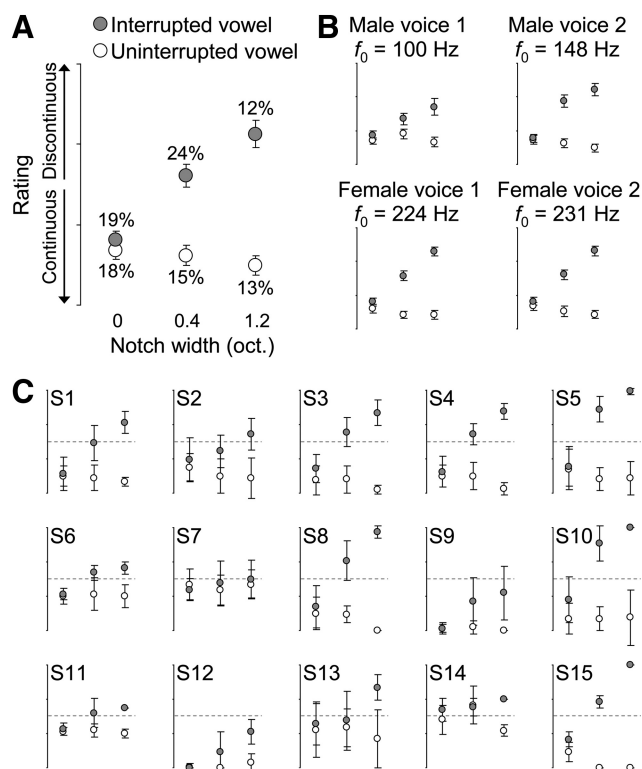


Figure 2. Behavioral results and perceptual switches. **A**, The rating scores (circles, means; bars, SEM across 15 listeners) indicate that stronger masking (induced by narrower notches in the noise) elicited more salient continuity illusions of the interrupted vowel (gray). A smaller opposite effect was found for the uninterrupted vowel (white). The percentages indicate the amount of variance within each stimulus condition (averaged across listeners; for details, see Materials and Methods). oct., Octaves. **B**, Notch width \times gap interactions were obtained for all four voices, suggesting that the effects in **A** were speaker independent. **C**, The length of the error bars indicates the variance of ratings across blocks of physically identical stimuli. This variance was maximal for the interrupted vowel in the 0.4 octave notch condition (see percentages in **A**); this ambiguous condition further evoked frequent switches between illusory continuity and veridical discontinuity [see the overlap of the error bars with the chance level (dotted line) in **C**].

($n = 11$) by excluding data from four participants in the present study who were selected based on the number of ratings they had missed (for details, see supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Second, the individual anatomies were normalized to a common spherical space using high-resolution cortex-based averaging methods that maximally preserve the anatomical specificity of the temporal cortex (Desai et al., 2005; Goebel et al., 2006). More specifically, all individual cortical surface representations were aligned with respect to their mean cortical curvature (obtained by averaging across both groups), and all individual unsmoothed functional data were resampled onto the aligned (normalized) surface representations. Third, for each group, a statistical surface map revealing rating \times gap interactions was computed as before (see above, Contrast analysis). The most significant vertex clusters in the resulting maps were defined as group surface ROIs (size, 25 mm^2). Finally, β values associated with the rating predictors for the interrupted stimuli were extracted for each of the two group surface ROIs.

In addition, regions involved in vowel continuity illusions were compared with regions showing more sensitivity to intact complex sounds than to intact simple sounds. The latter regions were localized by contrasting uninterrupted vowel stimuli versus energy-matched uninterrupted tone stimuli in the no-noise condition. The GLM that was used for these analyses included an additional binary predictor coding for the two groups. β values were obtained as described above. Surface maps were computed using independent samples t tests. All MRI data processing was performed using Brain Voyager QX (Brain Innovation) and Matlab (MathWorks).

Results

Behavioral results

Analysis of listeners' continuity ratings revealed that the interrupted vowel evoked more salient continuity illusions when the missing portion was masked more strongly by the noise (i.e., when the spectral notch in the noise was narrower; $F_{(2,13)} = 28.4$, $p < 10^{-5}$) (Fig. 2A). The uninterrupted vowel showed a smaller opposite effect; these non-illusory stimuli were rated less continuous when their middle portion was masked more strongly ($F_{(2,13)} = 7.0$, $p < 10^{-3}$), in line with previous results on tones (Riecke et al., 2007, 2008). The ratings exhibited similar notch \times gap interactions (all $F_{(2,13)} > 5.2$, $p < 0.05$) (Fig. 2B) for all four voices.

Listeners adjusted their ratings most frequently in the notched-noise condition that was designed to elicit ambiguous percepts (interrupted stimuli, 0.4 octave notch). This was reflected by the interblock variability (see percentages in Fig. 2A) being on average larger in this ambiguous condition than in all the non-illusory conditions (paired t tests, $t_{(14)} > 2.6$, $p < 0.03$). The ambiguous condition further evoked bistability, i.e., frequent switches between reports of illusory continuity and reports of veridical discontinuity (see the overlap of the error bars with the chance level in Fig. 2C).

The perceptual reference stimuli (no-noise conditions) were consistently rated as most continuous and most discontinuous, respectively (data not shown). Because these conditions comprised no noise and were perceptual invariant, they were not considered in the statistical analyses above, which did not affect the overall results.

Imaging results

Mapping results: regions reflecting continuity illusions and acoustic discontinuities

fMRI data analysis revealed that the vowels without noise (no-noise stimuli) evoked widespread activity in bilateral AC, insula, thalamus, putamen, and inferior frontal cortex compared with baseline (Fig. 3A, red regions). A more specific analysis of these sound-sensitive regions based on listeners' continuity ratings revealed a significant rating \times gap interaction in an anterolateral portion of the right-sided Heschl's gyrus (alHG) (Fig. 3B, green region). This effect was unrelated to changes in the notch width (see Materials and Methods, Orthogonal factors). Additional analysis of the sound-related regions based on notch widths revealed a significant notch width \times gap interaction in a region on the left-sided posterior superior temporal gyrus (pSTG) (Fig. 3C, left, blue region). A similar effect was localized in a more anterior region in the right-sided insula (Fig. 3C, right, blue region). All of the above effects were obtained at a corrected significance level of $p < 0.05$ ($t_{(14)} = 2.6$, excluding voxel clusters $< 64 \text{ mm}^3$).

ROI results: activity patterns accompanying continuity illusions and acoustic discontinuities

The interactions that were localized in the group analyses (Fig. 3B,C) were further characterized using analyses of individual ROIs (Fig. 4A). For the region in the right alHG, listeners' ratings of the interrupted vowel were found associated with positive β values (Fig. 4B). Thus, activity in this region decreased as a function of more salient continuity illusions (for details, see Materials and Methods, General linear model). To validate that these parametric decreases reflected genuine bistabilities in listeners' ratings of continuity illusions, switches between continuity reports and discontinuity reports were compared across individual blocks of the ambiguous stimuli. These supplemental analyses

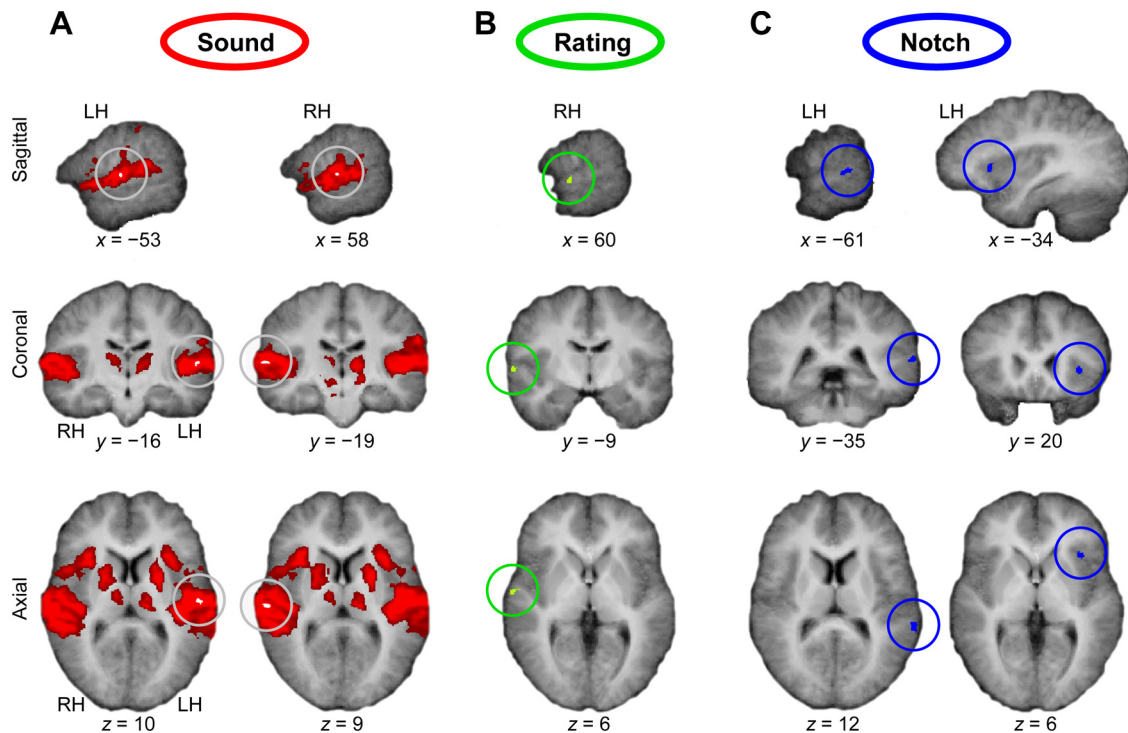


Figure 3. Cortical regions sensitive to sound (**A**), illusory continuity (**B**), and acoustic discontinuity (**C**). Statistical activation maps are projected onto listeners' average anatomical images in Talairach space. **A**, Vowels without noise evoked widespread bilateral activity in AC, insula, thalamus, putamen, and inferior frontal cortex compared with baseline. This sound-evoked activity was strongest (white spot, encircled) in the vicinity of HG. **B**, The aHMG showed effects (rating \times gap interactions) that reflected illusory continuity (as inferred from listeners' perceptual reports of the interrupted vowel) but not veridical continuity (as inferred from reports of the uninterrupted vowel). Importantly, these effects were unrelated to changes in the notch width. **C**, The pSTG and the left anterior insula showed effects that reflected the amount of acoustic discontinuities in and around f_0 (notch width \times gap interactions). LH and RH, Left and right hemisphere; x , y , z , Talairach coordinates in millimeters.

confirmed that reports of continuity illusions were associated with weaker activity than reports of veridical discontinuity percepts of physically identical stimuli (see supplemental Fig. S3A, available at www.jneurosci.org as supplemental material). These stimulus-independent differences thus show that activity evoked by the interrupted vowel was reduced in the right alHG when listeners reported more salient continuity illusions (compared with reports of more salient discontinuity percepts of the same stimuli). Ratings of the non-illusory (uninterrupted) vowel were associated with negative β values (Fig. 4B). In summary, these results implicate that the rating-related effects for the interrupted vowel reflect the salience of continuity illusions rather than illusion-unrelated aspects of the stimuli or the task.

Analysis of the region in left pSTG revealed that the notch widths in the interrupted stimuli were associated with positive β values (Fig. 4B). Thus, activity in this region increased as a function of broader notches in the interrupted stimuli. To validate that these parametric increases reflected acoustic discontinuities in and around f_0 , interrupted and uninterrupted stimuli were compared across the three different noise conditions. These supplemental analyses revealed an activity pattern that closely resembled listeners' averaged continuity ratings (supplemental Fig. S3B, left, available at www.jneurosci.org as supplemental material) (Fig. 2A) and confirmed that broader notches evoked stronger activity in the left pSTG only for interrupted (i.e., acoustically discontinuous) stimuli. A similar activity pattern reflecting acoustic discontinuities was observed in the left anterior insula (supplemental Fig. S3B, right, available at www.jneurosci.org as supplemental material). The majority of the other investigated regions showed smaller positive β values associated with para-

metric notch width increases in the interrupted stimuli (Fig. 4B). In summary, the observed changes show that activity in the left pSTG and other AC regions increased when the stimulus comprised more acoustic discontinuities in and around f_0 .

Contralateral regions showed activity patterns that mostly resembled those reported above for the opposite hemispheres, although the contralateral patterns showed overall less clear parametric changes (Fig. 4B). The spatial variability of the individual ROIs differed slightly across the different ROI clusters (supplemental Fig. S4, available at www.jneurosci.org as supplemental material). No clear parametric effect was observed in the middle portion of HG (mHG) (Fig. 4B), a region that showed the strongest sound-evoked activity (as revealed by comparing the no-noise conditions vs baseline) and represented most the PAC (Hackett et al., 2001; Morosan et al., 2001; Formisano et al., 2003).

Comparison of vowels and tones: processing stages in AC for continuity illusions

The region in alHG that showed an effect related to vowel continuity illusions was further compared with a PAC region that was found previously related to tone continuity illusions (Riecke et al., 2007). Analysis of cortically aligned data from both studies based on listeners' continuity ratings revealed significant rating \times gap interactions in the right HG (Fig. 5A). More specifically, for vowels, this effect was localized in alHG (Fig. 5A, green region), whereas for energy-matched tones, it was localized in a more posteromedial portion in mHG (Fig. 5A, magenta region) ($t_{(10)} > 2.6$, $p < 0.05$, excluding vertex clusters $< 25 \text{ mm}^2$). Analogous analyses in Talairach space revealed a distance of ~ 17

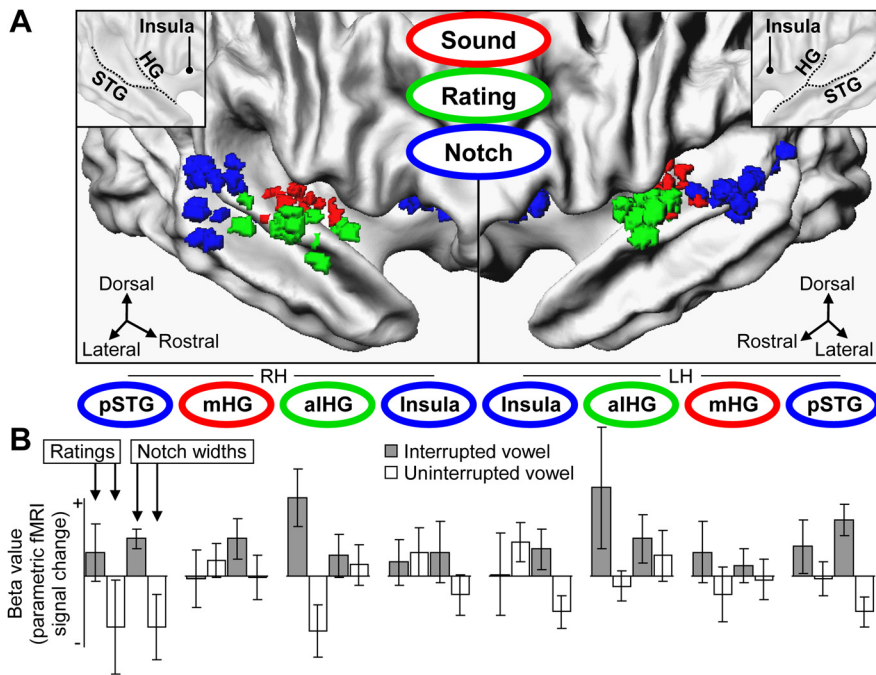


Figure 4. Activity patterns in individual cortical regions sensitive to sound, illusory continuity, and acoustic discontinuity. **A**, Clusters of individual ROIs are projected onto listeners' average representations of the right and left cortical surface (for anatomical details, see inset and supplemental Fig. S4, available at www.jneurosci.org as supplemental material). LH and RH, Left and right hemisphere. **B**, β values (mean \pm SEM across listeners, arbitrary units) associated with continuity ratings or notch widths are shown for the interrupted vowel (gray) and the uninterrupted vowel (white), for each ROI cluster. In the right aIHG, ratings of the interrupted vowel were associated with positive β values, indicating weaker activity to reports of more salient continuity illusions (for details, see supplemental Fig. S3A, available at www.jneurosci.org as supplemental material). In the pSTG, notch widths in the interrupted stimuli were associated with positive β values, indicating stronger activity to more acoustic discontinuities in and around f_0 (for details, see supplemental Fig. S3B, available at www.jneurosci.org as supplemental material).

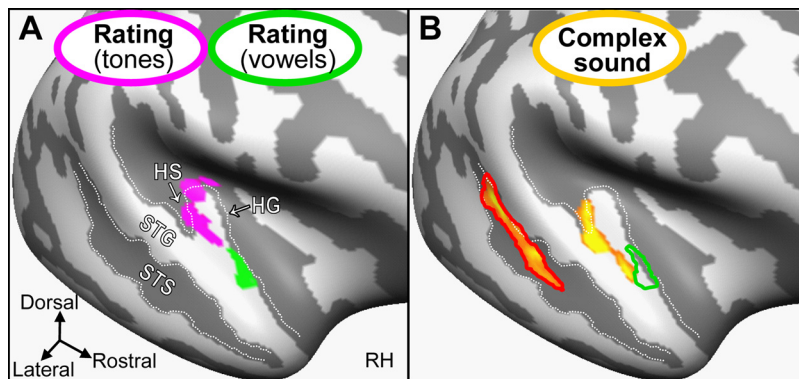


Figure 5. Cortical regions sensitive to the illusory continuity of differently complex sounds (**A**) and the complexity of intact sounds (**B**). **A**, Statistical activation maps obtained from the present study and a matching study on tone continuity illusions are projected onto listeners' unfolded average representation of the right cortical surface (light and dark gray indicate gyri and sulci). A region in aIHG (green region) showed effects related to continuity illusions for the interrupted vowel but not for interrupted tones. A more posteromedial region in mHG (magenta region) showed the opposite. Thus, continuity illusions of vowels and tones may involve distinct neural substrates in non-PAC and PAC, respectively. RH, Right hemisphere. **B**, Regions along the middle STS and in the lateral extensions of HG and HS, but not in aIHG (green outline), showed stronger activity to intact complex sounds (vowels) than to intact simple sounds (tones). The region in STS (red outline) was compatible with a previously reported voice-sensitive region and showed sensitivity to listeners' ratings of the interrupted vowel.

mm between the centers of mass of the two illusion-related regions. Additional analyses of surface ROIs based on listeners' ratings of the interrupted stimuli revealed no effect in mHG for vowels ($t_{(10)} = -0.4$, $p = 0.7$) and no effect in aIHG for tones ($t_{(10)} = 0.4$, $p = 0.7$). Therefore, these results indicate dissociation

between the anterolateral and posteromedial portions of the right HG regarding their sensitivity to continuity illusions of vowels and tones, respectively.

Finally, the region related to vowel continuity illusions was compared with regions that showed sensitivity to the complexity of intact sounds. Analysis of cortically aligned data revealed significantly stronger activity to the intact vowel than to energy-matched intact tones in several regions of the right superior temporal cortex (Fig. 5B, yellow regions) ($t_{(20)} = 2.2$, $p < 0.05$, excluding vertex clusters $< 25 \text{ mm}^2$), including the middle portion of superior temporal sulcus (STS) and the lateral extensions of HG and Heschl's sulcus (HS) but not the region in aIHG. The region in STS (Fig. 5B, red outline) was in agreement with a previously reported voice-sensitive region (Belin et al., 2000), and it showed an effect related to listeners' ratings of the interrupted vowel ($t_{(10)} = -2.7$, $p < 0.05$).

Additional repeated-measures analyses of data from one listener (S10) who participated in both studies revealed trends that were consistent with the above group comparison results (supplemental Fig. S5, available at www.jneurosci.org as supplemental material).

Discussion

The results demonstrate that listeners' reports of vowel continuity illusions are accompanied by significantly weaker activity in AC compared with reports of veridical discontinuity percepts. These illusion-specific reductions were found in the right aIHG under identical stimulus conditions and also in the left pSTG under varying stimulus conditions. The region in aIHG was further found dissociated from a region in PAC that showed an analogous effect related to tone continuity illusions and also from regions that were most sensitive to intact complex sounds. Overall, the results support the view that continuity illusions of simple sounds and complex sounds are facilitated by a common suppressive mechanism. They further suggest that this mechanism operates on different sound representations in AC.

Suppression as a common mechanism for continuity illusions

A common aspect of the effects of continuity illusions that were observed in the present study and in previous studies that also controlled for confounding stimulus effects (Riecke et al., 2007, 2009; Shahin et al., 2009) was their "suppressive" nature. In other words, neural activity to interrupted sounds in AC was likely weaker when these sounds were judged as continuous rather than discontinuous. It

should be noted that the term “suppression” here refers to a partial reduction of sound-evoked activity (i.e., above resting state activity) (supplemental Fig. S3A, available at www.jneurosci.org as supplemental material). Such continuity illusion-related reductions were presumably caused by forward suppression (Wehr and Zador, 2005) or adaptation (Ulanovsky et al., 2003, 2004) because these mechanisms in AC have been implicated also in possibly related auditory phenomena (Bregman et al., 1999) such as the integration of simple (Micheyl et al., 2005) or complex (Gutschalk et al., 2007) sound streams. It is conceivable that, in the case of continuity illusions, the early portion of the interrupted sound partially suppresses the neural response to the middle portion, provided that this middle portion is masked (see Introduction). For that case, the neural response to the perceptual mismatch between the early portion and the masked portion might not evoke a detectable change.

Sound representations in AC supporting continuity illusions

Our observation that the region in aHGM was dissociated from the region in PAC related to tone continuity illusions (Riecke et al., 2007) cannot be ascribed to differences in the task because both regions were obtained using identical tasks. Instead, this dissociation may reflect differences in the complexity of the illusory sounds (voices uttering a vowel vs tones). This dissociation has two main implications. First, continuity illusions of differently complex sounds may involve different regions in AC. Second, although the region for tone continuity illusions in mHGM fits with cytoarchitecturally defined medial subdivisions of human PAC, Te1.0 and Te1.1 (Morosan et al., 2001), the region for vowel continuity illusions in aHGM seems to extend laterally from primary area Te1.2 into the bordering nonprimary area Te3 (Morosan et al., 2005). Together with the notion that non-PAC constitutes a later stage in the primate’s auditory processing hierarchy than PAC (Rauschecker et al., 1997; Kaas and Hackett, 2000), these observations suggest that continuity illusions of differently complex sounds involve sound representations at different AC stages.

The results provide two indications regarding the specific sound representations that may have been involved in the present continuity illusions. First, the dissociation of the region in aHGM from regions that were most sensitive to intact complex sounds suggests that the latter regions were relatively little involved. A putative voice-sensitive region in the right STS showed an effect, suggesting that representations of intact voices were involved. However, interpretations in terms of sensitivity or specialization for voices or vowels remain tentative here because our methods were not suited for investigating individual voice-sensitive regions (Belin et al., 2000) or vowel-specific stimulus aspects [e.g., vowel identity (Carlyon et al., 2002)]. Second, in the literature, the aHGM and surrounding superior temporal regions have been implicated in the analysis of paralinguistic aspects of voices such as pitch (Belin et al., 2002; Warren and Griffiths, 2003; Barrett and Hall, 2006), saliency (Warren et al., 2005), and intelligibility (Davis and Johnsrude, 2003; Liebenthal et al., 2005; Scott et al., 2006). Specifically the right aHGM in non-PAC has been suggested to play a role in extracting the pitch of intact natural voices (Lattner et al., 2005) and synthetic broadband sounds (Patterson et al., 2002; Penagos et al., 2004). The region that we identified in right aHGM fits well with these previously reported pitch-sensitive regions (average Euclidean distance in Talairach space, 10 mm), suggesting that pitch representations were involved in the present continuity illusions. This notion is supported by the facts that our stimuli had a clear harmonic structure and continuity illusions

depended on f_0 . The f_0 is a determining factor for vocal pitch, and, possibly, our listeners exploited this cue for rating continuity (Plack and White, 2000). Thus, it is conceivable that continuity illusions of complex sounds with clear pitch such as the vowels used here depend on complex pitch representations in the right aHGM.

Although the effect in right aHGM could be ascribed to changes in reports of continuity illusions under identical stimulus conditions, it cannot be resolved what caused these changes. Presumably, non-acoustic factors that might influence the illusion, such as spontaneous fluctuations in neural activity (Micheyl et al., 2005), attention, or top-down expectancies (Sivonen et al., 2006), played a role (for an additional discussion, see Riecke et al., 2009).

Acoustic analysis of complex sounds in posterior temporal cortex

The effect of the spectral notch in the left pSTG indicates that this non-PAC region may be sensitive to acoustic discontinuities in f_0 , a determining factor for continuity illusions in the present study. Similar regions in pSTG, particularly in the left hemisphere, have been associated with the encoding of acoustic properties of speech (for review, see Hickok and Poeppel, 2000; Scott, 2005), such as voice formants (Lattner et al., 2005) and speech onsets or offsets (Celsis et al., 1999; Jancke et al., 1999; Harms et al., 2005). A region in the left middle temporal gyrus that has been associated with continuity illusions of complex sounds (see Introduction) showed a similar activity pattern in our study as the region we identified in pSTG (supplemental Fig. S6A, available at www.jneurosci.org as supplemental material). This corroborates the previous suggestion that vowel-sensitive posterior temporal regions may be sensitive to acoustic gaps in complex sounds (Heinrich et al., 2008). However, in our study, none of the previously reported regions exhibited a stimulus-unrelated effect of continuity illusions like the effect we observed in aHGM (supplemental Fig. S6, available at www.jneurosci.org as supplemental material). This discrepancy may be attributable to methodological differences, for example, our stimuli were more natural and probably evoked stronger pitch than some of the previously used stimuli (artificial formants, speech). Thus, our effects in aHGM may reflect illusory pitch of natural vocal sounds rather than vowel identity or speech-specific qualities, which are more likely restored in posterior temporal regions (Heinrich et al., 2008; Shahin et al., 2009).

Sensory-perceptual transformations of fragmented complex sounds in temporal cortex

It remains a matter of debate how temporal regions interact during continuity illusions of complex sounds. A possible explanation for the observed effect pattern is as follows. The fragmented sound and interrupting noise are integrated in pSTG and in proximate regions (Heinrich et al., 2008), provided that no discontinuities were detected in the sensory stimulus representation. This integrative process in posterior temporal cortex (Warren et al., 2005; Obleser et al., 2007) may be controlled by the insula, given the similar response profiles of these two regions associated with continuity illusions (Shahin et al., 2009). The outcomes of these preliminary sensory analyses could trigger subsequent illusory processes (Darwin, 2005). For fragmented speech, these processes could take place in posterior regions representing phonotactic schemas (Shahin et al., 2009), whereas for more tonal sounds they could take place in anterior AC regions subserving the identification of sound objects (Rauschecker and Scott, 2009). In the latter case, pitch representations in aHGM could

facilitate interpolation of f_0 during the noise and thus render the fragmented sound more continuous, provided that pitch-encoding neurons in this region (Bendor and Wang, 2005, 2006) failed to detect a pitch mismatch between that sound and the noise (Plack and White, 2000). The restoration process could be further influenced by non-acoustic factors for auditory object identification such as attention (Ahveninen et al., 2006; Alain and Bernstein, 2008).

Conclusions

Auditory continuity illusions may emerge from a common suppressive mechanism that counteracts interruptions of relevant sounds at multiple stages of sound representation in AC. For continuity illusions of simple sounds, such as tones, the mechanism may operate on primary frequency representations of the illusory sound. For continuity illusions of more complex sounds, such as vowels, the mechanism may operate on nonprimary and perceptually more integrated representations, such as complex pitch representations. This mechanism could facilitate smooth hearing of sounds that are interrupted by masking sounds.

References

- Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated “what” and “where” pathways in human auditory cortex. *Proc Natl Acad Sci U S A* 103:14608–14613.
- Alain C, Bernstein LJ (2008) From sounds to meaning: the role of attention during auditory scene analysis. *Curr Opin Otolaryngol Head Neck Surg* 16:485–489.
- Assmann PF, Paschall DD (1998) Pitches of concurrent vowels. *J Acoust Soc Am* 103:1150–1160.
- Barrett DJ, Hall DA (2006) Response preferences for “what” and “where” in human non-primary auditory cortex. *Neuroimage* 32:968–977.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403:309–312.
- Belin P, Zatorre RJ, Ahad P (2002) Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res* 13:17–26.
- Bendor D, Wang X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature* 436:1161–1165.
- Bendor D, Wang X (2006) Cortical representations of pitch in monkeys and humans. *Curr Opin Neurobiol* 16:391–399.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and non-speech sounds. *Cereb Cortex* 10:512–528.
- Boersma P (2001) Praat, a system for doing phonetics by computer. *Glott Int* 5:341–345.
- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Massachusetts Institute of Technology.
- Bregman AS, Colantonio C, Ahad PA (1999) Is a common grouping mechanism involved in the phenomena of illusory continuity and stream segregation? *Percept Psychophys* 61:195–205.
- Carlyon RP, Deeks J, Norris D, Butterfield S (2002) The continuity illusion and vowel identification. *Acustica Acta Acustica* 88:408–415.
- Celsis P, Boulanouar K, Doyon B, Ranjeva JP, Berry I, Nespoulous JL, Chollet F (1999) Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage* 9:135–144.
- Culling JF, Darwin CJ (1993) Perceptual separation of simultaneous vowels: within and across-formant grouping by F0. *J Acoust Soc Am* 93:3454–3467.
- Darwin CJ (2005) Simultaneous grouping and auditory continuity. *Percept Psychophys* 67:1384–1390.
- Davis MH, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. *J Neurosci* 23:3423–3431.
- Desai R, Liebenthal E, Possing ET, Waldron E, Binder JR (2005) Volumetric vs. surface-based alignment for localization of auditory cortex activation. *Neuroimage* 26:1019–1029.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R (2003) Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* 40:859–869.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. *Neuroimage* 7:30–40.
- Gaab N, Gabrieli JD, Glover GH (2007) Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Hum Brain Mapp* 28:703–720.
- Goebel R, Esposito F, Formisano E (2006) Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp* 27:392–401.
- Gutschalk A, Oxenham AJ, Micheyl C, Wilson EC, Melcher JR (2007) Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. *J Neurosci* 27:13074–13081.
- Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 441:197–222.
- Harms MP, Guinan JJ Jr, Sigalovsky IS, Melcher JR (2005) Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI. *J Neurophysiol* 93:210–222.
- 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.
- Hess W (1983) Pitch determination of speech signals. Berlin: Springer.
- Hickok G, Poeppel D (2000) Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci* 4:131–138.
- Houtgast T (1972) Psychophysical evidence for lateral inhibition in hearing. *J Acoust Soc Am* 51:1885–1894.
- Hwa Chen S (2007) Sex differences in frequency and intensity in reading and voice range profiles for Taiwanese adult speakers. *Folia Phoniatr Logop* 59:1–9.
- Jäncke L, Mirzazade S, Shah NJ (1999) Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266:125–128.
- Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A* 97:11793–11799.
- Lattner S, Meyer ME, Friederici AD (2005) Voice perception: sex, pitch, and the right hemisphere. *Hum Brain Mapp* 24:11–20.
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA (2005) Neural substrates of phonemic perception. *Cereb Cortex* 15:1621–1631.
- 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.
- Micheyl C, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48:139–148.
- Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13:684–701.
- Morosan P, Schleicher A, Amunts K, Zilles K (2005) Multimodal architectonic mapping of human superior temporal gyrus. *Anat Embryol (Berl)* 210:401–406.
- Obleser J, Zimmermann J, Van Meter J, Rauschecker JP (2007) Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex* 17:2251–2257.
- Overath T, Kumar S, von Kriegstein K, Griffiths TD (2008) Encoding of spectral correlation over time in auditory cortex. *J Neurosci* 28:13268–13273.
- Overath T, Kumar S, Stewart L, von Kriegstein K, Cusack R, Rees A, Griffiths TD (2010) Cortical mechanisms for the segregation and representation of acoustic textures. *J Neurosci* 30:2070–2076.
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36:767–776.
- Penagos H, Melcher JR, Oxenham AJ (2004) A neural representation of

- pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J Neurosci* 24:6810–6815.
- Petkov CI, O'Connor KN, Sutter ML (2007) Encoding of illusory continuity in primary auditory cortex. *Neuron* 54:153–165.
- Plack CJ, White LJ (2000) Perceived continuity and pitch perception. *J Acoust Soc Am* 108:1162–1169.
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724.
- Rauschecker JP, Tian B, Pons T, Mishkin M (1997) Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol* 382:89–103.
- Riecke L, van Opstal AJ, Goebel R, Formisano E (2007) Hearing illusory sounds in noise: sensory-perceptual transformations in primary auditory cortex. *J Neurosci* 27:12684–12689.
- Riecke L, Van Opstal AJ, Formisano E (2008) The auditory continuity illusion: a parametric investigation and filter model. *Percept Psychophys* 70:1–12.
- Riecke L, Esposito F, Bonte M, Formisano E (2009) Hearing illusory sounds in noise: the timing of sensory-perceptual transformations in auditory cortex. *Neuron* 64:550–561.
- Schweinberger SR, Casper C, Houthal N, Kaufmann JM, Kawahara H, Kloth N, Robertson DM, Simpson AP, Zäske R (2008) Auditory adaptation in voice perception. *Curr Biol* 18:684–688.
- Scott SK (2005) Auditory processing: speech, space and auditory objects. *Curr Opin Neurobiol* 15:197–201.
- Scott SK, Rosen S, Lang H, Wise RJ (2006) Neural correlates of intelligibility in speech investigated with noise vocoded speech: a positron emission tomography study. *J Acoust Soc Am* 120:1075–1083.
- Shah NJ, Steinhoff S, Mirzazade S, Zafiris O, Grosse-Ruyken ML, Jäncke L, Zilles K (2000) The effect of sequence repeat time on auditory cortex stimulation during phonetic discrimination. *Neuroimage* 12:100–108.
- Shahin AJ, Bishop CW, Miller LM (2009) Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44:1133–1143.
- Sivonen P, Maess B, Lattner S, Friederici AD (2006) Phonemic restoration in a sentence context: evidence from early and late ERP effects. *Brain Res* 1121:177–189.
- Sugita Y (1997) Neuronal correlates of auditory induction in the cat cortex. *Neuroreport* 8:1155–1159.
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
- 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.
- Warren JD, Griffiths TD (2003) Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J Neurosci* 23:5799–5804.
- Warren JD, Jennings AR, Griffiths TD (2005) Analysis of the spectral envelope of sounds by the human brain. *Neuroimage* 24:1052–1057.
- 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.
- Wehr M, Zador AM (2005) Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron* 47:437–445.
- Wessinger CM, VanMeter J, Tian B, Van Lare J, Pekar J, Rauschecker JP (2001) Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci* 13:1–7.
- Wilf HS (1962) Mathematics for the physical sciences. New York: Dover Publications.