

Hearing Illusory Sounds in Noise: Sensory-Perceptual Transformations in Primary Auditory Cortex

Lars Riecke,¹ A. John van Opstal,² Rainer Goebel,¹ and Elia Formisano¹

¹Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, 6200 MD, Maastricht, The Netherlands, and ²Department of Biophysics, Institute for Neuroscience, Radboud University Nijmegen, 6500 HB, Nijmegen, The Netherlands

A sound that is interrupted by silence is perceived as discontinuous. However, when the silence is replaced by noise, the target sound may be heard as uninterrupted. Understanding the neural basis of this continuity illusion may elucidate the ability to track sounds of interest in noisy auditory scenes, but yet little is known. In the present functional magnetic resonance imaging study in humans we report that activity in primary auditory cortex reflects perceived continuity of illusory tones in noise. Exploiting a parametric manipulation of the illusory stimuli, we show that stimulus-evoked activity does not correlate with the basic acoustic properties of tones or noises, but rather with the abstract dependencies among them. Importantly, changes of neural responses to acoustically identical stimuli parallel changes of listeners' report of perceived continuity of these same stimuli, thus confirming the perceptual nature of these responses. Our findings show that, beyond the sensory representation of an auditory scene, primary auditory areas play a constructive role in the grouping of scene segments into unified auditory percepts.

Key words: illusion; hearing; perceptual grouping; auditory cortex; fMRI; object recognition

Introduction

We can recognize sounds of interest remarkably well even when these are masked by background noise. For example, when a tone is interrupted by noise, the interrupted sound may be heard as continuing through the noise. This continuity illusion illustrates the constructive nature of perception that may serve to enhance sensitivity to expected signals and to ensure robustness against background noise. Psychoacoustic research investigated the continuity illusion for various sounds (tones, sweeps, melodies, vocalizations, and speech), and in several species [birds (Braaten and Leary, 1999), cats (Sugita, 1997), monkeys (Miller et al., 2001; Petkov et al., 2003), humans (e.g., Warren, 1999)], suggesting that common perceptual mechanisms may operate at multiple levels of abstraction. The current view holds that the gap induces sudden energy changes in the frequency channel of the interrupted sound (target) which are therefore perceptually interpreted as target off- and onsets. When such changes in the so-called on-frequency band are masked by another sound (masker), they become less salient and the masker may be interpreted as containing the target. The continuity of the target after the gap facilitates the emergence of the illusory target during the interruption (Bregman, 1990).

The neural correlates of the continuity illusion are poorly un-

derstood. Previous research identified neurons in cats' primary auditory cortex (AC) presumably involved in the processing of the spectral-temporal properties of targets and maskers (Sugita, 1997). A recent electrophysiology study (Petkov et al., 2007) reported that neurons in monkey core area AI respond to illusory tones interrupted by noise as if these tones were continuous. These results thus suggest that activity in primary AC reflects perceptual rather than acoustic stimulus properties.

In humans, the only neurophysiological study (Micheyl et al., 2003) so far demonstrated that changes in the on-frequency band of the target tone are accompanied by changes in the level of the mismatch-negativity. The mismatch-negativity is a short-latency event-related scalp potential thought to originate from AC; it may reflect the processing of an acoustic short-term buffer that detects unexpected acoustic deviances in predictable stimulus sequences (Näätänen et al., 2001). Therefore, these results suggest that stimuli evoking continuity illusions recruit areas in AC involved in preattentive detection of unexpected on-frequency band changes. Unfortunately, neither study could relate brain activities to listeners' actual percepts as behavioral data were not obtained during the physiological measurements.

In the present human functional magnetic resonance imaging (fMRI) study we measure blood oxygenation level dependent (BOLD) responses in AC to illusory and nonillusory tones (targets) with different levels of masking and we identify the regions in which activity levels covary with the subjective report of perceived continuity. The masking of the omitted on-frequency bands was varied across four parametric levels, and listeners rated the targets' overall perceived continuity on a four-point scale while functional images were collected. To dissociate brain regions related to continuity illusions from those related to nonillusory continuity percepts we presented interrupted and uninter-

Received June 14, 2007; revised Sept. 3, 2007; accepted Oct. 1, 2007.

This work was supported by the Netherlands Organization for Scientific Research (Cognitie program Grant 05104020 and VIDI Grant 452-04-330). We thank Armin Heinecke, Nick Kilian-Huetten, Daniel Mendelsohn, and Fabrizio Esposito for their contributions, and the anonymous reviewers for their comments.

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

DOI:10.1523/JNEUROSCI.2713-07.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/2712684-06\$15.00/0

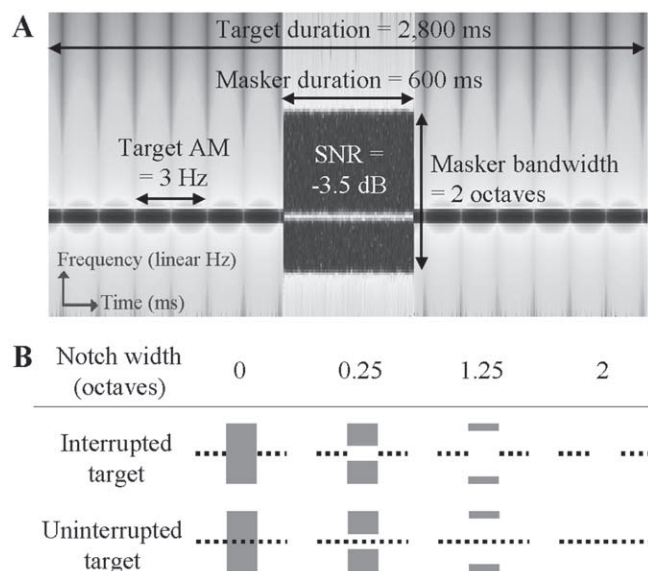


Figure 1. Stimulus and experimental design. **A**, Spectrogram and parameter settings, exemplified for a target stimulus of 1732 Hz, interrupted for 600 ms by a noise burst containing a 0.25 octave notch. The noise and the notch were logarithmically centered on the target frequency. AM, Amplitude modulation; SNR, signal-to-noise ratio. **B**, Schematic stimulus spectrograms and labels of all eight stimulus conditions. Dots represent AM of tone targets which contain a temporal gap (top) or not (bottom). Gray rectangles represent noise maskers comprising a spectral notch of varying width.

rupted targets. We optimized stimuli and task for fMRI demands in a preceding psychoacoustics study (Riecke et al., 2007).

Materials and Methods

Participants

Eleven human volunteers (mean age, 26 years, seven women) with normal hearing abilities participated after providing informed consent. The local ethics committee approved the experimental protocol.

Stimuli, design, and task

Tones of varying frequency (500, 930, 1732, 3223 Hz) with sine-modulated amplitude (3 Hz) were used as targets. Broadband Gaussian noise bursts were used as maskers. Noise bursts were bandpassed and band-stopped (notched) with both filters centered in the on-frequency band. To vary the masking level, the width of the spectral notch was parameterized across 0, 0.25, 1.25, and 2 octaves. Maskers were superimposed in the spectral-temporal center of targets on a logarithmic-linear scale. In interrupted target stimuli, onsets and offsets of gaps and maskers were synchronized. All onsets and offsets were linearly ramped with 3 ms rise/fall times. Uninterrupted target stimuli were spectrotemporally matched, except that no gaps were inserted in the target. All stimuli were matched for overall spectral power and sampled with 16 bits at 44.1 kHz in Matlab 7.0.1 (MathWorks, Natick, MA). Stimulus parameter settings are summarized in Figure 1A.

Before imaging, listeners were trained to attend to targets and to rate their overall perceived continuity on a four-point scale (labeled with “most likely continuous,” “probably continuous,” “probably discontinuous,” “most likely discontinuous”). Experimental conditions (Fig. 1B) were presented during blocks of stimulation, each of which comprised four targets of either low (500, 930 Hz) or high (1732, 3223 Hz) frequencies in randomized order delivered diotically via Commander XG headphones (Resonance Technology, Northridge, CA) at 70 dB sound pressure level (SPL). Different stimulation blocks were matched and randomized with respect to their number and order so that successive blocks always differed. Stimulation blocks alternated with stimulation-free baseline blocks of same duration (20 s) during which listeners rated the preceding stimuli.

Imaging and data preprocessing

Images were collected with a Siemens (Erlangen, Germany) Allegra 3-Tesla magnetic resonance imaging system. BOLD signal changes were measured with a head coil using a gradient echo planar imaging (EPI) sequence [time to echo (TE), 30 ms; acquisition time, 2200 ms; repetition time (TR), 5000 ms; field of view, 256×256 mm²; matrix size, 128×128 ; slice thickness, 2 mm]. During each experiment, 134 volumes were collected, each comprising 28 axial contiguous slices centered on the Sylvian fissure and covering the entire AC. In total, 40 functional runs were obtained (seven listeners by four runs, four listeners by three runs). Structural T1-weighted volumes optimized for gray-white matter contrast were obtained using three-dimensional modified-driven equilibrium with Fourier transform pulse sequences (voxel resolution, $1 \times 1 \times 1$ mm³).

The clustered volume EPI technique that was applied allowed for presentation of auditory stimuli in silence between subsequent volume acquisitions (Jäncke et al., 2002; Van Atteveldt et al., 2004). To increase statistical detection power and support listeners' attentiveness (Shah et al., 2000; Gaab et al., 2007) we used a relatively short TR, compared with that of other sparse temporal sampling techniques (Hall et al., 1999). To minimize the resulting risk of the overlapping between stimulus-related BOLD responses and residual BOLD responses to preceding acoustic scanner noise that may arise from usage of such short TR (Belin et al., 1999; Hall et al., 1999; Talavage and Edmister, 2004) we took the following steps. First, we used carrier frequencies that avoided the peaks of the EPI noise spectrum (fundamental frequency at ~ 1250 Hz) and presented the stimuli in silence, which resulted in clear spectral-temporal separation between the acoustic stimuli and the scanner noise. Second, our listeners wore double ear protection (ear plugs and ear muffs) which attenuated the EPI noise level by ~ 50 dB SPL and allowed for stimulus presentation at a level perceived as much higher than that of the preceding EPI noise. Finally, listeners were instructed to attend to carrier frequencies, but not to the ambient EPI noise, which may have further enhanced the contrast-to-noise (CNR) ratio of the measured BOLD signals (Jäncke et al., 1999; Hall et al., 2000). The results (supplemental Fig. 1A, available at www.jneurosci.org as supplemental material) indicate appropriate CNR in our BOLD data.

The collected fMRI data were analyzed using Brain Voyager QX 1.7 (Brain Innovation, Maastricht, The Netherlands) and Matlab. Preprocessing included head motion correction (no head motion exceeded 1.5 mm), interslice acquisition time correction, temporal high-pass filtering (cutoff, seven cycles per time course), functional-anatomical image coregistration, normalization to Talairach space, and cortical surface reconstruction. To maximally preserve anatomical specificity, group analysis was performed using cortical alignment methods (Goebel et al., 2006) and nonsmoothed functional data. Analysis was also performed in conventional Talairach space using functional data spatially smoothed with a 4 mm full-width half-maximum Gaussian kernel.

Statistical analyses

General linear model analysis. Listeners' perceived continuity rating data were averaged across block repetitions and analyzed in SPSS 12.0.1 (SPSS, Chicago, ILL) using a general linear model (GLM) and an ANOVA for repeated measures. For fMRI data, group statistical maps were obtained with a vertex-by-vertex (or voxel-by-voxel, for Talairach space-based analysis) two-level random-effect analysis of the BOLD signal time series. At the first-level, a GLM was computed for each experiment using separate predictors for each listener. The predicted time courses were adjusted for the hemodynamic response delay by convolution with a double-gamma hemodynamic response function. At the second level, group contrast *t* maps were obtained based on the parameter estimates (β values) derived from the first level analysis.

In the stimulus-based analysis of the fMRI data, the GLM included two predictors for interrupted and uninterrupted targets, respectively: A binary predictor coded for target presentation, and a linearly scaled predictor coded for the notch width parameterization, respectively. The contrast between the notch width parameterization for interrupted versus uninterrupted targets was computed to identify voxels exhibiting significant masker by gap interactions. The subsequent analyses in regions of

interest (ROIs) involved assessment of BOLD response differences between individual conditions, based on a GLM that included a binary predictor for each condition (Fig. 1*B*).

In the subsequent percept-based analysis of the fMRI data, two linearly scaled predictors that coded for listeners' ratings of perceived continuity of interrupted and uninterrupted targets, respectively, were added to the GLM. To allow for detection of perceptual changes that were unrelated to stimulus changes, the two predictors were orthogonalized with respect to all other stimulus-related predictors in the model as described in the following.

Orthogonalization. A general problem in physiological studies of perception stems from the fact that percept changes are typically induced by concomitant stimulus changes (stimulus-related changes). Given this collinearity, correlated BOLD signal changes can be attributed to changes in the stimulus and in the concomitant percept. However, percept changes may also occur in the absence of concomitant stimulus changes (stimulus-unrelated changes) (e.g., the same stimulus may be perceived differently, depending on subjective factors like the amount of attention that the listener paid to the stimulus). BOLD signal changes that correlate with such stimulus-unrelated changes can be attributed exclusively to percept changes but not to stimulus changes.

In the current study, to test whether there was such additional stimulus-unrelated variance that could be explained exclusively by subjective changes (i.e., by perceived continuity changes), stimulus-related variance was first removed from the measured BOLD responses. To this end, the stimulus-unrelated variance from listeners' perceived continuity data were extracted, using a linear algebra procedure called Gram–Schmidt orthogonalization (Wilf, 1962). Because this procedure operates in vector space, the variance in each predictor (see GLM analysis) was first transformed into its vector representation (Bandettini et al., 1993). The Gram–Schmidt process then extracted the component of a given vector that was orthogonal to all the components of another given vector. Specifically, the procedure extracted the part of variance in one predictor that was linearly uncorrelated to the variance in another predictor. This allowed for computation of two perceived continuity predictors (one for interrupted and one for uninterrupted targets) that were unrelated to the respective notch width predictors. Each of these linearly scaled predictors thus allowed identifying voxels with significant rating-related BOLD changes that occurred without concomitant changes in the respective notch width predictor. In other words, these predictors allowed for an analysis of partial correlations, (i.e., the calculation of the relative amount of variance in the BOLD response that was explained exclusively by changes in the perceived continuity).

Thresholding. Random-effect maps were thresholded based on a three-dimensional extension of the randomization procedure described by Forman et al. (1995). First, a voxelwise threshold was set to $t_{10} = 3.5$ (uncorrected $p < 0.005$). Thresholded maps were then submitted to a whole-brain correction criterion based on the estimate of the map's spatial smoothness and on Monte Carlo simulations for estimating cluster-level false-positive rates after 1000 iterations. Maps were applied the minimum cluster size threshold which yielded a corrected cluster-level false-positive rate of $p < 0.05$. Multiple comparison-corrected maps were superimposed on the cortical surface that was obtained from alignment of individual cortices (cortex-based analysis), or on the average volume (Talairach space-based analysis). The most significant clusters were defined as ROIs.

ROI time course analysis. ROI analysis of BOLD signal time courses was based on mean amplitudes defined by a time window around the peak of the condition-related average time courses. Mean amplitudes associated with the different notch width conditions for interrupted and uninterrupted target stimuli were normalized with respect to the mean amplitude in the respective zero-octave notch condition.

Best-frequency mapping. To identify the locations of auditory cortical fields (ACFs), a best-frequency mapping technique (Formisano et al., 2003) was applied to the cortically aligned functional data. The analysis considered only vertices in which activity differed significantly between targets versus baseline (supplemental Fig. 1*A*, available at www.jneurosci.org as supplemental material). The two lower and upper carrier frequencies were binned into a low-frequency and high-frequency con-

dition, respectively. Using GLM analysis, a frequency-selectivity index FS of the two associated predictors was computed as the ratio of their subtracted β values to their summed β values. Ratios were color-coded at each vertex with hue and saturation coding for the field sign and value of FS , respectively.

Results

Consistent with our recent results obtained outside the scanner (Riecke et al., 2007), behavioral data analysis revealed that perceived continuity increased significantly with masking levels for only interrupted targets but not for uninterrupted targets (masker by gap interaction, $F_{(3,79)} = 164.4$; $p < 10^{-34}$) (Fig. 2*B*). fMRI data analyses on the cortical sphere and in Talairach space yielded consistent results (see below).

Stimulus-based analysis

Interrupted and uninterrupted targets evoked widespread activities in bilateral AC relative to baseline ($t_{10} = 4.6$, cluster size >100 ; corrected $p < 0.05$) (supplemental Fig. 1*A*, available at www.jneurosci.org as supplemental material). The analysis of the parametric manipulation of the masker in these regions (based on the contrast between the parametric notch width predictors for interrupted versus uninterrupted targets) revealed that activation levels in the right lateral transverse gyrus [Heschl's gyrus (HG)] and in adjacent superior temporal gyrus (STG) correlated significantly differently with the masking levels of interrupted versus that of uninterrupted targets (masker by gap interaction, $t_{10} = 2.6$; cluster size, >50 ; corrected $p < 0.05$) (Fig. 2*A*, supplemental Figs. 1*B*, 2, blue regions, available at www.jneurosci.org as supplemental material). Comparable effects were also observed in the left hemisphere ($t_{10} = 3.5$; cluster size, >100 ; corrected $p < 0.05$) (supplemental Fig. 1*B*, available at www.jneurosci.org as supplemental material). The most significant clusters in the right hemisphere were defined as ROIs and the patterns of their masker by gap interactions were further specified (Fig. 2*A*, supplemental Fig. 1*B*, blue arrows, available at www.jneurosci.org as supplemental material). For interrupted targets, increases in masking were associated with significant increases in the strength of the continuity illusion ($F_{(3,79)} = 98.4$; $p < 10^{-26}$) (Fig. 2*B*, supplemental Fig. 3, available at www.jneurosci.org as supplemental material) and significant decreases in brain activity in lateral HG and STG (Fig. 2*B,C*, supplemental Fig. 3, available at www.jneurosci.org as supplemental material) (ROI analysis, $t_{10} = 2.6$ and 4.6; $p = 0.02$ and 0.0009), whereas for uninterrupted targets no such parametric effects were observed (ROI analysis, $t_{10} = -1.1$ and -0.2 ; $p = 0.3$ and 0.8). At the highest masking level (no-notch condition), interrupted targets were perceived as continuity illusions, and were not associated with significant activity differences in STG compared with the nonillusory continuity percepts of uninterrupted targets (ROI analysis, $t_{10} = -1.9$; $p = 0.08$).

Percept-based analysis

The analysis so far suggested that lateral HG and STG play a role in the detection of changes in the on-frequency band as well as in the concomitant change in the perception of the overall continuity of the stimulus. To investigate whether the same regions were also involved in changes in listeners' actual percepts alone (i.e., without concomitant on-frequency band changes), another analysis was performed based on the reports of listeners' perceived continuity that had been collected during the functional measurements (see Materials and Methods). Strikingly, this analysis revealed that the changes in perceived continuity that were unre-

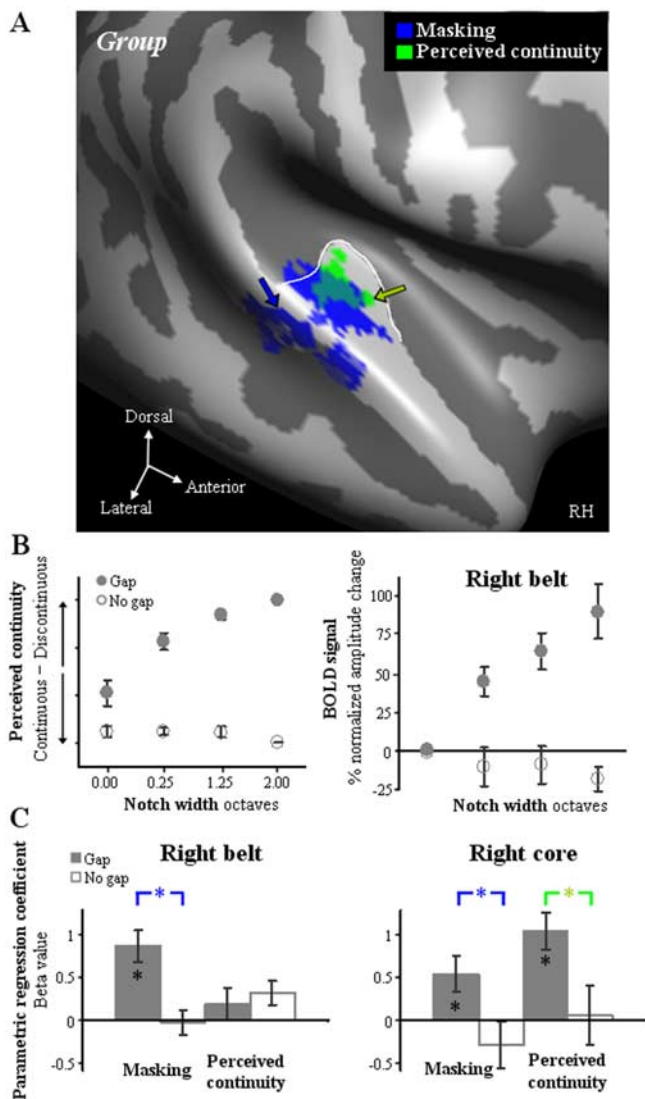


Figure 2. Neuroimaging and behavioral group results. *A*, Parametric activation maps projected onto the inflated cortical surface of the right hemisphere (RH) after cortical alignment across listeners. Activity in core and adjacent belt areas correlated with masking (blue regions) and perceived continuity (green regions) for only interrupted targets (random-effect analysis, corrected $p < 0.05$). The white line delineates the location of Heschl's gyrus. *B*, Perceived continuity (left) and BOLD signal changes (right) in belt areas (*A*, blue arrow) showed parametric effects of masking for interrupted (solid circles), but not for uninterrupted (open circles) targets across listeners. Circles and bars represent mean value and SEM across listeners. To allow for comparison between interrupted and uninterrupted targets, mean BOLD signals were normalized to the respective no-notch conditions. *C*, Parametric changes in masking had significant effects on BOLD responses to interrupted (solid rectangles), but not to uninterrupted (open rectangles) targets in core and belt areas. In the core areas (*A*, green arrow), perceived continuity of only interrupted targets significantly explained additional variance in BOLD responses. Rectangles and bars represent mean value and SEM of the parametric regression coefficients across listeners. Asterisks indicate significance at $p < 0.05$. The differential effects of masking (blue asterisks) and perceived continuity (green asterisk) defined the activated brain regions shown in *A*.

lated to acoustic stimulus changes correlated significantly with activity in a lateral portion of the right HG for only interrupted targets (rating by gap interaction, $t_{10} = 2.6$; cluster size, >50 ; corrected $p < 0.05$) (Fig. 2*A*, green regions) (ROI analysis for interrupted and uninterrupted targets, $t_{10} = 4.9$ and -0.6 ; $p = 0.0005$ and 0.5 , respectively) (Fig. 2*C*, ROI locations are as indicated in *A*, supplemental Fig. 1 *B*, green arrows, available at www.jneurosci.org as supplemental material).

Comparable effects were also observed in the left hemisphere ($t_{10} = 3.5$; cluster size, >100 ; corrected $p < 0.05$) (supplemental Fig. 1 *B*, available at www.jneurosci.org as supplemental material). As these changes in the BOLD responses were not related to stimulus-driven changes, they likely reflected spontaneous changes in stimulus interpretation. Note that these BOLD response changes were largest for the most illusory percepts (interrupted targets, no-notch condition) (supplemental Fig. 3, available at www.jneurosci.org as supplemental material).

Best-frequency mapping

Best-frequency mapping identified a low frequency-selective cluster on HG that was surrounded by high frequency-selective clusters in more medial, rostral, and caudal regions (Fig. 3). This result is compatible with results from previous human tonotopy studies that support the notion of multiple, mirror-symmetric primary ACFs along HG (Formisano et al., 2003; Talavage et al., 2004). Projection of the cluster that we found to covary with listeners' perceived continuity (Fig. 2*A*, green region) onto the best-frequency map (Fig. 3, green outline) revealed that this region overlapped with the low-frequency selective region in HG, suggesting its inclusion within the ACFs of the primary core. The other cluster that we found to covary with the broadband masker for only interrupted targets was located in a more lateral region in STG (Fig. 2*A*, lateral blue region), which may correspond to the ACFs of the secondary belt (Hackett et al., 2001; Sweet et al., 2005).

Discussion

Our finding that masking parametrically affects BOLD responses in the putative human core and belt areas for only interrupted

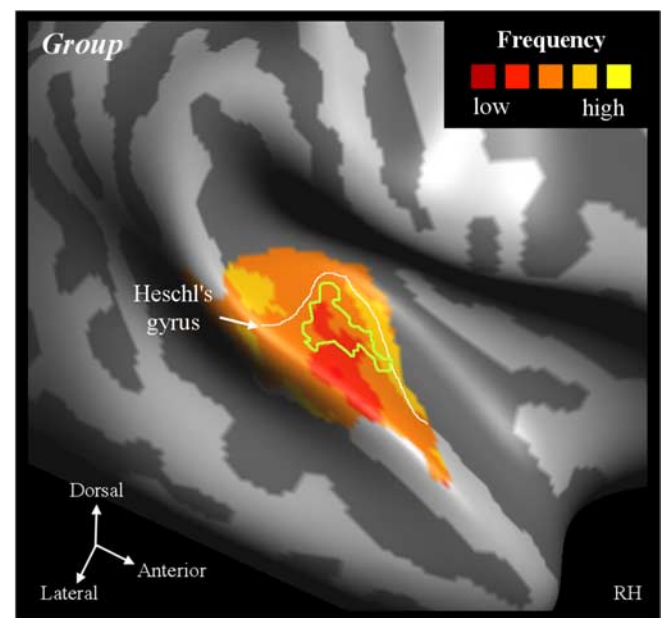


Figure 3. Best-frequency mapping group results. Best-frequency group maps projected onto the inflated cortical surface of the right hemisphere (RH) after cortical alignment across listeners, showing cortical clusters of different frequency selectivity (corrected $p < 0.05$; $F5 > 0.29$; red, low frequencies; yellow, high frequencies). The region that was found to covary with listeners' perceived continuity (Fig. 2*A*, green region) was projected (green outline) onto the frequency selectivity map, revealing partial overlap with a low-frequency cluster in lateral Heschl's gyrus that may be part of the human primary core (Hackett et al., 2001; Morosan et al., 2001; Rademacher et al., 2001; Sweet et al., 2005) (for details, see Results). The gray line delineates the location of Heschl's gyrus.

targets suggests that neural activations in these areas do not reflect masking per se, but rather the salience of on-frequency band changes, which depends on the level of masking. Thus, these modulations point to perceptual differences between the two stimuli (i.e., perceived continuity differences, which varied across masking levels) rather than to spectral-temporal differences between interrupted and uninterrupted targets (i.e., differences in the gap, which were constant across masking levels). In line with current views on auditory object formation (Griffiths and Warren, 2004), we interpret these effects as reflecting the extraction of abstract relations between the basic spectral-temporal properties of targets and maskers and their transformation into perceptual and task-related representations. This interpretation is consistent with a previous electrophysiological study of the continuity illusion which suggested that neurons in the monkey core represent tones and interrupting noise bursts as if these were integrated (Petkov et al., 2007). The result that discontinuity percepts evoke stronger BOLD responses than continuity illusions suggests that the observed activities primarily reflect the perception of onsets of new auditory objects. This interpretation is consistent with recent evidence from human neuroimaging studies (Mustovic et al., 2003; Warren et al., 2005; Herdener et al., 2007; Wilson et al., 2007) that used nonillusory stimuli.

The finding that percept changes without concomitant stimulus changes also correlate with BOLD responses in lateral HG is consistent with previous evidence from studies on auditory streaming, a form of perceptual grouping that may be related to the continuity illusion (Bregman et al., 1999). Perceived changes in streaming in the absence of stimulus changes were reported to correlate with neuromagnetic scalp signals that likely arise from lateral HG (Gutschalk et al., 2005), and with single-cell responses in monkey core (Michey et al., 2005). During perceptual switches, these and other (Cusack, 2005) brain regions may be modulated by top-down processes. Similar regions have been shown to be involved in the processing of illusory pitch (Patel and Balaban, 2001; Bendor and Wang, 2005), especially in the right hemisphere (Zatorre, 1988; Zatorre et al., 2002).

Regarding the precise functional localization of our effects, the results from the best-frequency mapping suggest that listeners' ratings of perceived continuity are included in a region that is most likely part of the human auditory core. However, the impossibility to compute the tonotopic gradient and the exact borders between the ACFs in the present data (related to the small number of frequencies and the group analysis) prevents specifying of particular core field(s). Similar conclusions can be drawn by comparing our results with previous cytoarchitecture studies that reported that the human homolog of the primary core is likely located in HG (Hackett et al., 2001; Morosan et al., 2001; Sweet et al., 2005), although the exact locations of primary area borders may vary in or around the gyrus (Rademacher et al., 2001). In fact, the location of the region that we found to covary with listeners' perceived continuity appears in good agreement with the cytoarchitecturally defined human primary core area Te1.0 in HG (Morosan et al., 2001) (supplemental Fig. 1*B*, available at www.jneurosci.org as supplemental material). Similarly, the regions that we localized in more lateral regions on STG probably coincide with the human secondary belt which may surround the core (Hackett et al., 2001; Sweet et al., 2005). Together, our physiological and anatomical considerations suggest that the detected regions in HG and in STG (Fig. 2*A*) are likely located in primary core and secondary belt, respectively.

It should be noted that our analyses complemented each other in differentiating two different types of cortical processing asso-

ciated with our auditory task: the stimulus-based analysis identified cortical regions involved in the perceptual processing of abstract spectral-temporal stimulus properties. The percept-based analysis, however, revealed regions involved in the perceptual processing of cognitive and stimulus-unrelated factors. This differentiation was enabled by using the orthogonalization procedure (see Materials and Methods). Together our results indicate that (1) the activity of early auditory cortical regions in humans does not correlate with the basic acoustic properties of tones or noise bursts alone, but also with the complex dependencies among them, and that (2) changes of neural responses in a core region (lateral HG) to acoustically identical sounds parallel listeners' reports of perceived continuity to these same sounds.

Together with the consideration that the continuity illusion is disrupted by removal of the target after the noise (Bregman, 1990), our findings suggest a two-stage model for cortical processing of stimuli evoking the continuity illusion. At a first stage, neuronal populations in the core and belt areas entail a representation of the auditory scene (tone plus noise) based on abstract dependencies between the individual scene segments, such as spectral edges between the tone and the noise or changes in the on-frequency band. The result of this stage may be regarded as the formulation of a sensory-constrained perceptual hypothesis (Bregman, 1990), with the two possible outcomes of segregation of the scene into two (one continuous tone and noise) or three (two discontinuous tones and noise) perceptual objects. This hypothesis is resolved at a second stage in the core, in which the representation of the scene is re-evaluated based on the sensory evidence provided by the postnoise interval. Whereas the first processing stage is possibly fast and automatic, the second may be influenced by contextual, attentional and task-related processes, thus explaining the covariance of neural and behavioral responses to the same illusory stimuli. The idea of a duplex analysis of the auditory scene in early auditory areas is in line with recent electrophysiological evidence in cats showing that the same primary AC neurons can represent stimuli at multiple time scales (Nelken, 2004; Ulanovsky et al., 2004). The latter may provide the neural basis for an acoustic short-term buffer that would be required for retention of the perceptual hypotheses in our model. Consistently, the mismatch-negativity results from a previous study (Michey et al., 2003) suggest that the analysis of changes in the target on-frequency band recruits a buffer in auditory cortical areas.

In conclusion, our findings provide strong evidence that beyond sensory representation of an auditory scene, primary auditory cortical areas also play a constructive role in the grouping of scene segments into unified percepts. Furthermore, they put forward a model for neural computation of auditory scenes that may inform analyses of physiological data obtained at high temporal resolution, such as with invasive or noninvasive electrophysiology and magnetoencephalography.

References

- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993) Processing strategies for time-course data sets in functional MRI of the human brain. *Magn Reson Med* 30:161–173.
- Belin P, Zatorre RJ, Hoge R, Evans AC, Pike B (1999) Event-related fMRI of the auditory cortex. *Neuroimage* 10:417–429.
- Bendor D, Wang X (2005) The neural representation of pitch in primate auditory cortex. *Nature* 436:1161–1165.
- Braaten RF, Leary JC (1999) Temporal induction of missing birdsong segments in Eur starlings. *Psychol Sci* 10:162–166.
- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: MIT.

- 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.
- Cusack R (2005) The intraparietal sulcus and perceptual organization. *J Cogn Neurosci* 17:641–651.
- 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.
- Gaab N, Gabrieli JDE, 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 Brain Voyager 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.
- Griffiths TD, Warren JD (2004) What is an auditory object? *Nat Rev Neurosci* 5:887–892.
- Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci* 25:5382–5388.
- Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 411:197–221.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) “Sparse” temporal sampling in auditory fMRI. *Hum Brain Mapp* 7:213–223.
- Hall DA, Haggard MP, Akeroyd MA, Summerfield AQ, Palmer AR, Elliott MR, Bowtell RW (2000) Modulation and task effects in auditory processing measured using fMRI. *Hum Brain Mapp* 10:107–119.
- Herdener M, Esposito F, Di Salle F, Lehmann C, Bach DR, Scheffler K, Seifritz E (2007) BOLD correlates of edge detection in human auditory cortex. *Neuroimage* 36:194–201.
- 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.
- Jäncke L, Wüstenberg T, Scheich H, Heinze HJ (2002) Phonetic perception and the temporal cortex. *Neuroimage* 15:733–746.
- 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.
- Miller CT, Dibble E, Hauser MD (2001) Amodal completion of acoustic signals by a nonhuman primate. *Nat Neurosci* 4:783–784.
- 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.
- Mustovic H, Scheffler K, Di Salle F, Esposito F, Neuhaus JG, Hennig J, Seifritz E (2003) Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *Neuroimage* 20:429–434.
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) “Primitive intelligence” in the auditory cortex. *Trends Neurosci* 24:283–288.
- Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480.
- Patel AD, Balaban E (2001) Human pitch perception is reflected in the timing of stimulus-related cortical activity. *Nat Neurosci* 4:839–844.
- Petkov CI, O’Connor KN, Sutter ML (2003) Illusory sound perception in macaque monkeys. *J Neurosci* 23:9155–9161.
- Petkov CI, O’Connor KN, Sutter ML (2007) Encoding of illusory continuity in primary auditory cortex. *Neuron* 54:153–165.
- 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 Opstal AJ, Formisano E (2007) The auditory continuity illusion: a parametric investigation and filter model. *Percept Psychophys*, in press.
- 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.
- Sugita Y (1997) Neuronal correlates of auditory induction in the cat cortex. *NeuroReport* 8:1155–1159.
- Sweet RA, Dorph-Petersen KA, Lewis DA (2005) Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol* 491:270–289.
- Talavage TM, Edmister WB (2004) Nonlinearity of fMRI responses in human auditory cortex. *Hum Brain Mapp* 22:216–228.
- Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM (2004) Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *J Neurophysiol* 91:1282–1296.
- Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24:10440–10453.
- Van Atteveldt N, Formisano E, Goebel R, Blomert L (2004) Integration of letters and speech sounds in the human brain. *Neuron* 43:271–282.
- 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.
- Wilf HS (1962) Mathematics for the physical sciences. New York: Dover.
- Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical fMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. *J Neurophysiol* 97:2230–2238.
- Zatorre RJ (1988) Pitch perception of complex tones and human temporal-lobe function. *J Acoust Soc Am* 84:566–572.
- Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37–46.