

Adaptation in the sensory cortex drives bistable switching during auditory stream segregation

Nathan C. Higgins^{1,†}, Alexandra N. Scurry², Fang Jiang², David F. Little³, Claude Alain^{4,‡}, Mounya Elhilali^{3,§}, Joel S. Snyder^{5,¶}

¹Department of Communication Sciences and Disorders, University of South Florida, 4202 E. Fowler Avenue, PCD1017, Tampa, FL 33620, USA; ²Department of Psychology, University of Nevada, 1664 N. Virginia Street Mail Stop 0296, Reno, NV 89557, USA; ³Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA; ⁴Rotman Research Institute, Baycrest Health Sciences, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada; ⁵Department of Psychology, University of Nevada, 4505 Maryland Parkway Mail Stop 5030, Las Vegas, NV 89154, USA

[†]Nathan C. Higgins, <http://orcid.org/0000-0002-1219-6448>

[‡]Claude Alain, <http://orcid.org/0000-0003-4459-1538>

[§]Mounya Elhilali, <http://orcid.org/0000-0003-2597-738X>

[¶]Joel S. Snyder, <http://orcid.org/0000-0002-5565-3063>

*Correspondence address. Department of Communication Sciences and Disorders, University of South Florida, 4202 E. Fowler Ave., Tampa, FL 33620, USA.

Tel: +860-301-9112; E-mail: higgins1@usf.edu

Abstract

Current theories of perception emphasize the role of neural adaptation, inhibitory competition, and noise as key components that lead to switches in perception. Supporting evidence comes from neurophysiological findings of specific neural signatures in modality-specific and supramodal brain areas that appear to be critical to switches in perception. We used functional magnetic resonance imaging to study brain activity around the time of switches in perception while participants listened to a bistable auditory stream segregation stimulus, which can be heard as one integrated stream of tones or two segregated streams of tones. The auditory thalamus showed more activity around the time of a switch from segregated to integrated compared to time periods of stable perception of integrated; in contrast, the rostral anterior cingulate cortex and the inferior parietal lobule showed more activity around the time of a switch from integrated to segregated compared to time periods of stable perception of segregated streams, consistent with prior findings of asymmetries in brain activity depending on the switch direction. In sound-responsive areas in the auditory cortex, neural activity increased in strength preceding switches in perception and declined in strength over time following switches in perception. Such dynamics in the auditory cortex are consistent with the role of adaptation proposed by computational models of visual and auditory bistable switching, whereby the strength of neural activity decreases following a switch in perception, which eventually destabilizes the current percept enough to lead to a switch to an alternative percept.

Keywords: bistable perception; auditory cortex; adaptation; auditory stream segregation

Introduction

Most theoretical and empirical studies on the neural basis of conscious perception have traditionally been focused on the visual modality (Crick and Koch 1995; Tong et al. 2006; Koch and Tsuchiya 2007; Melloni et al. 2021) and to a lesser extent on the auditory modality (Snyder et al. 2012; Dykstra et al. 2017). Although there are many theories about the neural basis of conscious perception, they mostly agree that it arises at least in part from activity in sensory cortical pathways. For example, upon becoming subjectively aware of a familiar face or voice in a crowded social gathering, feed-forward and feedback activity in a ventral cortical pathway might be sufficient to explain the conscious percept (Lamme and Roelfsema 2000; Hochstein and Ahissar 2002; DiCarlo et al. 2012). However, alternate theories say that while such activity in ventral areas is “necessary” for conscious perception, it is not “sufficient” (Dehaene et al. 2003; Dehaene and Changeux 2011; Lau and Rosenthal 2011; Brown et al. 2019); such theories propose

that activity in frontal and/or parietal areas is also necessary for conscious perception to occur. Another class of theories takes the position that any cortical network that has certain computational properties can give rise to conscious perception (Tononi et al. 1994, 2016).

Different neurophysiological techniques have been used to study the neural correlates of conscious perception in both human and non-human primates. For example, studies of binocular rivalry in old-world monkeys showed that the proportion of neurons that change their firing rate around the time of perceptual switches is especially large in higher-level visual areas in the ventral stream, compared to mid-level areas (V4 and middle temporal) and early visual cortex (V1/V2) (Logothetis and Schall 1989; Leopold and Logothetis 1996; Sheinberg and Logothetis 1997). Consistent with these findings, event-related brain potential (ERP) studies in humans show responses to reversals in perception of bistable stimuli (Tononi et al. 1998; Kornmeier and Bach 2005; Pitts et al. 2007, 2008, 2010) that likely arise from mid- and high-level

Received: 25 July 2022; Revised: 17 October 2022; Accepted: 26 December 2022

© The Author(s) 2023. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

visual cortical areas in the ventral stream (Pitts et al. 2009). Evidence from functional magnetic resonance imaging (fMRI) studies in humans supports the importance of low-level visual thalamic and sensory areas in resolving perceptual ambiguity (Polonsky et al. 2000; Tong and Engel 2001; Lee et al. 2005; Meng et al. 2005; Wunderlich et al. 2005), as well as higher-level visual areas and frontal and parietal areas (Kleinschmidt et al. 1998; Tong et al. 1998; Sterzer et al. 2002; Sterzer and Rees 2008). One limitation of prior studies is that they seek to identify neural correlates of the contents of conscious perception, without probing the mechanisms reflected by switches in conscious perception. For example, one objection to observations of frontal and parietal activations during changes in conscious perception is that they reflect pre-perceptual mechanisms such as attention or post-perceptual mechanisms such as decision-making and motor preparation (Aru et al. 2012). No-report paradigms in which participants' conscious perception is determined without requiring decision-making and motor responses during the key perceptual experience support the notion that frontal and parietal areas may indeed not be directly involved in determining conscious perception (Shafto and Pitts 2015; Cohen et al. 2020).

Another limitation is that most studies of conscious perception have been performed using visual stimuli, and resulting theories have largely disregarded evidence from other sensory modalities. Yet, there are a growing literature on conscious auditory perception (Snyder et al. 2012; Dykstra et al. 2017) and mounting evidence that mechanisms may be similar across the visual and auditory domains (Pressnitzer and Hupé 2006; Hupé et al. 2008; Davidson and Pitts 2014; Snyder et al. 2015a,b; Higgins et al. 2021). One type of auditory-based paradigm presents an isochronous stream of target tones that participants are asked to detect among a "cloud" of distractor tones. Such studies show that successful detection of the target tones evokes larger responses from 50 to 250 ms, compared to when the targets are not detected (Gutschalk et al. 2008; Königs and Gutschalk 2012; Wiegand and Gutschalk 2012; Dykstra et al. 2016); shorter-latency neural activity in the primary auditory cortex meanwhile is evoked similarly regardless of detection success (Gutschalk et al. 2008). Studies of auditory stream segregation typically present alternating patterns of low and high tones that can be perceived as a single stream of tones (low-high-low-high ...) or as two segregated streams (low-low ... and high-high ...). With longer sequences of the tone patterns, perception spontaneously switches back and forth in a bistable fashion (Denham and Winkler 2006; Pressnitzer and Hupé 2006). Larger frequency separations between the low and high tones result in a greater likelihood of perceiving segregated (two streams) and larger responses in the auditory cortex (Gutschalk et al. 2005; Snyder et al. 2006; Gutschalk et al. 2007; Wilson et al. 2007; Snyder et al. 2009b; Weintraub et al. 2012). Likewise, when a moderate frequency separation is presented, making the pattern ambiguous, auditory cortex responses were larger when the percept was segregated compared to integrated (Gutschalk et al. 2005; Hill et al. 2011; Billig et al. 2018; Curtu et al. 2019), and activity from the auditory cortex enabled above-chance classification of the reported percepts (Billig et al. 2018; Sanders et al. 2018; Curtu et al. 2019). Another study, however, only found increased activity for segregated compared to integrated in the intraparietal sulcus (Cusack 2005), while a more recent study found that measures of integration and differentiation in frontal-parietal networks could distinguish perception of integrated versus segregated (Canales-Johnson et al. 2020). During switches from integrated to segregated or vice versa, studies have found activations in subcortical and cortical areas, including inferior colliculus,

medial geniculate nucleus of the thalamus, and auditory cortex (Kondo and Kashino 2009; Schadwinkel and Gutschalk 2011). A study of bistable switching in a verbal transformation task found activity in the left inferior prefrontal cortex and the anterior cingulate around the time of switches (Kondo and Kashino 2007). Similarly, perceived pitch change direction has been associated with not only an early brain response that likely comes from the auditory cortex but also a later brain response hypothesized to come from wider frontal and parietal areas, although source analysis was not performed to verify this (Davidson and Pitts 2014). Together, these studies highlight the role of a neural network that could be critical to auditory perception, including subcortical areas, modality-specific areas, and supramodal areas.

A promising approach is to consider the process of how bistable perception switches occur in sensory systems at a more mechanistic level. Conceptual and computational models of visual bistability have proposed several interrelated mechanisms that can be tested empirically using behavioral and neurophysiological data. These mechanisms include grouping of features into larger structures, adaptation of neural representations of stimulus features and/or higher-level percepts, competitive inhibition of alternative interpretations, attention, and neural noise (Wilson 2003; Tong et al. 2006; Noest et al. 2007; Brascamp et al. 2008; Grossberg et al. 2008; Li et al. 2017). These mechanisms have recently been used to explain auditory bistable perception using a model of the primary auditory cortex like tonotopic maps (Rankin et al. 2015, 2017; Rankin and Rinzel 2019). Few studies, however, have provided direct empirical evidence in support of inhibitory mechanisms operating during bistable perception, with the exception of studies of magnetic resonance spectroscopy measures of excitatory and inhibitory neurotransmitters (van Loon et al. 2013; Kondo et al. 2018).

Two recent papers from our groups provide valuable context for the present study. We published a new computational model that embodies the general mechanisms of adaptation, inhibition, and noise described earlier and applies it to the auditory system (Little et al. 2020). The model included a peripheral stage of processing with an array of frequency-tuned neuron-like units, a primary auditory cortex with units having different frequency bandwidths, and a secondary auditory cortex stage with units reflecting different interpretations of integrated or segregated streams of tones. Each of the three stages of processing incorporated competitive inhibition between units, adaptation of active units, and noise fluctuations. A range of parameter settings for each of these mechanisms was used to determine the best settings to accurately reproduce human behavioral patterns of percept switching and percept durations. The findings showed that a larger range of parameter settings enabled good fits with human behaviors at the secondary auditory cortex stage of processing, with fewer sets of parameters that fit human data at the lower levels of processing. This general framework of adaptation, inhibition, and noise can account for auditory bistability, just as it does for visual bistability. It also highlights the role of higher-level auditory processes in determining bistability. The model also inspires the search for empirical evidence supporting the importance of adaptation, inhibition, and noise in determining bistable perception.

As a first step, we conducted an ERP study that intermittently presented three repetitions of a low-high-low pattern of sounds, followed by a brief silent period for participants to respond with how they perceived the three tone triplets (Higgins et al. 2020). This allowed us to remove some of the ambiguity in associating

particular cortical responses to sounds with particular perceptual interpretations. We found converging evidence with the model by showing that a sustained negative response that arises from a secondary auditory cortex region in the ventral auditory processing pathway is enhanced during segregated perception. Dipole modeling showed that the sustained response also had a small parietal source, in addition to the main ventral auditory cortex source. We also found this response to be enhanced following switches from one percept to the other. This latter finding is consistent with the role of neural adaptation. Following a switch in perception, neurons processing the new percept should be most active, followed by a gradual decrease in their activity because of neural adaptation.

The current study further tests the general framework of adaptation, inhibition, and noise—and the specific computational model presented in [Little et al. \(2020\)](#)—using fMRI. We use the same intermittent stimulus presentation as in our ERP study and measure activity across the whole brain (including in the auditory thalamus) during periods of integration, during periods of segregation, after switches from integrated to segregated, and after switches from segregated to integrated. We also track the dynamics of brain activity in the auditory cortex following perceptual switches and leading up to perceptual switches, which allow us to provide additional evidence for the role of neural adaptation in the auditory cortex in bistable perception.

Methods

Ethics statement and participants

All research techniques and procedures were approved by the University of Nevada, Reno Institutional Review Board, and the University of Nevada, Las Vegas Institutional Review Board, and aimed to include representative human populations in terms of sex, age, and ethnicity. Prior to the experiment, all participants provided written informed consent followed by a standard hearing screening to ensure that audiometric thresholds did not exceed 25 dB hearing level. Thirteen normal-hearing adults (8 males) with an average age of 30.1 (SD = 7.3) years were recruited from the community in and around the University of Nevada, Reno.

Intermittent response paradigm

As shown in [Fig. 1](#), participants were presented with repeating A and B tones, organized into 650-ms triplets with an ABA₁ configuration, where A and B were pure tones at 400 and 565.5 Hz, and the ₁ represents a missing B tone with equivalent duration. A and B tones were 75-ms duration, presented with 162.5-ms interstimulus interval for a total of 650 ms for each triplet. Each trial consisted of three consecutively presented triplets followed by 650-ms silence during which participants indicated their perception with a button press: Button 1 for integrated (one stream) and

Button 2 for segregated (two streams) in a two-alternative forced-choice paradigm. The three triplets and the equivalent-duration silent-response period had a total duration of 2.6 s for each trial.

Stimulus familiarization

Prior to the imaging session, the stimulus paradigm was explained to each participant. They were presented with variations of the paradigm that consisted of exaggerated large (12 semitones) or small (three semitones) separation between the A and B tones to provide familiarity with both the segregated and integrated perceptions. Then, they completed one 152-trial practice session with intermediate frequency separation (six semitones) where they indicated their perception with a button press for each trial. Stimuli were presented outside the fMRI scanner over headphones for the familiarization and practice periods.

Stimulus presentation

Acoustic stimuli were generated and presented using custom MATLAB (Mathworks) and PsychToolBox routines and delivered via insert earphones (Sensimetrics) enclosed in circumaural ear defenders at a sound level of 75 dB at a sampling rate of 44 100 Hz. Trials were grouped into consecutively presented segments of 38 trials; four segments made up one full presentation run of 152 trials (~7-min duration). Each of the four segments was separated by silent “blank trials” equal to the duration of three trials (7.8 s). In total, five presentation runs were presented to each participant with a short rest between each.

Definition of perceptual events

As described earlier, each run of data collection consisted of three segments of 38 sequential trials. Based on participant responses, each trial was designated as an integrated or segregated percept and as a switch or no-switch. Each switch was further designated as a switch from integrated to segregated or segregated to integrated. Perceptual phases of a continuous percept were defined by the number of trials between switches, and each trial within a phase was designated by its sequential location within the phase (i.e. first, second, third, third-to-last, second-to-last, and last trials). The first trial of a perceptual phase was designated a switch trial; all others were designated as stable, no-switch trials. Perceptual phases that encompassed an entire stimulus segment were excluded due to concern that no switching activity indicated lack of attention by the participant.

fMRI data acquisition

Scanning was performed at the Neuroimaging Facility of Renown Health Hospital in Reno, NV, on a 3T Philips Ingenia scanner using a 32-channel digital SENSE head coil (Philips Medical Systems, Best, Netherlands). Data were collected over the course of a single imaging session per participant, including a high-resolution

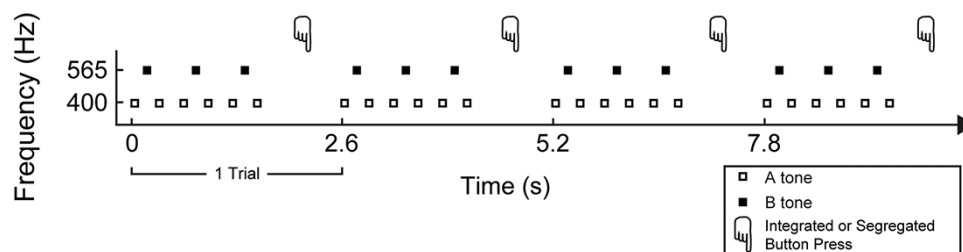


Figure 1. The intermittent streaming paradigm. Three consecutive ABA triplets along with a behavioral response constitute a single trial in this paradigm

T1-weighted whole-brain structural scan (MPRAGE, 1-mm³ isotropic voxels), and five echo-planar functional scans, each lasting 340 s. Functional imaging consisted of a continuous acquisition paradigm with a repetition time of 2 s, echo time of 35 ms, and slice thickness of 3 mm for a voxel size of 2.75 × 2.75 mm² over 40 slices.

fMRI data preprocessing and analysis

fMRI data were processed using fMRI Expert Analysis Tool (FEAT) (FSL 6.0; FMRIB) and consisted of high-pass filtering (100 s), motion correction, and skull stripping (Woolrich et al. 2001). An event-related general linear model was implemented in FEAT by using the onset of each trial as an event. Here, the trial types corresponded to sound onsets or the silent onsets (or blank trials in between segments). The results were used to identify voxels with a significant response to sound compared to silence (initial cluster definition threshold $Z > 2.3$, corrected $P < 0.05$). Results from this analysis were used to create a voxel-based sound-sensitive mask used in the adaptation analyses.

Following initial preprocessing steps, custom MATLAB scripts were used to run a parallel analysis where the blood oxygenation level-dependent (BOLD) time course for each voxel was temporally interpolated and normalized to Z-score values. Linear regression of a custom hemodynamic response function (HRF) was used to estimate single-trial responses. The HRF was defined as the difference of two gamma density functions adapted from Glover (1999) with the onset of each triplet treated as an event. The following parameters defined the custom HRF: 4.5 (peak 1), 5.2 [full width at half maximum FWHM 1], 10.8 (peak 2), 7.35 (FWHM 2), 0.15 (dip) where 1 and 2 correspond to the first and second gamma functions and FWHM defines the full width at half-maximum of each function. The resulting beta weight from the linear regression quantified the evoked BOLD response for each sound trial, comprised the functional 3D dataset for each run for each participant, and retained the sequential order of voxel responses used for adaptation analyses.

Regions of interest and cortical surface topography

For each participant and for each run, the high-resolution T1-weighted anatomical image was used with FreeSurfer software (<http://surfer.nmr.mgh.harvard.edu/>) to register 3D functional datasets to common spherical anatomical space using curvature-based alignment across participants. From that point, parallel processing pipelines were used to (i) project functional data to 2D cortical surface or (ii) isolate voxels specific to cortical regions of interest (ROIs) in 3D space.

With respect to Processing Pipeline 1, functional (trial-based) data were retained through the transformation to the cortical surface for statistical analyses. Following surface projection, topography was smoothed with a resolution of 10-mm FWHM and flattened to a 500 × 1000-pixel map using the Mollweide projection (Woods et al. 2009).

For Processing Pipeline 2, ROIs (i.e. voxels within ROIs) corresponding to the auditory cortex, rostral anterior cingulate, and inferior parietal lobe (IPL) were isolated via FreeSurfer and the Desikan et al.'s (2006) atlas. An additional ROI representing the auditory thalamus was defined using FreeSurfer's automatic sub-cortical segmentation of brain volume (Fischl et al. 2002). To isolate the auditory thalamus at the individual participant level, left and right hemisphere full-thalamus ROIs were each partitioned at the

midpoint of the medial–lateral dimension to create sections representing the lateral and medial thalami for right and left hemispheres. The medial partition, as the highest probability location of the medial geniculate nucleus, was retained to represent the auditory thalamus. The Desikan–Killiany atlas defines the superior temporal gyrus (STG) as a single ROI. Anterior and posterior STGs were designated as separate ROIs based on the boundary defined by the intersection with Heschl's gyrus (HG) (McLaughlin et al. 2016).

Adaptation analysis

Data used for adaptation analyses were selected using two separate cortical masks. The first mask identified ROIs that contain auditory cortex, specifically, HG and anterior and posterior sections of STG. The second mask used the sound–silence contrast to identify voxels within the auditory cortex for each subject that was sensitive to sound. The time course (i.e. the sequential order of trials) for each of these voxels was further analyzed as a function of perceptual phase duration and the sequential order of trials within each perceptual phase.

Results

Behavioral response patterns

Participant response patterns were consistent with those observed previously using a similar intermittent response paradigm (Higgins et al. 2020) or using traditional characterizations of stimuli that elicit bistable perception (e.g. Pressnitzer and Hupe, 2006; Denham et al. 2018). Perceptual responses at the beginning of trials were typically integrated, but average data across uninterrupted segments of the 38 trials within a segment converged toward roughly equal probability of integrated versus segregated perception (Fig. 2A). Phase duration of the integrated percept was positively correlated with duration of the segregated percept, such that individuals who reported longer phases for integrated also reported longer phases of segregated as well (Fig. 2B; $r^2 = 0.51$, $P = 0.006$). Another hallmark of perceptual bistability is a minimal-to-small relationship between a given phase duration and the next phase duration (Pressnitzer and Hupé 2006; Barniv and Nelken, 2015). This dynamic is observed here, whereby the phase duration of phase (N) is minimally (though significantly) correlated with the subsequent phase ($N + 1$) (Fig. 2C; $r^2 = 0.023$, $P < 0.001$). Finally, the shape of the distribution of phase durations was fit with log-normal and gamma functions. Relative to the observed distribution, the residuals for the log-normal function were significantly smaller than the gamma function ($t_{40} = 4.67$, $P < 0.001$, $d = 0.30$), indicating a better fit by the log-normal function, consistent with previous studies (Lehky 1995; Denham et al. 2018). This pattern of behavioral results confirms the conclusion of Higgins et al. (2020) that the intermittent response paradigm evokes bistable perception.

Cortical surface maps: sound sensitivity

The Mollweide projections of the cortical surface curvature averaged across all participants are presented in Fig. 3A for right and left hemispheres. Key regions corresponding to the auditory cortex (HG and STG), IPL, and cingulate gyrus (CG) used in later analyses are highlighted. In Fig. 3B, surface voxels from each participant with a significant response to sound summed together across common cortical areas illustrate regions of cortex most sensitive to the ABA-triplet stimuli compared to silence and correspond to the individual voxel masks used in later analyses. In both left and right hemisphere regions, areas corresponding to

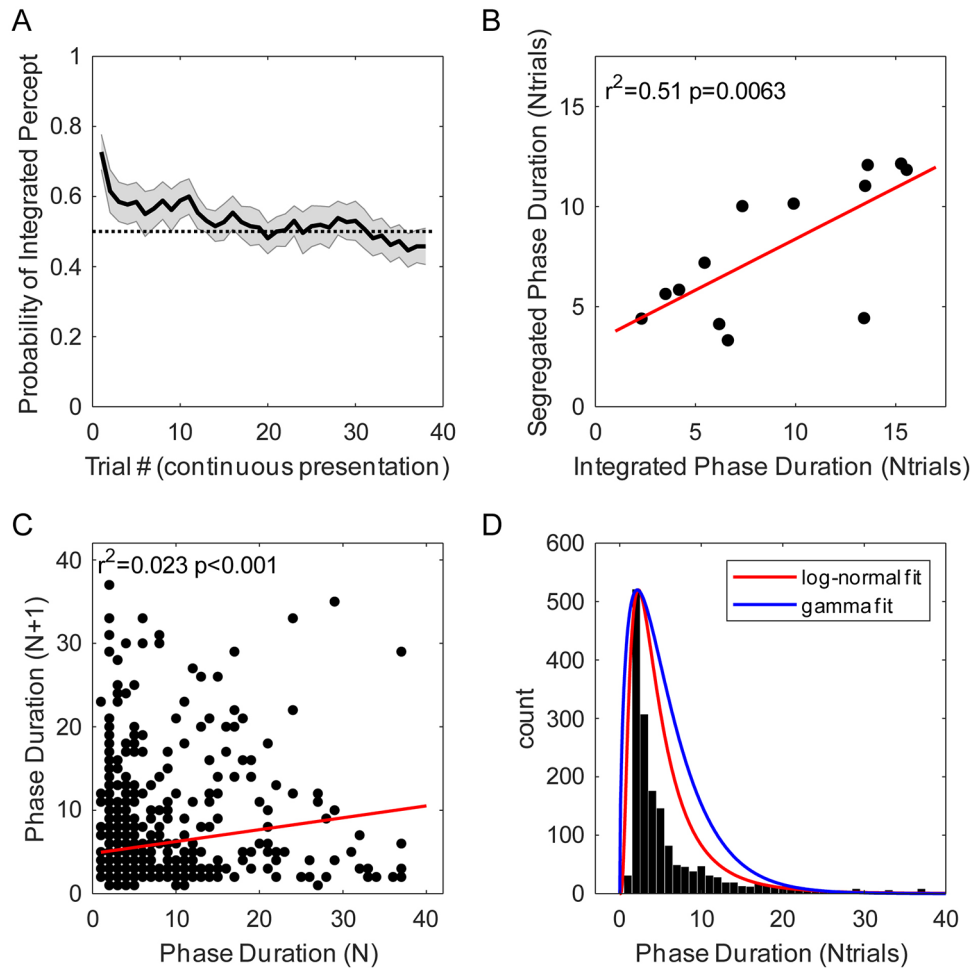


Figure 2. Behavioral data. (A) The time course of the probability of perceiving an integrated percept. (B) A scatterplot and linear regression showing a correlation between durations of integrated and segregated percepts. (C) The relationship between the phase duration of each percept versus the next sequential percept; the line represents the best-fit linear regression. (D) The histogram of percept durations and fits for log-normal and gamma functions

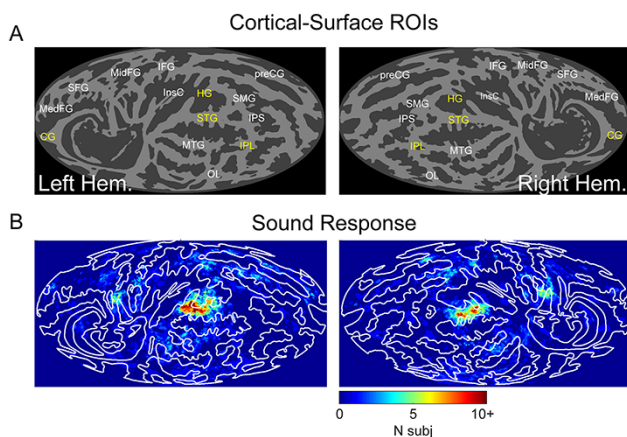


Figure 3. (A) Surface projections of cortical anatomy in common space, averaged across subjects for left and right hemispheres. Light gray regions indicate gyri, and dark gray regions correspond to sulci. (B) Functional surface topography of significant sound-sensitive surface voxels summed across participant. Anatomical labels: CG, cingulate gyrus; HG, Heschl's gyrus; InsC, insular cortex; IFG, inferior frontal gyrus; IPL, inferior parietal lobe; IPS, inferior parietal sulcus; MedFG, medial frontal gyrus; MidFG, mid-frontal gyrus; MTG, middle temporal gyrus; OL, occipital lobe; preCG, pre-central gyrus; SFG, superior frontal gyrus; SMG, supramarginal gyrus; STG, superior temporal gyrus

lateral HG and posterior STG (or planum temporale) were most consistently active across participants. Note that this is a full brain analysis, and negligible sound sensitivity was observed outside of the auditory cortex.

Cortical surface maps: switch versus no-switch

Beta weights averaged across participants for all no-switch trials in the left and right hemispheres were slightly higher around the auditory cortex region compared to silence (Fig. 4A) but were relatively low compared to the switch trials (Fig. 4B). The beta weight topography for switch trials had a much higher range with peaks around the auditory cortex, inferior frontal gyrus, and medial frontal gyrus but minima around the IPL and CG. Comparison between maps for no-switch and switch trials thresholded using random field theory ($\alpha=0.05$) (Brett *et al.* 2003) and only revealed significant differences for clusters of surface voxels in regions corresponding to the CG and the IPL in both left and right hemispheres (Fig. 4C; red outlined panels).

Cortical surface maps: integrated percept versus switch from segregated to integrated

To separate out the contributions of different switch types, a comparison between trials where a stable-integrated percept was reported was compared to the initial trial of a perceptual phase

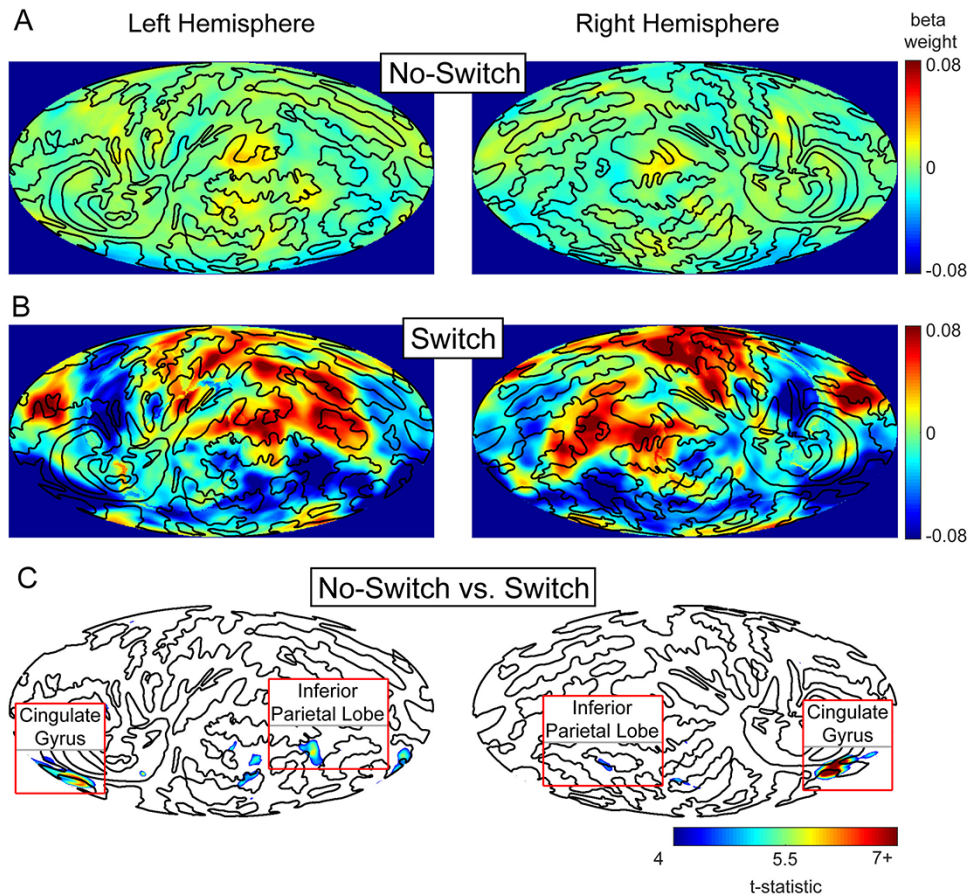


Figure 4. No-switch versus switch activation maps. (A) Surface topography showing beta weight values averaged across participants for trials where no-switch in perception (i.e. stable) was reported. (B) Surface topography showing beta weight values averaged across participants for trials where a switch in perception was reported. (C) Surface topography of significant t-statistic values contrasting no-switch minus switch trials (no-switch minus switch). Contour lines outline cortical anatomy. Panels indicate ROIs where large regions of significant differences were observed.

where an integrated percept was reported (i.e. trials where a switch to integrated occurred). As shown in Fig. 4, Fig. 5 shows less activity during stable trials (Fig. 5A) compared to switch trials (Fig. 5B) but no significant differences between stable and switch trials (Fig. 5C). In a similar but separate analysis (unable to project subcortical thalamic activity to surface map), significantly larger beta weights were observed in the auditory thalamus ROI during switch compared to stable trials (all voxels included per subject; paired t-test: $t(12) = 2.57$, $P = 0.024$, $d = 0.73$).

Cortical surface maps: stable segregated versus switch from integrated to segregated

Similar to Fig. 5, Fig. 6 shows less activity during stable trials (Fig. 6A) compared to switch trials (Fig. 6B), but for this contrast, significant differences were identified bilaterally in CG and left inferior parietal lobule (Fig. 6C). Beta weights from all voxels in the thalamic ROI were compared for stable segregated versus switch from integrated to segregated. No significant difference was observed (all voxels included per subject; paired t-test: $t(12) = 0.15$, $P = 0.88$, $d = 0.057$).

Neural dynamics following and preceding switches

To determine the pattern of neural activity following and preceding switches, beta weights corresponding to perceptual phases

that consisted of six trials or greater were extracted from sound-sensitive voxels in three different portions of the auditory cortex: HG (which contains primary auditory cortex), anterior STG (which contains the ventral auditory pathway), and posterior STG (which contains the dorsal auditory pathway). Beta weights corresponding to the first (switch trial), second, and third trials in the phase and the third, second, and first trials from the end of the phase (preceding a switch) are plotted in Fig. 7A. Figure 7B includes the fourth and fifth trials following and preceding a switch. It must be noted that expanding this internal range results in potential overlap between trials. For example, for a perceptual phase consisting of seven trials, the fourth trial after the following the switch is also the fourth trial preceding a switch (Fig. 7B, gray shading indicates the region of potential overlap). All three auditory cortical ROIs showed similar temporal dynamics following and preceding switches, with a small decrease in activation following switches from one percept to another and large rebounds in activity prior to a switch. The results were similar for relatively short perceptual phases (Fig. 7A) and relatively long perceptual phases (Fig. 7B). No differences were observed for integrated compared to segregated phases and were combined into a single dataset for statistical purposes. Beta weights were averaged across the three brain regions and subjected to single-factor repeated-measures analyses of variance with time points following or leading up to a change in percept as the factor. Data for separate ROIs were retained in the figure for illustrative purposes. For each of the four

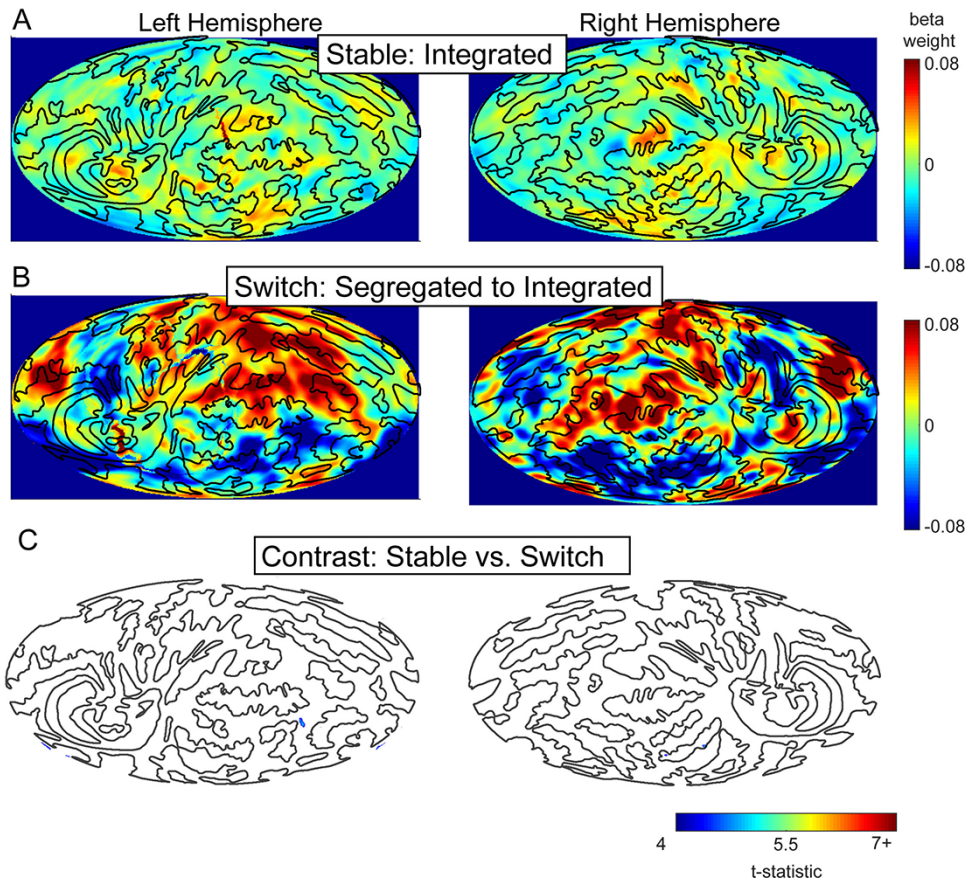


Figure 5. Stable-integrated versus switch to integrated activation maps. (A) Surface topography showing beta weight values averaged across participants for trials where stable, integrated, no-switch in perception (i.e. stable) was reported. (B) Surface topography showing beta weight values averaged across participants for trials where a switch from segregated to integrated perception was reported. (C) No significant cortical regions of t-statistic values were observed comparing stable-integrated versus switch to integrated trials. Contoured lines outline cortical anatomy.

combinations of time course length, and phase start or phase ending, a similar effect size was obtained (three trials at the start of phase: $F(2,24) = 3.752$, $P = 0.038$, $\eta_p^2 = 0.238$; three trials at the end of phase: $F(2,24) = 3.244$, $P = 0.057$, $\eta_p^2 = 0.213$; five trials at the start of phase: $F(4,48) = 3.644$, $P = 0.011$, $\eta_p^2 = 0.233$; five trials at the end of phase: $F(4,48) = 3.639$, $P = 0.011$, $\eta_p^2 = 0.233$). It may be interesting to note that an identical analysis that used every voxel in the HG, anterior STG, and posterior STG regions (as separate ROIs or combined) failed to show adaptation effects. As a result, comparable analyses in other regions revealed in the whole-brain mapping could not be assessed, as there were little-to-no sound-sensitive voxels found in CG and IPL.

Discussion

In this study, an intermittent response paradigm in combination with fMRI was used to investigate the neural correlates of spontaneous switches in perception in response to a bistable auditory stimulus. Behavioral response patterns showed standard measures of auditory bistability: approximately equal proportion of both percepts (Fig. 2A), minimal correlation between phase duration and subsequent phase duration (Fig. 2C), and a log-normal distribution of phase duration (Fig. 2D; Pressnitzer and Hupé 2006; Denham et al. 2014; Barniv and Nelken 2015; Brascamp et al. 2015). We found that the distribution of perceptual phase durations was better fit with a log-normal function than a gamma function,

which appears consistent with a theory of stochasticity in the neural processes underlying spontaneous switches in perception. The hypothesis reflected by this characteristic is that phase duration distributions best fit with a gamma function reflect a single stochastic process, while a log-normal function reflects multiple independent stochastic processes (see Cao et al. 2016; Denham et al. 2018). The results presented here shed light on the regions of the brain implicated in these processes and the neural representation of the effect of these processes on the time course of perception in the auditory cortex.

The whole-brain analyses presented here show several brain areas with enhanced activity around the time of switches in perception, but the areas showing enhancements depended on the switch direction. Only the auditory thalamus showed significant activation following a switch from segregated to integrated, while only the anterior CG and inferior parietal lobule showed significant activation following a switch from integrated to segregated. The current study showed qualitative enhancements of activity in the auditory cortex around the time of perceptual switches, but these were not statistically significant (Fig. 4C). This contrasts with a prior study that focused on activity in the thalamus and did show greater activity in the auditory cortex when a switch occurred, in addition to finding that the relative timing of these two regions depended on whether perception switched to a more or less dominant percept (Kondo and Kashino 2009). A study on auditory stream segregation using interaural timing differences instead of frequency separation as the cue for

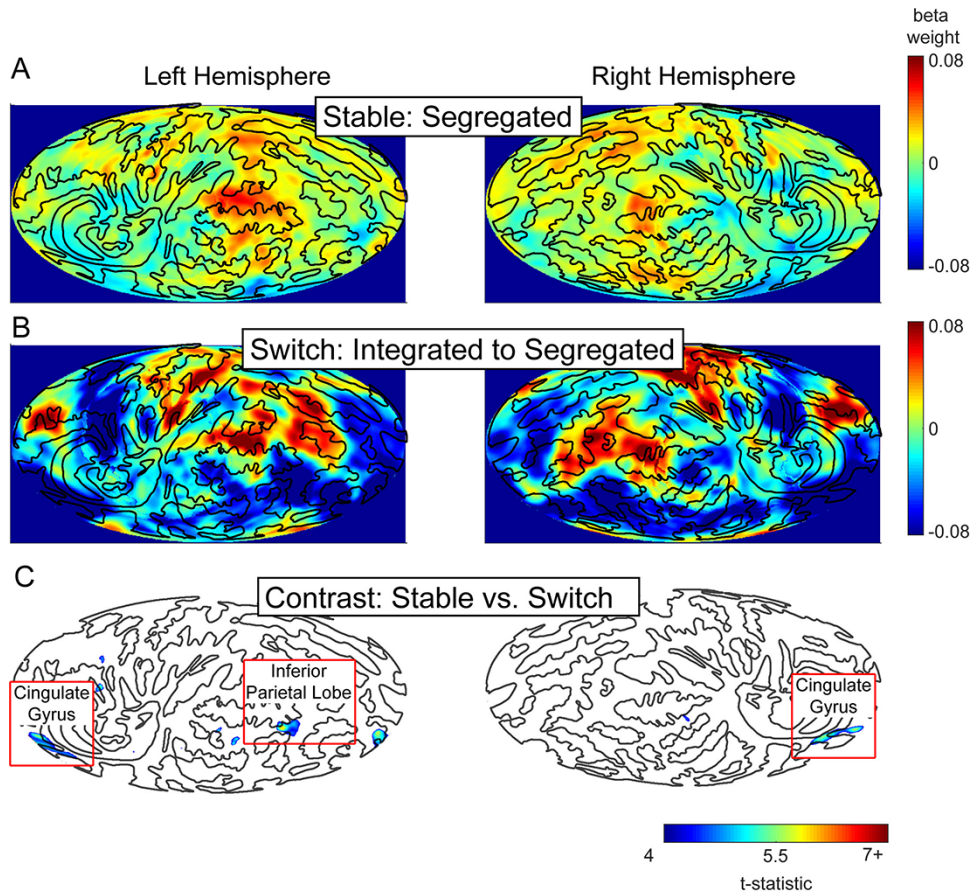


Figure 6. Stable segregated versus switch to segregated. (A) Surface topography showing beta weight values averaged across participants for trials where a segregated, stable no-switch in perception (i.e. stable) was reported. (B) Surface topography showing beta weight values averaged across participants for trials where a switch to segregated perception was reported. (C) Surface topography of significant t-statistic values contrasting stable-segregated minus switch to segregated trials (stable minus integrated). Contoured lines outline cortical anatomy. Panels indicate ROIs where large regions of significant differences were observed.

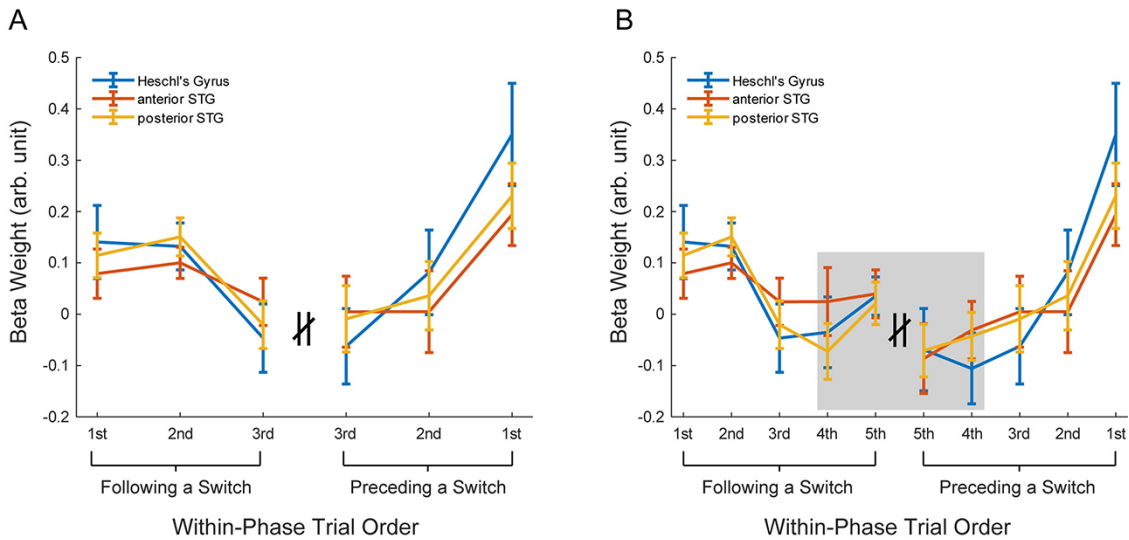


Figure 7. Evidence for neural adaptation. (A) The magnitude of neural responses for prolonged perceptual phases (minimum of six consecutive trials) for the first, second, and third trials following a switch and the third, second, and first trials preceding a switch. (B) The magnitude of prolonged perceptual phases including the fourth and fifth trials following and preceding a switch. Gray shading indicates the region of potential trial overlap for phase durations of six to nine trials

segregation found activity in inferior colliculus and auditory cortex around the time of perceptual switches (Schadwinkel and Gutschalk 2011).

The current findings are consistent with a past study showing anterior cingulate cortex and thalamic activations during verbal transformations (Kondo and Kashino 2007). That study also

found intraparietal sulcus activity during switches, which is near the inferior parietal lobule, the parietal region activated during switches in the current study. Other studies of bistable perception using auditory segregation paradigms have also indicated an important role for parietal areas during bistable perception, with one study showing more activity in the intraparietal sulcus during perception of segregated compared to perception of integrated (Cusack 2005) and another study showing greater activity in the intraparietal sulcus during passive listening when more powerful segregation cues were present (Teki et al. 2011). Both of these past studies failed to find significant effects in or near the primary auditory cortex for whole-brain comparisons between conditions, as in the current study. It should also be noted that this study with $N = 13$ may be underpowered for observing differential effects in the auditory cortex in the presence of ongoing sound, when the only differences are cognitive. With an increased number of participants, some of the weaker trends, such as the difference between no-switch and switch conditions in the auditory cortex, may have reached significance.

It is difficult to evaluate the extent to which the whole-brain analysis findings support one type of theory of conscious perception or another. Enhanced activity in brain areas outside of auditory cortex around the time of perceptual switches could support global workspace theory and higher-order theories (Dehaene et al. 2003; Dehaene and Changeux 2011; Lau and Rosenthal 2011; Brown et al. 2019). However, the fact that only the thalamus was activated for switches to integrated while the anterior cingulate cortex and inferior parietal lobule were activated during switches to segregated complicates the interpretation. Furthermore, the lack of precise timing with fMRI and the experimental paradigm means that it is unclear the exact relationship between conscious perception and the cingulate and parietal activities.

To interpret the results of whole-brain analyses, it is helpful to consider some of the analytical details of this study. The parameters of the canonical HRF used to calculate voxel-response beta weights were chosen to maximize BOLD response sensitivity to acoustic stimuli and have a peak at 4.5 s (Glover, 1999). Topographic maps of beta weights presented in Figs 4B, 5B, and 6C show regions of dark blue in ROIs corresponding to the CG and IPL, indicating large deviations from the BOLD response tuned to primary auditory cortex. These large deviations are responsible for the significant differences observed in the contrast maps in Figs 4C and 6C. For now, we can only speculate that regional differences in temporal dynamics are responsible for the reduced activity observed in the CG and IPL.

Finally, the fact that prefrontal cortex areas were not activated during perceptual switches limits the support for higher-order theories, which claim that prefrontal cortex must contain second-order representations of sensory information for conscious perception to occur (Lau and Rosenthal 2011; Brown et al. 2019). In contrast, global workspace theory says that both prefrontal and parietal areas are important for conscious perception (Dehaene et al. 2003; Dehaene and Changeux 2011). Future studies of auditory conscious perception using no-report paradigms in which decision-making and button presses are decoupled from conscious perception would likely help resolve these issues (Shafto and Pitts 2015; Cohen et al. 2020) and possibly provide support for sensory-only theories of conscious perception (Lamme and Roelfsema 2000; Hochstein and Ahissar 2002; Aru et al. 2012; DiCarlo et al. 2012). Studies of neural processing during perceptual switches induced by abrupt changes to stimulus properties (Haywood and Roberts 2010, 2013; Rankin et al. 2017; Higgins et al. 2021), rather than spontaneous switches,

could also be used to help dissociate conscious perception from task demands.

Our second set of findings showed decreased activity in sound-responsive areas following a switch from one percept to another but increased activity preceding switches in percept. This pattern of dynamic activity in the auditory cortex is consistent with our previous finding that the auditory sustained potential was larger following a switch compared to no-switch trials (Higgins et al. 2020), although that study did not track the time course of change like in the current study. Our findings of neural adaptation are also consistent with standard models that explain switches in perception as the result of neural adaptation that destabilizes dominance of the current percept, allowing an alternative percept to gradually overcome inhibition and take over as a new dominant percept (Wilson 2003; Tong et al. 2006; Noest et al. 2007; Brascamp et al. 2008; Grossberg et al. 2008; Rankin et al. 2015; Li et al. 2017; Rankin et al. 2017; Rankin and Rinzel 2019; Little et al. 2020). Finally, we found similar neural dynamics of adaptation in multiple parts of the superior temporal lobe, including HG, posterior STG, and anterior STG. This is consistent with our computational model (Little et al. 2020), as well as a prior conceptual model (Tong et al. 2006), which proposed similar neural dynamics underlying bistable perception at various levels of sensory hierarchies.

An unexpected feature of the adaptation analysis is the observation that the trial immediately preceding the trial where a switch was reported exhibited a large response, even larger on average than the switch trial. There are a few possible explanations of this finding that relate to two choices made by us. The first is our analytic decision that the responses must be from perceptual phases of at least six-trial durations (or 15.6 s), in order to establish a pattern of adaptation representing (at minimum) three trials at the beginning and end of the perceptual phase. Given long duration percepts, it is possible that the large response prior to the switch may reflect (i) indecision as to whether a switch occurred, resulting in misassignment of the switch trial or (ii) volitional control of perception exerted prior to the switch in perception, a cognitive phenomenon known to effect cortical responses (Billig et al. 2018). Separate from these possible explanations, it has been widely speculated in the framework of perceptual rivalry that when perception switches, there is release from inhibition of one percept combined with strengthening of the other percept (Shapiro et al. 2009; Little et al. 2020). It is possible the underlying neural dynamics contributed to the increased BOLD signal observed in Fig. 7 prior to a switch.

Models of bistable perception also typically feature inhibition as a central mechanism for switching and for the dominance of a single percept at any given time. However, there have been few studies that investigated this directly. An exception is from magnetic resonance spectroscopy studies, which found that higher *in vivo* levels of gamma-aminobutyric acid in modality-specific cortical areas correlate with longer percept durations while observing auditory and visual bistable stimuli (van Loon et al. 2013; Kondo et al. 2018). This is consistent with the standard models because more inhibition of the non-dominant percept would likely prolong the duration of the dominant percept. In contrast, the ability to volitionally control perceptual switching was associated with gamma-aminobutyric acid levels in the posterior parietal cortex for both auditory and visual bistable stimuli (Kondo et al. 2018). Finally, ingestion of lorazepam, an agonist of gamma-aminobutyric acid A receptors, also had the effect of prolonging percept durations, again consistent with the standard model of bistable perception (van Loon et al. 2013). While these

magnetic resonance spectroscopy studies directly assessed the role of inhibition in bistable switching, our fMRI data are the first, to our knowledge, to directly test the importance of neural adaptation in perceptual switches. Furthermore, it is possible that the increase in BOLD activity prior to a switch in perception reflects aspects of inhibition; in particular, it is possible that as neural adaptation of the current percept proceeds, a likewise reduction in the inhibition of the other percept leads to overall more activity just prior to a switch.

An important limitation of the adaptation results, as well as past modeling work, is that they ignore possible contributions of attention (but see Rankin and Rinzel 2022), which prior studies have shown to be important in enabling the first switch from integration to segregation during auditory streaming paradigms (Carlyon et al. 2001; Cusack et al. 2004; Snyder et al. 2006). We did not measure or manipulate attention, so it is impossible to know whether some of the perceptual switches and concomitant neural dynamics may have been the result of distraction from the scanner or internal thoughts. Therefore, it would be interesting to see in future studies how the observed neural dynamics from the current study behave following attention shifts or during bottom-up effects of stimulus change on perception that have also been observed in behavioral studies of auditory streaming (Anstis and Saida 1985; Rogers and Bregman 1998; Haywood and Roberts 2013; Rankin et al. 2017; Higgins et al. 2021). Finally, it would also be interesting to observe how effects of prior stimulus characteristics and prior perception, known to influence the likelihood of perception to change from one trial to the next, modulate the neural dynamics we observed in the auditory cortex (Snyder et al. 2008, 2009a,b; Haywood and Roberts 2010, 2011; Weintraub and Snyder 2015).

Conclusion

Neural correlates of switches in perception were examined using an intermittent bistable auditory stream segregation paradigm using fMRI. A whole-brain analysis found that the auditory thalamus had enhanced activity during switches to the integrated percept, but the anterior CG and inferior parietal lobule had enhanced activity during switches to the segregated percept. An analysis of sound-responsive areas only showed patterns of declining BOLD activity following switches in perception and increasing strength prior to a switch in perception, consistent with computational models of bistable switching that rely on neural adaptation and inhibitory competition.

Data availability

Our institutional review board protocol and consent forms did not allow for sharing of data with third parties, so we are not able to make participant behavioral and fMRI data available. Our code is available on the Open Science Framework: https://osf.io/ytsn5/?view_only=1b81842940c9402997499d3f16b96255.

Funding

This work was supported by the National Institutes of Health (grant number P20GM103650) and the Office of Naval Research (grant number N00014-16-1-2879).

Conflict of interest statement

None declared.

References

- Anstis S, Saida S. Adaptation to auditory streaming of frequency-modulated tones. *J Exp Psychol Hum Percept Perform* 1985;**11**:257–71.
- Aru J, Bachmann T, Singer W et al. Distilling the neural correlates of consciousness. *Neurosci Biobehav Rev* 2012;**36**:737–46.
- Barniv D, Nelken I. Auditory Streaming as an Online Classification Process with Evidence Accumulation. *PLoS One* 2015;**10**:e0144788.
- Billig AJ, Davis MH, Carlyon RP. Neural decoding of bistable sounds reveals an effect of intention on perceptual organization. *J Neurosci* 2018;**38**:2844–53.
- Brascamp JW, Klink PC, Levelt WJ. The ‘laws’ of binocular rivalry: 50 years of Levelt’s propositions. *Vision Res* 2015;**109**:20–37.
- Brascamp JW, Knapen TH, Kanai R et al. Multi-timescale perceptual history resolves visual ambiguity. *PLoS One* 2008;**3**:e1497.
- Brett M, Penny W, Kiebel S. Introduction to random field theory. In: Ashburner J, Friston K, Penny W (eds), *Human Brain Function*, 2nd edn. New York: Academic, 2003, 867–80.
- Brown R, Lau H, LeDoux JE. Understanding the higher-order approach to consciousness. *Trends Cogn Sci* 2019;**23**:754–68.
- Canales-Johnson A, Billig AJ, Olivares F et al. Dissociable neural information dynamics of perceptual integration and differentiation during bistable perception. *Cereb Cortex* 2020;**30**:4563–80.
- Cao R, Pastukhov A, Mattia M et al. Collective activity of many bistable assemblies reproduces characteristic dynamics of multistable perception. *J Neurosci* 2016;**36**:6957–72.
- Carlyon RP, Cusack R, Foxtan JM et al. Effects of attention and unilateral neglect on auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2001;**27**:115–27.
- Cohen MA, Ortego K, Kyroudis A et al. Distinguishing the neural correlates of perceptual awareness and postperceptual processing. *J Neurosci* 2020;**40**:4925–35.
- Crick F, Koch C. Are we aware of neural activity in primary visual cortex? *Nature* 1995;**375**:121–3.
- Curtu R, Wang X, Brunton BW et al. Neural signatures of auditory perceptual bistability revealed by large-scale human intracranial recordings. *J Neurosci* 2019;**39**:6482–97.
- Cusack R. The intraparietal sulcus and perceptual organization. *J Cogn Neurosci* 2005;**17**:641–51.
- Cusack R, Deeks J, Aikman G et al. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J Exp Psychol Hum Percept Perform* 2004;**30**:643–56.
- Davidson GD, Pitts MA. Auditory event-related potentials associated with perceptual reversals of bistable pitch motion. *Front Hum Neurosci* 2014;**8**:572.
- Dehaene S, Changeux JP. Experimental and theoretical approaches to conscious processing. *Neuron* 2011;**70**:200–27.
- Dehaene S, Sergent C, Changeux JP. A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc Natl Acad Sci USA* 2003;**100**:8520–5.
- Denham S, Bohm TM, Bendixen A et al. Stable individual characteristics in the perception of multiple embedded patterns in multistable auditory stimuli. *Front Neurosci* 2014;**8**:25.
- Denham SL, Farkas D, van Ee R et al. Similar but separate systems underlie perceptual bistability in vision and audition. *Sci Rep* 2018;**8**:7106.
- Denham SL, Winkler I. The role of predictive models in the formation of auditory streams. *J Physiol Paris* 2006;**100**:154–70.
- Desikan RS, Segonne F, Fischl B et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 2006;**31**:968–80.
- DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? *Neuron* 2012;**73**:415–34.

- Dykstra AR, Cariani PA, Gutschalk A. A roadmap for the study of conscious audition and its neural basis. *Phil Trans R Soc B* 2017;**372**:20160103.
- Dykstra AR, Halgren E, Gutschalk A et al. Neural correlates of auditory perceptual awareness and release from informational masking recorded directly from human cortex: a case study. *Front Neurosci* 2016;**10**:472.
- Fischl B, Salat DH, Busa E et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002;**33**:341–55.
- Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 1999;**9**:416–29.
- Grossberg S, Yazdanbakhsh A, Cao Y et al. How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Res* 2008;**48**:2232–50.
- Gutschalk A, Micheyl C, Melcher JR et al. Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci* 2005;**25**:5382–8.
- Gutschalk A, Micheyl C, Oxenham AJ. Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol* 2008;**6**:e138.
- Gutschalk A, Oxenham AJ, Micheyl C et al. Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. *J Neurosci* 2007;**27**:13074–81.
- Haywood NR, Roberts B. Build-up of the tendency to segregate auditory streams: resetting effects evoked by a single deviant tone. *J Acoust Soc Am* 2010;**128**:3019–31.
- Haywood NR, Roberts B. Effects of inducer continuity on auditory stream segregation: comparison of physical and perceived continuity in different contexts. *J Acoust Soc Am* 2011;**130**:2917–27.
- Haywood NR, Roberts B. Build-up of auditory stream segregation induced by tone sequences of constant or alternating frequency and the resetting effects of single deviants. *J Exp Psychol Hum Percept Perform* 2013;**39**:1652–66.
- Higgins NC, Little DF, Yerkes BD et al. Neural correlates of perceptual switching while listening to bistable auditory streaming stimuli. *Neuroimage* 2020;**204**:116220.
- Higgins NC, Monjaras AG, Yerkes BD et al. Resetting of auditory and visual segregation occurs after transient stimuli of the same modality. *Front Psychol* 2021;**12**:720131.
- Hill KT, Bishop CW, Yadav D et al. Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. *BMC Neurosci* 2011;**12**:85.
- Hochstein S, Ahissar M. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 2002;**36**:791–804.
- Hupé JM, Joffo LM, Pressnitzer D. Bistability for audiovisual stimuli: perceptual decision is modality specific. *J Vis* 2008;**8**:1–15.
- Kleinschmidt A, Buchel C, Zeki S et al. Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc Biol Sci* 1998;**265**:2427–33.
- Koch C, Tsuchiya N. Attention and consciousness: two distinct brain processes. *Trends Cogn Sci* 2007;**11**:16–22.
- Kondo HM, Kashino M. Neural mechanisms of auditory awareness underlying verbal transformations. *Neuroimage* 2007;**36**:123–30.
- Kondo HM, Kashino M. Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. *J Neurosci* 2009;**29**:12695–701.
- Kondo HM, Pressnitzer D, Shimada Y et al. Inhibition-excitation balance in the parietal cortex modulates volitional control for auditory and visual multistability. *Sci Rep* 2018;**8**:14548.
- Königs L, Gutschalk A. Functional lateralization in auditory cortex under informational masking and in silence. *Eur J Neurosci* 2012;**36**:3283–90.
- Kornmeier J, Bach M. The Necker cube—an ambiguous figure disambiguated in early visual processing. *Vision Res* 2005;**45**:955–60.
- Lamme VA, Roelfsema PR. The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 2000;**23**:571–9.
- Lau H, Rosenthal D. Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* 2011;**15**:365–73.
- Lee SH, Blake R, Heeger DJ. Traveling waves of activity in primary visual cortex during binocular rivalry. *Nat Neurosci* 2005;**8**:22–3.
- Lehky SR. Binocular rivalry is not chaotic. *Proc Biol Sci* 1995;**259**:71–6.
- Leopold DA, Logothetis NK. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 1996;**379**:549–53.
- Li HH, Rankin J, Rinzel J et al. Attention model of binocular rivalry. *Proc Natl Acad Sci USA* 2017;**114**:E6192–201.
- Little DF, Snyder JS, Elhilali M. Ensemble modeling of auditory streaming reveals potential sources of bistability across the perceptual hierarchy. *PLoS Comput Biol* 2020;**16**:e1007746.
- Logothetis NK, Schall JD. Neuronal correlates of subjective visual perception. *Science* 1989;**245**:761–3.
- McLaughlin SA, Higgins NC, Stecker GC. Tuning to binaural cues in human auditory cortex. *J Assoc Res Otolaryngol* 2016;**17**:37–53.
- Melloni L, Mudrik L, Pitts M et al. Making the hard problem of consciousness easier. *Science* 2021;**372**:911–2.
- Meng M, Remus DA, Tong F. Filling-in of visual phantoms in the human brain. *Nat Neurosci* 2005;**8**:1248–54.
- Noest AJ, van Ee R, Nijs MM et al. Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J Vis* 2007;**7**:10.
- Pitts MA, Gavin WJ, Nerger JL. Early top-down influences on bistable perception revealed by event-related potentials. *Brain Cogn* 2008;**67**:11–24.
- Pitts MA, Martinez A, Hillyard SA. When and where is binocular rivalry resolved in the visual cortex? *J Vis* 2010;**10**:25.
- Pitts MA, Martinez A, Stalmaster C et al. Neural generators of ERPs linked with Necker cube reversals. *Psychophysiology* 2009;**46**:694–702.
- Pitts MA, Nerger JL, Davis TJ. Electrophysiological correlates of perceptual reversals for three different types of multistable images. *J Vis* 2007;**7**:6.
- Polonsky A, Blake R, Braun J et al. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat Neurosci* 2000;**3**:1153–9.
- Pressnitzer D, Hupé JM. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr Biol* 2006;**16**:1351–7.
- Rankin J, Osborn Popp PJ, Rinzel J. Stimulus pauses and perturbations differentially delay or promote the segregation of auditory objects: psychoacoustics and modeling. *Front Neurosci* 2017;**11**:198.
- Rankin J, Rinzel J. Computational models of auditory perception from feature extraction to stream segregation and behavior. *Curr Opin Neurobiol* 2019;**58**:46–53.
- Rankin J, Rinzel J. Attentional control via synaptic gain mechanisms in auditory streaming. *Brain Res* 2022;**1778**:147720.
- Rankin J, Sussman E, Rinzel J. Neuromechanistic model of auditory bistability. *PLoS Comput Biol* 2015;**11**:e1004555.
- Rogers WL, Bregman AS. Cumulation of the tendency to segregate auditory streams: resetting by changes in location and loudness. *Percept Psychophys* 1998;**60**:1216–27.
- Sanders RD, Winston JS, Barnes GR et al. Magnetoencephalographic correlates of perceptual state during auditory bistability. *Sci Rep* 2018;**8**:976.

- Schadwinkel S, Gutschalk A. Transient bold activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. *J Neurophysiol* 2011;**105**:1977–83.
- Shafto JP, Pitts MA. Neural signatures of conscious face perception in an inattentive blindness paradigm. *J Neurosci* 2015;**35**:10940–8.
- Sheinberg DL, Logothetis NK. The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci USA* 1997;**94**:3408–13.
- Shapiro A, Moreno-Bote R, Rubin N et al. Balance between noise and adaptation in competition models of perceptual bistability. *J Comput Neurosci* 2009;**27**:37–54.
- Snyder JS, Alain C, Picton TW. Effects of attention on neuroelectric correlates of auditory stream segregation. *J Cogn Neurosci* 2006;**18**:1–13.
- Snyder JS, Carter OL, Hannon EE et al. Adaptation reveals multiple levels of representation in auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2009a;**35**:1232–44.
- Snyder JS, Carter OL, Lee S-K et al. Effects of context on auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2008;**34**:1007–16.
- Snyder JS, Gregg MK, Weintraub DM et al. Attention, awareness, and the perception of auditory scenes. *Front Psychol* 2012;**3**:15.
- Snyder JS, Holder WT, Weintraub DM et al. Effects of prior stimulus and prior perception on neural correlates of auditory stream segregation. *Psychophysiology* 2009b;**46**:1208–15.
- Snyder JS, Schwiedrzik CM, Vitela AD et al. How previous experience shapes perception in different sensory modalities. *Front Hum Neurosci* 2015a;**9**:594.
- Snyder JS, Yerkes BD, Pitts MA. Testing domain-general theories of perceptual awareness with auditory brain responses. *Trends Cogn Sci* 2015b;**19**:295–7.
- Sterzer P, Rees G. A neural basis for percept stabilization in binocular rivalry. *J Cogn Neurosci* 2008;**20**:389–99.
- Sterzer P, Russ MO, Preibisch C et al. Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *Neuroimage* 2002;**15**:908–16.
- Teki S, Chait M, Kumar S et al. Brain bases for auditory stimulus-driven figure-ground segregation. *J Neurosci* 2011;**31**:164–71.
- Tong F, Engel SA. Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* 2001;**411**:195–9.
- Tong F, Meng M, Blake R. Neural bases of binocular rivalry. *Trends Cogn Sci* 2006;**10**:502–11.
- Tong F, Nakayama K, Vaughan JT et al. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 1998;**21**:753–9.
- Tononi G, Boly M, Massimini M et al. Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci* 2016;**17**:450–61.
- Tononi G, Sporns O, Edelman GM. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci USA* 1994;**91**:5033–7.
- Tononi G, Srinivasan R, Russell DP et al. Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proc Natl Acad Sci USA* 1998;**95**:3198–203.
- van Loon AM, Knapen T, Scholte HS et al. GABA shapes the dynamics of bistable perception. *Curr Biol* 2013;**23**:823–7.
- Weintraub DM, Ramage EM, Sutton G et al. Auditory stream segregation impairments in schizophrenia. *Psychophysiology* 2012;**49**:1372–83.
- Weintraub DM, Snyder JS. Evidence for high-level feature encoding and persistent memory during auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2015;**41**:1563–75.
- Wiegand K, Gutschalk A. Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. *Neuroimage* 2012;**61**:62–9.
- Wilson EC, Melcher JR, Micheyl C et al. Cortical fMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. *J Neurophysiol* 2007;**97**:2230–8.
- Wilson HR. Computational evidence for a rivalry hierarchy in vision. *Proc Natl Acad Sci USA* 2003;**100**:14499–503.
- Woods DL, Stecker GC, Rinne T et al. Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One* 2009;**4**:e5183.
- Woolrich MW, Ripley BD, Brady M et al. Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* 2001;**14**:1370–86.
- Wunderlich K, Schneider KA, Kastner S. Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nat Neurosci* 2005;**8**:1595–602.