



Neural correlates of perceptual switching while listening to bistable auditory streaming stimuli

N.C. Higgins^{a,*}, D.F. Little^b, B.D. Yerkes^a, K.M. Nave^a, A. Kuruville-Mathew^a, M. Elhilali^b, J.S. Snyder^a

^a Department of Psychology, University of Nevada, Las Vegas, 4505 South Maryland Parkway, Las Vegas, NV, 89154, USA

^b Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 North Charles Street, Barton Hall, Baltimore, MD, 21218, USA

ABSTRACT

Understanding the neural underpinning of conscious perception remains one of the primary challenges of cognitive neuroscience. Theories based mostly on studies of the visual system differ according to whether the neural activity giving rise to conscious perception occurs in modality-specific sensory cortex or in associative areas, such as the frontal and parietal cortices. Here, we search for modality-specific conscious processing in the auditory cortex using a bistable stream segregation paradigm that presents a constant stimulus without the confounding influence of physical changes to sound properties. ABA triplets (i.e., alternating low, A, and high, B, tones, and _ gap) with a 700 ms silent response period after every third triplet were presented repeatedly, and human participants reported nearly equivalent proportions of 1- and 2-stream percepts. The pattern of behavioral responses was consistent with previous studies of visual and auditory bistable perception. The intermittent response paradigm has the benefit of evoking spontaneous perceptual switches that can be attributed to a well-defined stimulus event, enabling precise identification of the timing of perception-related neural events with event-related potentials (ERPs). Significantly more negative ERPs were observed for 2-streams compared to 1-stream, and for switches compared to non-switches during the sustained potential (500–1000 ms post-stimulus onset). Further analyses revealed that the negativity associated with switching was independent of switch direction, suggesting that spontaneous changes in perception have a unique neural signature separate from the observation that 2-stream percepts evoke more negative ERPs than 1-stream. Source analysis of the sustained potential showed activity associated with these differences originating in anterior superior temporal gyrus, indicating involvement of the ventral auditory pathway that is important for processing auditory objects.

1. Introduction

The moment-to-moment conscious states we all experience represent an enormous variety of experiences, due to our capacity to process many different types of stimuli while also incorporating internal and external contextual factors into our perceptual representations. According to the global workspace theory (Baars, 1988; Changeux and Dehaene, 2008; Dehaene and Changeux, 2011), individual sensory pathways process stimulus features unconsciously, until they arrive in frontal and parietal cortical areas that enable the widespread sharing of information about different features within and across modalities. In contrast, the information integration theory is more agnostic about where exactly in the cortex consciousness is generated, simply specifying that it can occur in any area that can generate different states in which the contents of awareness are integrated (Tononi et al., 2016). Still other theories hypothesize that basic forms of sensory awareness are generated in sensory cortex pathways, such as the ventral visual stream in the inferior occipital and temporal lobe (DiCarlo et al., 2012; Hochstein and Ahissar, 2002; Milner and Goodale, 2008; Pitts et al., 2012; Tong et al., 2006), without need for processing in

associative areas such as the frontal and parietal cortex. Thus, there is still considerable debate about the neural basis of consciousness, including where in the brain it is generated. Moreover, almost all of these theories have been generated on the basis of visual studies, making it vitally important to also study auditory conscious processing to test the generality of these theories (Dykstra et al., 2017; Snyder et al., 2015).

Bistable stimuli provide an ideal means for experimentally manipulating consciousness because they induce mutually exclusive percepts that switch back and forth despite unchanging physical stimulus parameters. At the neural level, the standard model for bistable perception proposes that at any given time frame the current percept is destabilized over time due to adaptation, eventually reaching a threshold whereby the second percept becomes active and suppresses the first (Brascamp et al., 2018; Rankin et al., 2015, 2017). Thus, competitive inhibition of both percepts results in a subjective experience of multiple percepts switching back and forth over time. In studies of binocular rivalry for example, two dissimilar images are presented simultaneously to each eye resulting in perception of one image or the other, spontaneously switching over time. Recordings of action potentials from individual neurons implicate the ventral visual

* Corresponding author.

E-mail addresses: nathan.higgins@unlv.edu (N.C. Higgins), david.frank.little@gmail.com (D.F. Little), yerkesb@unlv.nevada.edu (B.D. Yerkes), nave@unlv.nevada.edu (K.M. Nave), amat527@aucklanduni.ac.nz (A. Kuruville-Mathew), mounya@jhu.edu (M. Elhilali), joel.snyder@unlv.edu (J.S. Snyder).

<https://doi.org/10.1016/j.neuroimage.2019.116220>

Received 12 June 2019; Received in revised form 19 August 2019; Accepted 19 September 2019

Available online 20 September 2019

1053-8119/© 2019 Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

pathway as the most likely locus for representation of the active percept (Leopold and Logothetis, 1996), although functional imaging studies in humans implicate earlier visual areas (Tong and Engel, 2001) and frontal and parietal networks (Lumer et al., 1998; Tong et al., 1998).

In both visual and auditory systems, there is ample evidence for diverging dorsal (“where”) and ventral (“what”) processing pathways (Arnott et al., 2004; Goodale and Milner, 1992; Lomber and Malhotra, 2008; Rauschecker and Tian, 2000). The ventral pathway, therefore, is a logical candidate for object identification, and in the case of complex scenes, segregation of separate objects and resolution of perceptual ambiguities. In the auditory literature, however, with the exception of an intracranial study by Curtu et al. (2019), there is little evidence for involvement of the ventral pathway in bistable perception comparable to observations in the visual domain. Human imaging studies investigating bistable auditory stimuli have implicated primary and secondary auditory cortex in and around Heschl’s gyrus, as well as parietal cortical regions (Billig et al., 2018; Curtu et al., 2019; Cusack, 2005; Gutschalk et al., 2008; Kondo et al., 2018). Moreover, the majority of these studies have focused on differences between the contents of perception or mechanisms of a switch in perception but not both (Kondo and Kashino, 2009; Sanders et al., 2018; Snyder et al., 2006). Finally, in both visual and auditory studies, the most direct evidence connecting neural adaptation and inhibition to perception comes from a study using magnetic resonance spectroscopy. Kondo et al. (2018) demonstrated a link between GABA/glutamate ratios in primary sensory cortices and percept duration during spontaneous fluctuations in perception, whereas prefrontal and parietal regions were linked to volitional control of perception. More specifically, the higher the GABA-to-glutamate ratio in frontal and parietal areas, the longer a percept was maintained, providing valuable insight into the neural dynamics between the active and alternate percept. However, questions remain about how the contents of perception are modulated, and how a perceptual switch is initiated relative to the global network responsible for conscious perception.

To answer these questions, we devised an experiment that uses an established bistable auditory stream segregation paradigm (Bregman, 1990; Van Noorden, 1975), but with intermittently presented stimuli (Kornmeier and Bach, 2004; Pitts et al., 2008). This paradigm presents triplets of ABA_n tones where A corresponds to a low tone, B to a high tone, and the blank, _n, to the absence of a tone. When presented repetitively, these triplets can be perceived as either a single “galloping” auditory stream, or two separate “metronome” streams. Typically, participants hold down one button (1-stream) or a second button (2-streams) to continuously indicate their perception. In this experiment, however, every third triplet is followed by a brief pause during which the participant presses the button to indicate their perception about the prior three triplets. The benefits of this approach are two-fold. First, it tightens the temporal link between components of the EEG and what a participant determines to be a 1- or 2-stream percept, potentially allowing for the separation of components related to the contents of perception and those related to the switch in perception. Secondly, it also changes the morphology of the event-related potentials (ERPs). In particular, the introduction of 700 ms of silence provides a well-defined baseline period, enabling clearer identification of the negative sustained potential (500–1000 ms), an auditory ERP that arises from the ventral auditory pathway that is linked to auditory object perception (Scherg et al., 1989). The sustained potential is therefore a component of the ERP expected to reveal effects of adaptation of the dominant percept according to standard theories of bistable perception (Brascamp et al., 2018; Rankin et al., 2015; Tong et al., 2006).

2. Materials and methods

2.1. Participants

Thirty normal-hearing adults (11 male) with average age of 22.3 years (18–36 years) were recruited from the community in and around

the University of Nevada, Las Vegas. All techniques and procedures were approved by the University of Nevada, Las Vegas Institutional Review Board. Experimental data, protocols, and analytical routines will be made available at https://osf.io/b4qrh/?view_only=81a1f5038e304822978d6d147ae70b3d, and upon direct request to the corresponding author. Prior to the experiment all participants provided informed consent followed by a standard hearing screening to ensure that audiometric thresholds did not exceed 25 dB hearing level at 0.25, 0.5, 1, 2, 4, and 8 kHz. An additional 23 individuals who participated in the experiment were excluded due to a scarcity of trials in which a switch in perception was reported. Fourteen of these (out of $N_{\text{total}} = 53$) reported fewer than 20 total switches in perception throughout the experiment, overwhelmingly reporting 1-stream perception for the entire experiment. The remaining nine participants had noisy EEG data to a degree that less than 20 switch trials remained following the automatic epoch rejection that is described below. Therefore, the elevated number of excluded participants might be attributed to increased difficulty perceiving the 2-stream percept in this paradigm. In this case, our data might not provide insight for a subset of listeners who require different stimulus parameters to perceive 2-streams.

2.2. Intermittent response paradigm

A variation of the classic ABA_n auditory stream segregation paradigm was used. Participants were presented with repeating triplets of A and B tones, a stimulus that elicits alternating percepts of a single “galloping” auditory stream, or two separate “metronome” streams. Each 700 ms triplet consisted of A (400 Hz) and B (565.5 Hz) tones (6 semi-tone separation) presented in an ABA_n pattern with 175 ms separation between tone onsets, and a silent interval substituted for the 2nd B tone (Fig. 1). Tones were 73 ms in duration. Each trial (2.8 s total) was defined by three triplets presented in sequence followed by a 700 ms silent period designated for responding. Prior to the experiment, participants were familiarized with the task and practiced conveying their perceptual response with a button press (button 1 for 1-stream, or button 2 for 2-streams; Cedrus response pad) during the 700 ms period following the 3 ABA_n triplets. The entire experiment was divided into 8 blocks of 75 trials presented in each block. Short breaks were provided to participants in between blocks.

2.3. Stimulus presentation

Auditory stimuli were presented to listeners via insert earphones (E-A-RTONE 3A Insert Earphones) at 65 dB SPL while sitting in a sound attenuation chamber. Prior to the experiment, participants were instructed to keep their eyes focused on a white fixation cross on a gray background presented in the center of a computer screen and to report their perception with a button press. Participants were instructed to allow their perception to fluctuate without trying to hear the pattern one way or the other. All experimental stimuli were presented and responses recorded using routines written in the Julia programming language.

2.4. EEG data collection and analysis

During the task, EEG data were recorded using the BIOSEMI Active-Two system (512 Hz A/D rate) from 72 electrodes, including 64 electrodes in an EEG cap and 8 additional face electrodes. EEG data were processed using EEGLAB (Delorme and Makeig, 2004) and custom Matlab routines. Individual participant data were referenced to the average of the two mastoid channels, bandpass filtered (0.01–30 Hz), and subjected to *infomax* independent component analysis (ICA) decomposition using the *-extended* and *-runica* options (Jung et al., 2000). The results were used to manually select and remove components related to ocular artifacts. Continuous data were then epoched for each 2.8 s trial and automatic epoch rejection (*pop_autorej*) was used to remove epochs that exceeded a threshold of 120 μV . Participants had to meet an inclusion

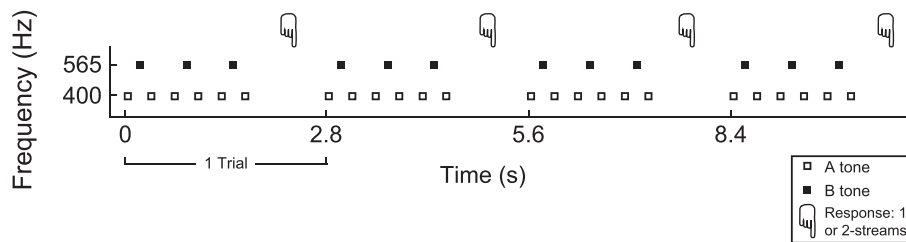


Fig. 1. Stimulus presentation and intermittent response paradigm. Schematic depicting 4 trials of the intermittent bistable auditory stimulus. Each trial consisted of three triplets composed of pure tones presented in a low-high-low (ABA) sequence. Participants indicated perception of 1- or 2-streams with a single button press during a 700 ms silent period at the end of each trial. Seventy-five consecutive trials made up an experimental block.

criterion of at least 20 epochs retained that corresponded to a perceptual switch. Specifically, a participant must have indicated a switch in perception via button press in at least 20 trials, and at least 20 of those epoched trials must have survived automatic epoch rejection. Each epoch was then defined by perceptual state: 1-stream, 2-streams, switch, no-switch, switch from 1- to 2-streams, and switch from 2- to 1-stream. Trials designated as 1- or 2-streams did not include switch trials.

2.5. Statistical analysis

Statistical comparison of waveforms was conducted using a non-parametric cluster-based analysis (Maris and Oostenveld, 2007). The first step in this process was to generate a test-statistic for comparison of perceptual states by clustering spatially adjacent channels (radius = 4 cm) for each perceptual state and conducting a paired *t*-test across subjects for each time point in the waveform (0–2100 ms; duration of sound presentation). Contiguous time points that exceeded the specified threshold ($\alpha = 0.001$ for all comparisons except $\alpha = 0.05$ for comparison of 1- versus 2-streams) were identified and *t*-values within each cluster were summed together to create cluster-level statistics based on spatial and temporal adjacency. The largest of these statistics (summed *t*-values) for each tested pair served as the test-statistic for comparison with a null-distribution (next step).

The second step generates a permutation, or null distribution, by resampling each waveform-comparison via random partitioning; a process that scrambles the labels and resamples the data into two equal-sized new, or permuted datasets. This process was repeated 1000 times for each channel and participant, effectively resulting in 1000 resampled waveforms nominally corresponding to each perceptual state. The outcome of this process generates 1000, 30-subject permuted datasets used to calculate a Monte Carlo estimate of the *p*-value. The contiguous (spatially and temporally adjacent) cluster analysis described above was then performed on each permuted-population dataset, resulting in a permuted distribution of 1000 summed *t*-values representing the largest contiguous region of significance.

In the final step, a Monte Carlo estimated *p*-value was calculated based on the number of instances the permuted distribution from step 2 exceeded the test-statistic from step 1. If the probability was less than 0.05 (50 out of 1000), the difference was considered significant.

2.6. Source analysis

Separate grand average waveforms (averaged across participants) for each of five perceptual states (all combined, 1-stream, 2-streams, switch 1-stream to 2-streams, switch 2-streams to 1-stream) were imported into BESA (Brain Electrical Source Analysis, Gräefelfing, Germany) software for dipole source analysis. The grand average that included a combination of all perceptual states was used to find a general solution that accounted for the scalp data over a time range (90–2100 ms) that encompassed the sustained potential. First, two symmetric dipoles were allowed to fit to the source to maximize the variance explained, then a second set of symmetric dipoles were allowed to fit themselves to further

maximize the variance explained. At this stage, over 96% of the variance was accounted for and we determined that any additional dipoles would likely be fitting noise. The first set of dipoles (anterior superior temporal gyrus [STG]) alone accounted for just under 95% of the variance. The second set (in medial parietal lobe), though only minorly contributing to the solution, were retained and reported due to precedence in the literature for involvement of parietal areas during auditory stream segregation (Curtu et al., 2019; Cusack, 2005; Teki et al., 2011). The two solutions (inclusion of all four dipoles, and just the initial two) were then applied to each of the other perceptual states, source waveforms were extracted for each and used to qualitatively reconstruct perceptual-state comparisons observed in the scalp data.

3. Results

3.1. Behavioral response patterns

Response patterns reflecting perception of 1- or 2-streams were collected and analyzed from 30 participants. In an effort to establish that the intermittent presentation strategy employed here resulted in a similar pattern of bistable perception as the conventional continuous presentation paradigm, a number of perceptual characteristics were examined. First, in accord with previous studies, participants typically reported an initial bias to perceive 1-stream (Bregman, 1978; Pressnitzer and Hupé, 2006) followed by convergence towards an equivalent chance of reporting 1- or 2-streams (Sanders et al., 2018), approximately 10 trials into the block (Fig. 2A, black line). A similar measurement of switch probability measured over time revealed a consistent rate of switching around 0.2 (a switch was observed on about 20% of trials) over the time course of the blocks (total switches per participant: $106.5_{\text{mean}} \pm 76_{\text{std}}$). In combination with the roughly equal probability of 1- versus 2-stream perception (1-stream probability: $0.55_{\text{mean}} \pm 0.1_{\text{std}}$), these results support the hypothesis that despite the intermittent nature of the paradigm, participants experienced stable perception over time, with occasional switches. If this were not the case, and the intermittent design failed to allow consistent perceptual buildup, the switch rate would likely be much higher, reflecting more frequent switches between the 1- and 2-streams percept due to interruptions in the ABA sequences (Cusack et al., 2004; Haywood and Roberts, 2013, 2010).

Secondly, characteristics of the duration of a given percept – conventionally referred to as a “phase” – were also in agreement with prior research. Here, each phase was defined by the number of consecutive trials that the same percept was reported (each trial is 2.8 s). The initial phase of each block was significantly longer in duration than following percepts (Fig. 2B; repeated measures ANOVA: $F_{6,29} = 4.01$, $p < 0.001$, $\eta^2_p = 0.12$; post-hoc *t*-test phase 1 vs. phase 2: $t_{29} = 3.0$, $p < 0.01$, $d = 0.54$), an observation believed to correspond to the buildup of segregation (Denham et al., 2013; Pressnitzer and Hupé, 2006). Due to the nature of the paradigm, phase measurements are necessarily a discrete variable with a minimum phase of 1 trial (or 2.8 s). Nevertheless, the distribution of phase durations approaches the shape of a logarithmic function (Fig. 2C), consistent with measures of bistable perception in

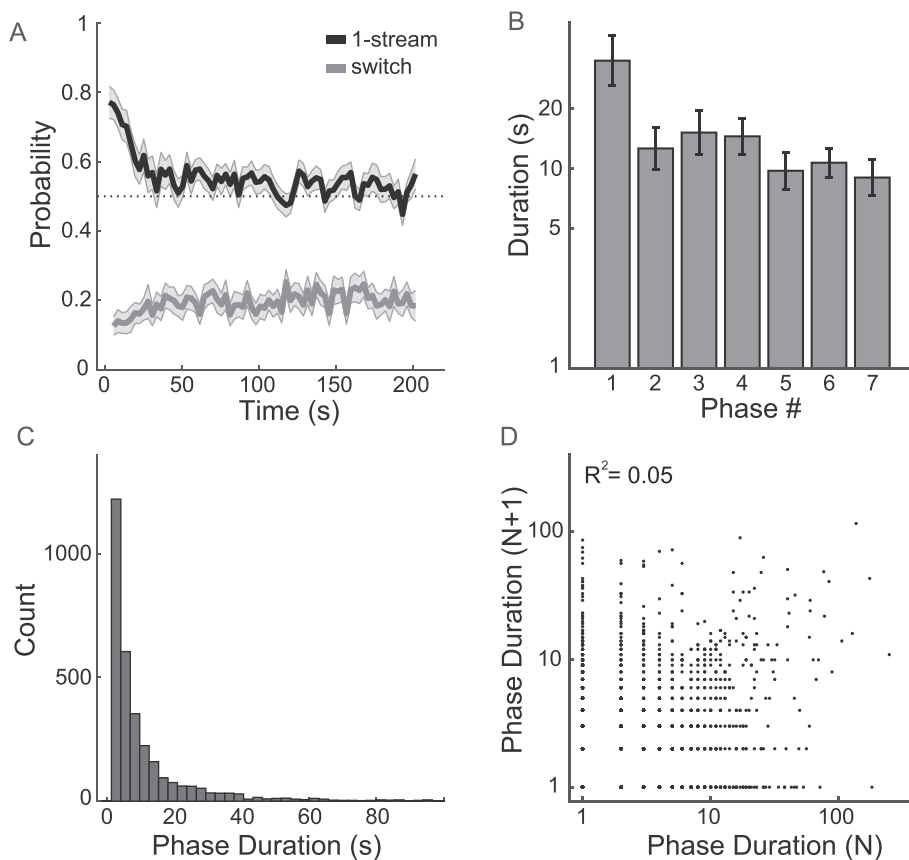


Fig. 2. Behavioral characteristics of bistable perception. A) Probability of 1-stream perception (black) and switch in perception (gray). Data represents 75 consecutive trials averaged across 8 blocks for each subject. B) Phase duration represents the number of consecutive 2.8 s trials the same percept was reported. The initial phase (length of time the same percept was reported) of each block exhibited longer duration than the subsequent seven phases. C) The distribution of phase durations pooled across the participant-dataset approximates a logarithmic function, despite the discrete nature of the variable. D) Phase duration of a given percept (N) is minimally correlated with the phase duration of the next percept (N+1). Error bars in A and B indicate SEM across subjects (N = 30).

both the auditory and visual domains (Farkas et al., 2018; Pressnitzer and Hupé, 2006). No significant difference was observed between the duration distributions for 1-stream versus 2-stream (Wilcoxon rank-sum test, $Z = 0.98$, $p = 0.33$). Lastly, the duration of a given phase (N) was minimally correlated with the duration of the following phase (N+1; Fig. 2D; $R^2 = 0.05$).

3.2. ERPs: 1-stream vs. 2-streams

ERPs were grouped into categories corresponding to 1-stream or 2-streams and switch or no-switch trials. Switch trials were subsequently separated into those in which perception switched from 1- to 2-streams or switched from 2- to 1-stream. Presentation of long duration auditory stimuli evokes a sustained negative potential that appears at frontal electrodes (Picton et al., 1978a, 1978b), and is localized to an area anterior to the portion of auditory cortex that generates the N1 (Scherg et al., 1989). This sustained potential was observed in ERPs during perception of both 1- or 2-streams, but had larger amplitude in response to 2-streams compared to 1-stream (Fig. 3A, top row). Significant differences were observed for one spatial cluster, and overall enhanced negativity for 2- versus 1-stream at channels around the top of the head (Fig. 3A, topography). This pattern of results is consistent with a number of EEG and MEG studies (Billig et al., 2018; Gutschalk et al., 2005; Snyder et al., 2006, 2009) using comparable ABA_ paradigms, with the prevailing explanation that a 2-stream percept represented by separate neural populations evokes greater activity at scalp electrodes.

3.3. ERPs: switch vs. no-switch

The first trial of each phase by definition was a trial in which a switch in perception must have occurred as it represents the stimulus period between the last indication of the previous percept and the first indication of the new percept. ERPs corresponding to switch trials were

compared to no-switch trials using the cluster-based permutation described above. The results revealed significantly greater negative responses in the sustained potential across 23 spatial clusters, for switch compared to no-switch trials (Fig. 3B, top). Differences were mainly located at right-frontal electrodes, as reflected in the difference topography (Fig. 3B, bottom). There are two potential reasons for this observed difference. The switch versus no-switch comparison does not distinguish between switching from 1- to 2-streams or switching from 2- to 1-stream, and could therefore reflect the fact that perceiving 2 streams results in larger activity, described above, regardless of the fact there was a switch. Alternatively, the observed differences could be due to switching independent of the percept. To address these possibilities two additional analyses were conducted. First, a comparison of switch type was made between switches from 1- to 2-streams and a switch from 2- to 1-stream (Fig. 4A). This revealed a temporal-spatial deviation, in which significant differences for four spatial clusters located on the top of the head (Fig. 4A, gray contour) were significantly different during an early part of the waveform (45–175 ms), whereas a large number of leftward channels displayed significant differences during a later part of the waveform (12 spatial clusters), temporally similar to the comparisons shown in Fig. 3. These observations suggest at least part of the switch versus no-switch difference is attributable to enhanced negativity associated with the perception of 2-streams versus 1-stream. Secondly, two additional comparisons were made in an effort to identify an effect of switch versus no-switch, while controlling for the already established effect of percept. Trials with a switch from 2-streams to 1-stream were compared to stable (no-switch) 1-stream percepts (Fig. 4B; eight significant clusters), and those with a switch from 1- to 2-streams were compared to stable (no-switch) 2-stream percepts (Fig. 4C; 16 significant clusters, including one at an early time range 104–162 ms, gray contour). In both cases, channels with significantly more negative potentials were observed for switch trials during the sustained potential portion of the ERP at frontal electrodes.

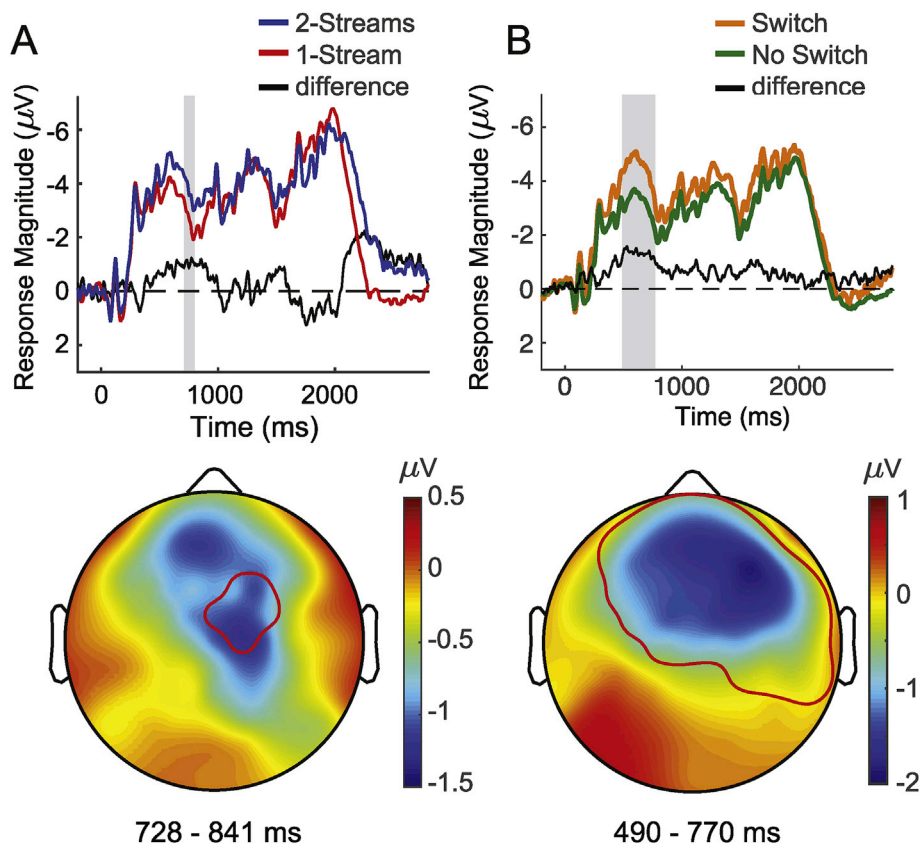


Fig. 3. Enhanced negative responses observed during the sustained potential for comparisons of A) stable 2-versus 1-stream percept, B) switch versus no-switch in perception. Top row: average ERP waveforms corresponding to contiguous clusters exhibiting significant differences (indicated by the red contour on the topographic map), and the difference wave for each comparison (black line). Gray shading indicates time range of significance differences for clusters within the red contour. Bottom row: topography of the difference between perceptual states. Red contour corresponds to regions of spatially adjacent clusters exhibiting a significant difference within the time range indicated for each map.

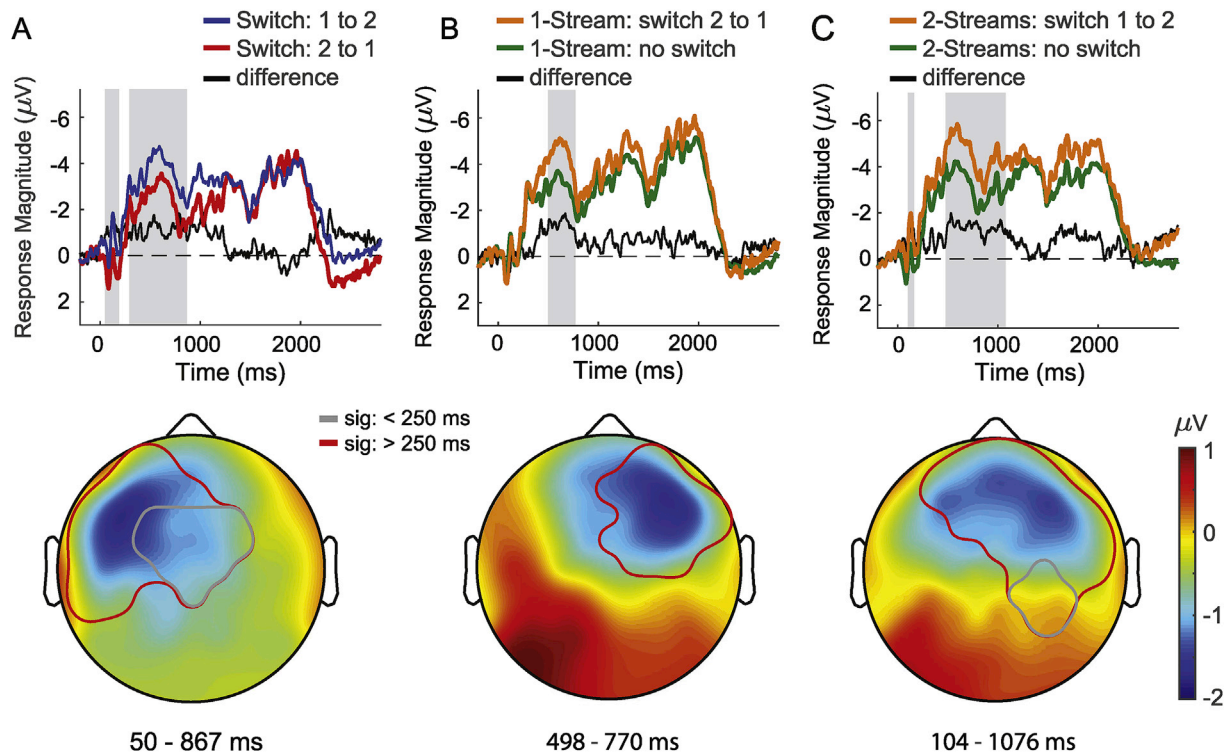


Fig. 4. Enhanced negative responses observed during the sustained potential for comparisons of A) switch from 1- to 2-streams versus 2- to 1-stream percept, B) switch from 2- to 1-stream versus stable 1-stream percept, C) switch from 1- to 2-streams versus stable 2-stream percept. Top row: average ERP waveforms corresponding to contiguous clusters exhibiting significant differences (indicated by the red/gray contour on the topographic map), and the difference wave for each comparison (black line). Gray shading indicates time range of significance differences for clusters within the red contour. Bottom row: topography of the difference between perceptual states. Red/gray contours correspond to regions of spatially adjacent clusters exhibiting a significant difference within time range indicated for each map.

3.4. Source analysis

Symmetric pairs of dipoles located bilaterally in auditory cortices and parietal lobes (Fig. 5A) accounted for 96.1% of the variance in the scalp data observed across all combined perceptual states measured over a large portion of the epoch (90–2100 ms post-stimulus). The time range includes transient responses (N1, P2) as well as the later sustained potential, encompassing all three triplets of the trial. This solution was applied to each of the individual perceptual states retaining the original dipole orientations and over the same time range: 1-stream, 2-stream, switch 1 to 2, and switch 2 to 1, and in all cases explained a large proportion of the variance (Explained Variance > 0.87; Table 1) for each perceptual state. Source waveforms qualitatively replicated the results presented in Figs. 3 and 4 (Fig. 5C, top row): 2-stream activity was greater than 1-stream, and switching perceptual states had more activity in the sustained potential than non-switching perceptual states (Fig. 5B).

Sources in anterior STG alone also accounted for a large portion of the overall sustained potential response variance (Explained Variance > 0.86; Fig. 5C, Table 1), but source waveforms isolated from these dipoles alone poorly reflect the difference between 1- and 2-stream percepts observed in Figs. 3A and 4A (Fig. 5C, middle row). Interestingly, differences between 1- and 2-stream percepts is best reflected in the parietal sources, specifically located in medial parietal cortex. The closest cortical areas to these sources are precuneus and posterior cingulate cortex, regions associated with Gestalt-type integration of features into coherent objects (Pflugshaupt et al., 2016), and the dorsal attention network (Raichle et al., 2001), respectively.

4. Discussion

To explore stable and dynamic aspects of conscious auditory perception, we performed an intermittent ABA₂ auditory streaming

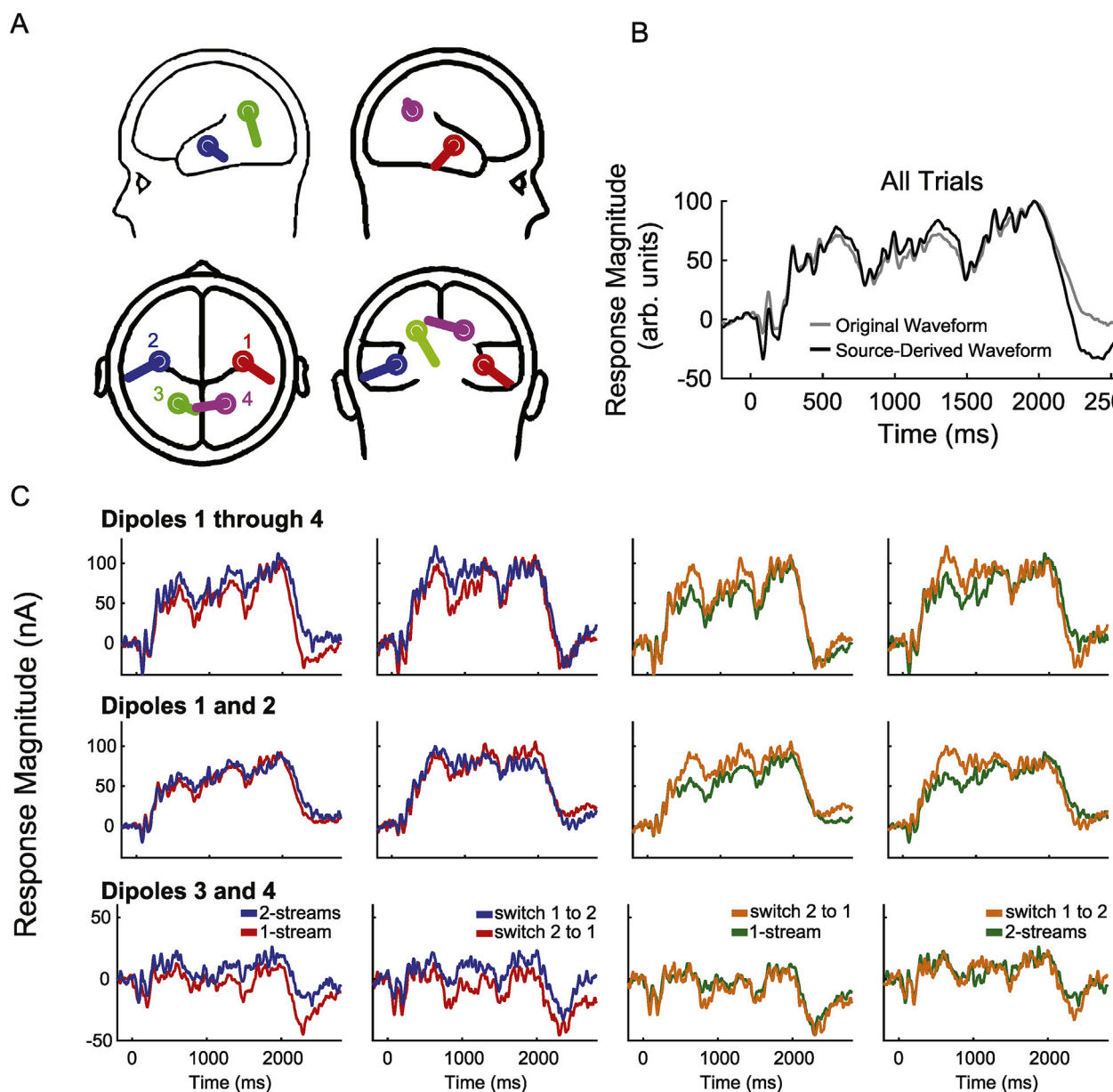


Fig. 5. Source analysis of the sustained potential. A) Symmetric dipoles located in anterior STG in left (blue) and right (red) hemispheres and parietal lobe in left (green) and right (cyan) hemispheres. B) GFP (Global Field Power) of the original waveform representing all combined perceptual states of the experiment (gray) and the source-derived waveform (black). C) GFP of source waveforms corresponding to separate perceptual states for sources in anterior STG and parietal lobe (dipoles 1, 2, 3, 4, top row), anterior STG only (dipoles 1 and 2, middle row), and parietal lobe only (dipoles 3 and 4, bottom row).

Table 1

Source analysis of scalp ERPs across all perceptual states yielded a solution with symmetric dipoles in bilateral anterior STG and bilateral parietal lobe (Fig. 5, dipoles 1, 2, 3, and 4), and a solution consisting solely of sources in bilateral anterior STG (Fig. 5, dipoles 1 and 2). These solutions when applied separately to each perceptual state, accounted for the indicated percentage of the variance.

Perceptual State	Variance Explained	
	anterior STG, Parietal Lobe	anterior STG
All Combined	96.12	94.74
1-Stream	91.04	89.08
2-Streams	87.05	85.61
Switch 2 to 1	89.25	86.76
Switch 1 to 2	91.44	88.62

experiment. Presenting the auditory stimuli in relatively discrete segments helped us identify modulations of the sustained potential during a switch in perception compared to stable periods, independent of switch direction. The sustained potential also reflected the contents of perception, namely whether participants were perceiving one vs. two streams, during the stable periods.

4.1. Behavioral response patterns

The ABA₁ streaming stimulus has been used extensively for experiments on auditory scene analysis and is commonly conducted using one of two general approaches. The first is a continuous-presentation design in which participants constantly indicate perception via button press over the time-course of multiple minutes (Anstis and Saida, 1985; Carl and Gutschalk, 2013; Denham et al., 2018; Pressnitzer and Hupé, 2006). The second typically consists of a two-part sequence with an induction period followed by a test period in which a manipulation of the ABA₁ stimulus along one or more dimensions (e.g., temporal, spectral, location) serves as a probe for perceptual continuity (Haywood and Roberts, 2010; Rogers and Bregman, 1993; Yerkes et al., 2019). The first approach accommodates for the observation of spontaneous switching of perception over an extended period of time, while the second provides a better-defined event associated with a perceptual switch. Despite the temporal discontinuity, the current findings follow established behavioral patterns characteristic of the continuous button-response paradigm: balanced time for each percept (Fig. 2A), an initial percept of 1-stream characterized by longer duration (Fig. 2A and B), a logarithmically shaped distribution of phase duration (Fig. 2C), and lack of correlation between sequential phase durations (Fig. 2D). As a result, the intermittent-response paradigm tested here incorporates benefits from each paradigm type, the observation of spontaneous switching behavior over time, and a well-defined stimulus event for linking perception to modulations of ERPs.

4.2. Sustained potential

Most of the ERP differences observed between 1- versus 2-streams (Figs. 3A and 4A) and switch versus non-switch perceptual states (Figs. 3B, 4B and 4C) were observed during the portion of the waveform considered to be the auditory sustained potential. This brain response is characterized by negative voltage at frontal scalp locations following presentation of continuous auditory stimulation (Kohler and Wegener, 1955; Picton et al., 1978a, 1978b; Scherg et al., 1989). Unlike earlier responses to sound onsets and offsets that exhibit more transient positive and negative deflections in the 75–200 ms range, the sustained potential is unaffected by mixed presentations of click and tone-burst stimuli, and in the context of auditory stream segregation has been shown to be sensitive to attention and features of the ABA₁ tones such as frequency separation (Snyder et al., 2006).

In prior work, source analysis of the underlying neural generators of the sustained potential revealed bilateral, vertically oriented dipoles in

anterior STG (Scherg et al., 1989). In agreement with this finding, the current study found optimized dipoles located bilaterally in anterior STG. This region has been linked to the “what” part of the “what/where” dual-pathway model for sound pattern identification (Ahveninen et al., 2013; Bizley and Cohen, 2013; Rauschecker and Tian, 2000; Zündorf et al., 2016). Thus, our results support the hypothesis that the differences observed in the sustained potential are related to auditory streaming, a process whereby stimulus features are integrated or segregated resulting in designation of a 1- or 2-auditory stream percept, respectively. Evidence for separate processes is observed in the ERPs, where differences between switch and no-switch trials (Fig. 3B) appear earlier and extend over a longer time range compared to differences between representation of 1- and 2-stream perception (Fig. 3A). This timing difference is also apparent where auditory cortical sources showed greater separation for switch versus no-switch perceptual states earlier in the waveform (Fig. 5C; middle row), whereas parietal sources exhibited the most separation for 1- versus 2-streams later in the waveform (Fig. 5C; bottom row). Parietal sources differentiating 1- versus 2-streams are consistent with fMRI studies (Cusack, 2005; Teki et al., 2011), however, the locations within the parietal lobe observed here are not in the same location as those fMRI studies, and it is important to note that the majority of variance-explained was accounted for by the sources in anterior STG (Table 1). In summary, the ERP differences presented here suggest separate processes underlie the neural signatures for the 1- versus 2-stream difference and the switch versus no-switch difference. The source analysis presented here, while not statistically comprehensive, does provide a firm qualitative assessment of the neural activity observed in this experiment, valuable for contextualizing the results with other studies using alternate auditory streaming paradigms and imaging modalities.

4.3. Neural representation of perceptual switches

Switch-related neural activity has been studied previously, notably using fMRI. Kondo and Kashino (2009) used an ABA₁ streaming stimulus to demonstrate brain region specific timing differences in the BOLD signal. Switches in perception from 2- to 1-stream evoked earlier activations in auditory thalamic voxels compared to switches in the opposite direction. Conversely, voxels in auditory cortex displayed earlier activity corresponding to switches from 1- to 2-streams. This finding is partially supported by a series of fMRI studies by Schadwinkel and Gutschalk that utilized the spatial cue carried by interaural time difference to cause stream segregation. Transient responses in the auditory cortical BOLD signal tended to be greater when switching from 1- to 2-streams, than for switches back to 1-stream (Schadwinkel and Gutschalk, 2010). In a different study also using ITD cues to promote stream segregation, transient responses in the inferior colliculus had large responses to switches in both directions (Schadwinkel and Gutschalk, 2011). Though somewhat mixed, the pattern of results from fMRI studies combined with the data presented here suggests the auditory cortex plays a prominent role in the representation of perceptual switching, and that there is likely a subcortical component related to switch direction.

4.4. Implications for theoretical models of bistability

In the face of ambiguous stimuli, the conventional dynamic model for bistable perception proposes that populations of neurons representing different states incorporate inhibition, adaptation, and noise to generate bistability or multistability (Brascamp et al., 2018; Rankin et al., 2017, 2015; Tong et al., 2006). Inhibition leads to competition between different states, and adaptation and noise allow for switches between the currently dominant state. As a steady percept is maintained, adaptation reduces responses until a threshold is crossed, at which time the non-dominant percept overtakes the dominant percept and a switch in perception is triggered.

Within the framework of this model, the results of this experiment may be interpreted as follows. The large potentials observed at the very

beginning of a percept, while termed a “switch” in this experiment, also correspond to a fresh percept prior to the forthcoming effects of adaptation. As the percept proceeds, adaptation quickly builds up until a switch occurs and another fresh percept (with a large ERP) emerges. This interpretation is supported by the observation in the results that significant differences related to a reversal in perception (regardless of direction) appear during the portion of the waveform corresponding to the first triplet (Figs. 3B, 4B–C). Studies measuring single-unit activity in response to repeated ABA triplets have demonstrated decreases in spike rate following the initial triplet, with the largest difference typically observed between the first and second triplets (Michey et al., 2005; Pressnitzer et al., 2008). The data presented here comport with this observation, in that the first triplet of each percept evokes the greatest response compared to the first triplet of subsequent trials. These studies also report additional spike rate decay (though not as extreme) in response to continued presentation of triplets. In the present study, evidence of progressively increasing adaptation over the percept time-course (apart from the initial adaptation) was not observed. This could be due to insufficient spatial sensitivity of EEG or the differing nature of what kind of neural activity is reflected in single-unit recordings vs. EEG.

From a modeling perspective, such a sequence of events could be used to differentiate switches from non-switch periods on the basis of the amount of adaptation and inhibition present. The present finding that the signature for switches in perception began earlier in the waveforms and in more sensory regions than the signature for the contents of perception (1-stream vs. 2-stream; Fig. 3A compared to 3B; Fig. 5C, bottom row compared to middle row) has implications for the locus of sources of adaptation, inhibition, and noise that may drive bistability. Several existing models of bistability in auditory streaming are consistent with an early locus for bistability (Rankin et al., 2017, 2015), while others assume bistability occurs as part of the process of identifying the number of sources in an auditory stream (Barniv and Nelken, 2015; Mill et al., 2013); these latter models do not appear to be consistent with our data because they predict a similar locus for switching and recognition of 1- vs. 2-stream percepts. Note that none of the existing auditory models account for the effects of the 700 ms break present in the current study, but could easily be modified on the basis of computational studies that do consider the effects of gaps (Noest et al., 2007; Rankin et al., 2017; Vattikuti et al., 2016). In summary, models that allow for early sensory sources of bistability (Noest et al., 2007; Vattikuti et al., 2016) or those that posit some form of top-down modulation of sensory competition (Brascamp et al., 2018; Kleinschmidt et al., 2012; Li et al., 2017) appear most consistent with our data.

5. Conclusion

In this study, we present data from an auditory streaming experiment using an intermittent stimulus paradigm that showed behavioral characteristics consistent with continuous bistable perception while maintaining control of the temporal dynamics important for recording ERPs. Consistent with previous studies, sustained auditory potentials associated with perception of 2-streams exhibited greater negative potentials than 1-stream. Unexpectedly, sustained potentials were significantly more negative when a perceptual switch occurred, regardless of the switch direction, leading to the conclusion that perceptual switches have a neural correlate unique from the overall representation of 1-stream or 2-streams. Importantly, the ability to tease apart the neural correlates associated with a) an internally derived event (a switch in perception) and b) an ongoing perceptual representation, can be attributed to the unique intermittent design employed in this experiment.

Acknowledgements

Supported by Office of Naval Research, United States: N00014-16-1-2879.

References

- Ahveninen, J., Huang, S., Nummenmaa, A., Belliveau, J.W., Hung, A.-Y., Jääskeläinen, I.P., Rauschecker, J.P., Rossi, S., Tiitinen, H., Raij, T., 2013. Evidence for distinct human auditory cortex regions for sound location versus identity processing. *Nat. Commun.* 4, 2585. <https://doi.org/10.1038/ncomms3585>.
- Anstis, S., Saida, S., 1985. Adaptation to auditory streaming of frequency-modulated tones. *J. Exp. Psychol. Hum. Percept. Perform.* 11 (3), 257–271. <https://doi.org/10.1037/0096-1523.11.3.257>.
- Arnott, S.R., Binns, M.A., Grady, C.L., Alain, C., 2004. Assessing the auditory dual-pathway model in humans. *Neuroimage* 22, 401–408. <https://doi.org/10.1016/j.neuroimage.2004.01.014>.
- Baars, B.J., 1988. *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Barniv, D., Nelken, I., 2015. Auditory streaming as an online classification process with evidence accumulation. *PLoS One* 10, e0144788. <https://doi.org/10.1371/journal.pone.0144788>.
- Billig, A.J., Davis, M.H., Carlyon, R.P., 2018. Neural decoding of bistable sounds reveals an effect of intention on perceptual organization. *J. Neurosci. Off. J. Soc. Neurosci.* 38, 2844–2853. <https://doi.org/10.1523/JNEUROSCI.3022-17.2018>.
- Bizley, J.K., Cohen, Y.E., 2013. The what, where and how of auditory-object perception. *Nat. Rev. Neurosci.* 14, 693–707. <https://doi.org/10.1038/nrn3565>.
- Brascamp, J., Sterzer, P., Blake, R., Knapen, T., 2018. Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annu. Rev. Psychol.* 69, 77–103. <https://doi.org/10.1146/annurev-psych-010417-085944>.
- Bregman, A.S., 1990. *Auditory Scene Analysis: the Perceptual Organization of Sound* (Dissertation). MIT, Cambridge, MA.
- Bregman, A.S., 1978. Auditory streaming: competition among alternative organizations. *Percept. Psychophys.* 23, 391–398.
- Carl, D., Gutschalk, A., 2013. Role of pattern, regularity, and silent intervals in auditory stream segregation based on inter-aural time differences. *Exp. Brain Res.* 224, 557–570. <https://doi.org/10.1007/s00221-012-3333-z>.
- Changeux, J.P., Dehaene, S., 2008. The neuronal workspace model: conscious processing and learning. <https://doi.org/10.1016/B978-012370509-9.00078-4>.
- Curtu, R., Wang, X., Brunton, B.W., Nourski, K.V., 2019. Neural signatures of auditory perceptual bistability revealed by large-scale human intracranial recordings. *J. Neurosci. Off. J. Soc. Neurosci.* <https://doi.org/10.1523/JNEUROSCI.0655-18.2019>.
- Cusack, R., 2005. The intraparietal sulcus and perceptual organization. *J. Cogn. Neurosci.* 17, 641–651. <https://doi.org/10.1162/0899829053467541>.
- Cusack, R., Deeks, J., Aikman, G., Carlyon, R.P., 2004. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 643–656. <https://doi.org/10.1037/0096-1523.30.4.643>.
- Dehaene, S., Changeux, J.-P., 2011. Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. <https://doi.org/10.1016/j.neuron.2011.03.018>.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Denham, S.L., Farkas, D., van Ee, R., Taranu, M., Kocsis, Z., Wimmer, M., Carmel, D., Winkler, I., 2018. Similar but separate systems underlie perceptual bistability in vision and audition. *Sci. Rep.* 8, 7106. <https://doi.org/10.1038/s41598-018-25587-2>.
- Denham, S.L., Gyimesi, K., Stefanics, G., Winkler, I., 2013. Perceptual bistability in auditory streaming: how much do stimulus features matter? *Learn. Percept.* 5, pp. 73–100. <https://doi.org/10.1556/LP.5.2013.Supp1.6>.
- DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? *Neuron* 73, 415–434. <https://doi.org/10.1016/j.neuron.2012.01.010>.
- Dykstra, A.R., Cariani, P.A., Gutschalk, A., 2017. A roadmap for the study of conscious audition and its neural basis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372. <https://doi.org/10.1098/rstb.2016.0103>.
- Farkas, D., Denham, S.L., Winkler, I., 2018. Functional brain networks underlying idiosyncratic switching patterns in multi-stable auditory perception. *Neuropsychologia* 108, 82–91. <https://doi.org/10.1016/j.neuropsychologia.2017.11.032>.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Gutschalk, A., Micheyl, C., Melcher, J.R., Rupp, A., Scherg, M., Oxenham, A.J., 2005. Neuromagnetic correlates of streaming in human auditory cortex. *J. Neurosci. Off. J. Soc. Neurosci.* 25, 5382–5388. <https://doi.org/10.1523/JNEUROSCI.0347-05.2005>.
- Gutschalk, A., Micheyl, C., Oxenham, A.J., 2008. Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol.* 6, e138. <https://doi.org/10.1371/journal.pbio.0060138>.
- Haywood, N.R., Roberts, B., 2013. 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.* 39, 1652–1666. <https://doi.org/10.1037/a0032562>.
- Haywood, N.R., Roberts, B., 2010. Build-up of the tendency to segregate auditory streams: resetting effects evoked by a single deviant tone. *J. Acoust. Soc. Am.* 128, 3019–3031. <https://doi.org/10.1121/1.3488675>.
- Hochstein, S., Ahissar, M., 2002. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.
- Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178.
- Kleinschmidt, A., Sterzer, P., Rees, G., 2012. Variability of perceptual multistability: from brain state to individual trait. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 988–1000. <https://doi.org/10.1098/rstb.2011.0367>.

- Kohler, W., Wegener, J., 1955. Currents of the human auditory cortex. *J. Cell. Physiol. Suppl.* 45, 25–54.
- Kondo, H.M., Kashino, M., 2009. Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. *J. Neurosci. Off. J. Soc. Neurosci.* 29, 12695–12701. <https://doi.org/10.1523/JNEUROSCI.1549-09.2009>.
- Kondo, H.M., Pressnitzer, D., Shimada, Y., Kochiyama, T., Kashino, M., 2018. Inhibition-excitation balance in the parietal cortex modulates volitional control for auditory and visual multistability. *Sci. Rep.* 8, 14548. <https://doi.org/10.1038/s41598-018-32892-3>.
- Kormmeier, J., Bach, M., 2004. Early neural activity in Necker-cube reversal: evidence for low-level processing of a gestalt phenomenon. *Psychophysiology* 41, 1–8. <https://doi.org/10.1046/j.1469-8986.2003.00126.x>.
- Leopold, D.A., Logothetis, N.K., 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379, 549–553. <https://doi.org/10.1038/379549a0>.
- Li, H.-H., Rankin, J., Rinzel, J., Carrasco, M., Heeger, D.J., 2017. Attention model of binocular rivalry. *Proc. Natl. Acad. Sci. U. S. A* 114, E6192–E6201. <https://doi.org/10.1073/pnas.1620475114>.
- Lomber, S.G., Malhotra, S., 2008. Double dissociation of “what” and “where” processing in auditory cortex. *Nat. Neurosci.* 11, 609–616. <https://doi.org/10.1038/nn.2108>.
- Lumer, E.D., Friston, K.J., Rees, G., 1998. Neural correlates of perceptual rivalry in the human brain. *Science* 280, 1930–1934.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Micheyl, C., Tian, B., Carlyon, R.P., Rauschecker, J.P., 2005. Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48, 139–148. <https://doi.org/10.1016/j.neuron.2005.08.039>.
- Mill, R.W., Böhm, T.M., Bendixen, A., Winkler, I., Denham, S.L., 2013. Modelling the emergence and dynamics of perceptual organisation in auditory streaming. *PLoS Comput. Biol.* 9, e1002925. <https://doi.org/10.1371/journal.pcbi.1002925>.
- Milner, A.D., Goodale, M.A., 2008. Two visual systems re-viewed. *Neuropsychologia* 46, 774–785. <https://doi.org/10.1016/j.neuropsychologia.2007.10.005>.
- Noest, A.J., van Ee, R., Nijs, M.M., van Wezel, R.J.A., 2007. Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J. Vis.* 7, 10. <https://doi.org/10.1167/7.8.10>.
- Pflugshaupt, T., Nösberger, M., Gutbrod, K., Weber, K.P., Linnebank, M., Brugger, P., 2016. Bottom-up visual integration in the medial parietal lobe. *Cereb. Cortex N. Y. N* 26, 943–949, 1991. <https://doi.org/10.1093/cercor/bhu256>.
- Picton, T.W., Woods, D.L., Proulx, G.B., 1978a. Human auditory sustained potentials. I. The nature of the response. *Electroencephalogr. Clin. Neurophysiol.* 45, 186–197.
- Picton, T.W., Woods, D.L., Proulx, G.B., 1978b. Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalogr. Clin. Neurophysiol.* 45, 198–210.
- Pitts, M.A., Gavin, W.J., Nerger, J.L., 2008. Early top-down influences on bistable perception revealed by event-related potentials. *Brain Cogn.* 67, 11–24. <https://doi.org/10.1016/j.bandc.2007.10.004>.
- Pitts, M.A., Martínez, A., Hillyard, S.A., 2012. Visual processing of contour patterns under conditions of inattentive blindness. *J. Cogn. Neurosci.* 24, 287–303. https://doi.org/10.1162/jocn_a.00111.
- Pressnitzer, D., Hupé, J.-M., 2006. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr. Biol. CB* 16, 1351–1357. <https://doi.org/10.1016/j.cub.2006.05.054>.
- Pressnitzer, D., Sayles, M., Micheyl, C., Winter, I.M., 2008. Perceptual organization of sound begins in the auditory periphery. *Curr. Biol. CB* 18, 1124–1128. <https://doi.org/10.1016/j.cub.2008.06.053>.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A* 98, 676–682. <https://doi.org/10.1073/pnas.98.2.676>.
- Rankin, J., Osborn Popp, P.J., Rinzel, J., 2017. Stimulus pauses and perturbations differentially delay or promote the segregation of auditory objects: psychoacoustics and modeling. *Front. Neurosci.* 11, 198. <https://doi.org/10.3389/fnins.2017.00198>.
- Rankin, J., Sussman, E., Rinzel, J., 2015. Neuromechanistic model of auditory bistability. *PLoS Comput. Biol.* 11, e1004555. <https://doi.org/10.1371/journal.pcbi.1004555>.
- Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A* 97, 11800–11806. <https://doi.org/10.1073/pnas.97.22.11800>.
- Rogers, W.L., Bregman, A.S., 1993. An experimental evaluation of three theories of auditory stream segregation. *Percept. Psychophys.* 53, 179–189. <https://doi.org/10.3758/BF03211728>.
- Sanders, R.D., Winston, J.S., Barnes, G.R., Rees, G., 2018. Magnetoencephalographic correlates of perceptual state during auditory bistability. *Sci. Rep.* 8, 976. <https://doi.org/10.1038/s41598-018-19287-0>.
- Schadwinkel, S., Gutschalk, A., 2011. Transient bold activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. *J. Neurophysiol.* 105, 1977–1983. <https://doi.org/10.1152/jn.00461.2010>.
- Schadwinkel, S., Gutschalk, A., 2010. Functional dissociation of transient and sustained fMRI BOLD components in human auditory cortex revealed with a streaming paradigm based on interaural time differences. *Eur. J. Neurosci.* 32, 1970–1978. <https://doi.org/10.1111/j.1460-9568.2010.07459.x>.
- Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the late human auditory evoked potentials. *J. Cogn. Neurosci.* 1, 336–355. <https://doi.org/10.1162/jocn.1989.1.4.336>.
- Snyder, J.S., Alain, C., Picton, T.W., 2006. Effects of attention on neuroelectric correlates of auditory stream segregation. *J. Cogn. Neurosci.* 18, 1–13. <https://doi.org/10.1162/089892906775250021>.
- Snyder, J.S., Holder, W.T., Weintraub, D.M., Carter, O.L., Alain, C., 2009. Effects of prior stimulus and prior perception on neural correlates of auditory stream segregation. *Psychophysiology* 46, 1208–1215. <https://doi.org/10.1111/j.1469-8986.2009.00870.x>.
- Snyder, J.S., Yerkes, B.D., Pitts, M.A., 2015. Testing domain-general theories of perceptual awareness with auditory brain responses. *Trends Cogn. Sci.* 19, 295–297. <https://doi.org/10.1016/j.tics.2015.04.002>.
- Teki, S., Chait, M., Kumar, S., von Kriegstein, K., Griffiths, T.D., 2011. Brain bases for auditory stimulus-driven figure-ground segregation. *J. Neurosci. Off. J. Soc. Neurosci.* 31, 164–171. <https://doi.org/10.1523/JNEUROSCI.3788-10.2011>.
- Tong, F., Engel, S.A., 2001. Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* 411, 195–199. <https://doi.org/10.1038/35075583>.
- Tong, F., Meng, M., Blake, R., 2006. Neural bases of binocular rivalry. *Trends Cogn. Sci.* 10, 502–511. <https://doi.org/10.1016/j.tics.2006.09.003>.
- Tong, F., Nakayama, K., Vaughan, J.T., Kanwisher, N., 1998. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21, 753–759.
- Tononi, G., Boly, M., Massimini, M., Koch, C., 2016. Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. <https://doi.org/10.1038/nrn.2016.44>.
- Van Noorden, L.P.A.S., 1975. *Temporal Coherence in the Perception of Tone Sequences (Unpublished Doctoral Dissertation)*. Eindhoven University of Technology, Eindhoven.
- Vattikuti, S., Thangaraj, P., Xie, H.W., Gotts, S.J., Martin, A., Chow, C.C., 2016. Canonical cortical circuit model explains rivalry, intermittent rivalry, and rivalry memory. *PLoS Comput. Biol.* 12, e1004903. <https://doi.org/10.1371/journal.pcbi.1004903>.
- Yerkes, B.D., Weintraub, D.M., Snyder, J.S., 2019. Stimulus-based and task-based attention modulate auditory stream segregation context effects. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 53–66. <https://doi.org/10.1037/xhp0000587>.
- Zündorf, I.C., Lewald, J., Karnath, H.-O., 2016. Testing the dual-pathway model for auditory processing in human cortex. *Neuroimage* 124, 672–681. <https://doi.org/10.1016/j.neuroimage.2015.09.026>.