

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/326813213>

# Neuronal correlates of auditory streaming in the auditory cortex of behaving monkeys

Article in *European Journal of Neuroscience* · August 2018

DOI: 10.1111/ejn.14098

CITATION

1

READS

138

4 authors:



**Elena Selezneva**

Leibniz Institute for Neurobiology

16 PUBLICATIONS 524 CITATIONS

SEE PROFILE



**Alexander Georgievich Gorkin**

Russian Academy of Sciences

32 PUBLICATIONS 106 CITATIONS

SEE PROFILE



**Eike Budinger**

Leibniz Institute for Neurobiology

44 PUBLICATIONS 1,485 CITATIONS

SEE PROFILE



**Michael Brosch**

Leibniz-Gemeinschaft

58 PUBLICATIONS 3,994 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Ischaemia and plasticity [View project](#)



Transcranial stimulation [View project](#)

# Neuronal correlates of auditory streaming in the auditory cortex of behaving monkeys

Elena Selezneva  | Alexander Gorkin | Eike Budinger | Michael Brosch

Leibniz Institut für Neurobiologie,  
Magdeburg, Germany

## Correspondence

Elena Selezneva, Speziallabor  
Primatenneurobiologie, Leibniz-Institut für  
Neurobiologie, Brenneckestraße 6, 39118  
Magdeburg, Germany.  
Email: brosch@lin-magdeburg.de

## Funding information

Deutsche Forschungsgemeinschaft, Grant/  
Award Number: SFB TR 31, He 1721/10-1  
and He 1721/10-2

## Abstract

This study tested the hypothesis that spiking activity in the primary auditory cortex of monkeys is related to auditory stream formation. Evidence for this hypothesis was previously obtained in animals that were passively exposed to stimuli and in which differences in the streaming percept were confounded with differences between the stimuli. In this study, monkeys performed an operant task on sequences that were composed of light flashes and tones. The tones alternated between a high and a low frequency and could be perceived either as one auditory stream or two auditory streams. The flashes promoted either a one-stream percept or a two-stream percept. Comparison of different types of sequences revealed that the neuronal responses to the alternating tones were more similar when the flashes promoted auditory stream integration, and were more dissimilar when the flashes promoted auditory stream segregation. Thus our findings show that the spiking activity in the monkey primary auditory cortex is related to auditory stream formation.

## KEYWORDS

audiovisual, bistable perception, population separation model, rivalry, temporal processing

## 1 | INTRODUCTION

Auditory scene analysis is the process by which the auditory system groups and segregates sounds of the environment into perceptual streams (Bregman, 1990). The rules governing primitive auditory stream formation are frequently tested with rapid sequences of alternating sounds. This has revealed that, depending on the similarity and the rate of presentation, the sounds may be perceived as being part of a single auditory stream or of two or more streams (Moore & Gockel, 2002). In studies in which, for example, sinusoidal tones are used, subjects tend to perceive two streams when differences

in frequency are large and presentation rates are high, while subjects tend to perceive a single stream when differences in frequency are small and rates are low. Similar observations are made when sounds differ in other features of sounds, for example, temporal envelope.

The highly influential study of Fishman and colleagues has revealed a neuronal correlate of auditory stream formation in the monkey primary auditory cortex (Fishman, Arezzo, & Steinschneider, 2004; Fishman, Reser, Arezzo, & Steinschneider, 2001). The researchers used sequences composed of sounds that alternated between a high- and a low-frequency tone and set one of them (say the high tone) at the best frequency of a neuron and the other (say the low tone) at a different frequency. In addition, the rate of presentation and the duration of tones were varied. When they computed the ratio of neuronal response amplitudes of the low tone to the high tone, they found that this ratio varied with the proportion of time humans hear either one or more auditory streams. Largest ratios were associated

**Abbreviations:** A/D, analog-digital; AV, audiovisual; BOLD, blood oxygenation level dependent; D/A, digital-analog; dB, decibel; HCl, hydrochloride; Hz, hertz; kHz, kilohertz; LED, light-emitting diode; PSTH, peri stimulus time histogram; SPL, sound pressure level.

Edited by Christoph M. Michel. Reviewed by Mitchell Steinchneider and Pierre Megevand.

All peer review communications can be found with the online version of the article.

with stimulus conditions in which subjects perceived a single auditory stream. Ratios were smaller under ambiguous stimulus conditions, in which the subjects' perception switched between one and two auditory streams. The smallest ratios were observed under stimulus conditions in which subjects perceived two auditory streams. Based on these observations, Fishman et al. (2001) put forward the population separation model of auditory streaming according to which a tone sequence is perceived as a single stream when the high and low tones activate the same neural population in primary auditory cortex. In contrast, a tone sequence is perceived as two streams when the high and low tones activate different neural populations along the tonotopic axis of primary auditory cortex.

A problem for the interpretation of the results obtained by Fishman et al. (2001) is that the correlation between response and auditory streaming was confounded using different auditory stimuli for the different percepts. Therefore it is not clear whether the response ratio in auditory cortex reflected the percept, or merely resulted from different frequencies. One way to distinguish neuronal activity related to perceptual processes from that related to physical stimulus characteristics is to use stimuli that allow more than one percept and to control the percept.

We have recently established a behavioral testing paradigm in which visual stimuli are used to assess, or modify, how humans and nonhuman primates perceptually organize tone sequences (Selezneva et al., 2012). In this paradigm, subjects are presented with a periodic sequence of light flashes and are required to report, by bar press, when the flashing is terminated. When the flash sequence is synchronized with a tone sequence alternating between two frequencies and when the tones are terminated before the flashes, subjects are able to respond faster to the termination of the flashes. This indicates that the subjects also attend to the auditory stimuli and use information from both the visual and the auditory modality for their behavioral responses. Most importantly, the perceptual merging of stimuli from the two modalities is associated with some modifications of the perception of unimodal stimuli. On the one hand, the perceptual organization of the tone sequence has an effect on how quickly the termination of the flash sequence is detected such that the change in response time allows to infer whether the tone sequence is perceived more frequently as one or two auditory streams. On the other hand, the visual stimuli affect the perceptual organization of the tone sequence: synchronizing the flashes with every second tone increases the probability that the tone sequence is perceived as two auditory streams. In contrast, synchronizing the flashes with every third tone increases the probability that the tone sequence is perceived as one auditory stream. These differential effects were explained by different strengths of perceptual binding between the tone sequence and the flash sequence, that is, by the match of the perceptual organizations of the two sequences.

Using this behavioral paradigm we recorded the firing of neurons from the auditory cortex of monkeys and tested whether the ratio of neuronal responses to alternating tones changes when the probability changes with which monkeys perceive a tone sequence as one or two auditory streams. Specifically, we hypothesized that the response ratio is higher in situations in which the tone sequence is more frequently perceived as one auditory stream as compared to situations in which the tone sequence is more frequently perceived as two auditory streams. We focused on primary auditory cortex because the original observations on neuronal correlates of auditory stream formation were made in this part of auditory cortex (Fishman et al., 2001). This area is also known to be affected by visual stimuli (Brosch, Selezneva, & Scheich, 2005, 2015; Schroeder & Lakatos, 2009). Here we report results from experiments that were optimized to compare response ratios when the perceptual organization of the tone sequence was modified by the flash sequences. The results obtained in experiments that were optimized to infer the perceptual organization of the tone sequence from the gain in behavioral response time will be presented elsewhere.

## 2 | METHODS

### 2.1 | Subjects

Experiments were performed on three adult male long-tailed macaques (*Macaca fascicularis*), monkey E, monkey W, and monkey C. All monkeys previously participated in behavioral studies using audiovisual sequences very similar to those used here (described in Rahne et al., 2008; Selezneva et al., 2012). As detailed in Brosch & Scheich, 2008, a headholder ("halo") device was implanted under deep anesthesia (Ketamine HCl [4 mg/kg] and Xylazine [5 mg/kg]) onto the monkeys' head to allow atraumatic head restraint. It consisted of three strong arches that closely encircled the occipital, supra-orbital and mid-sagittal ridges of the head. This helmet-like piece was firmly and permanently attached to the head by means of several counteracting stainless steel bolts with sharpened points, which were advanced by rotation through the intact skin and soft tissue until they lodged firmly against the skull. Subsequently, monkeys received a chamber implant, positioned in the left temporal regions of the skull, for microelectrode recordings from the auditory cortex. For the implantation, a piece of bone was removed with a trephine (diameter: 21 mm) and an externally threaded stainless steel cylinder was screwed into the slightly undersized hole. All surgical operations on the animals were performed under deep general anesthesia and were followed by a full course of antibiotic (Baytril, 0.2 ml/kg) and analgesic (Carprofen, 0.1 ml/kg) treatment during which they were monitored several times per day

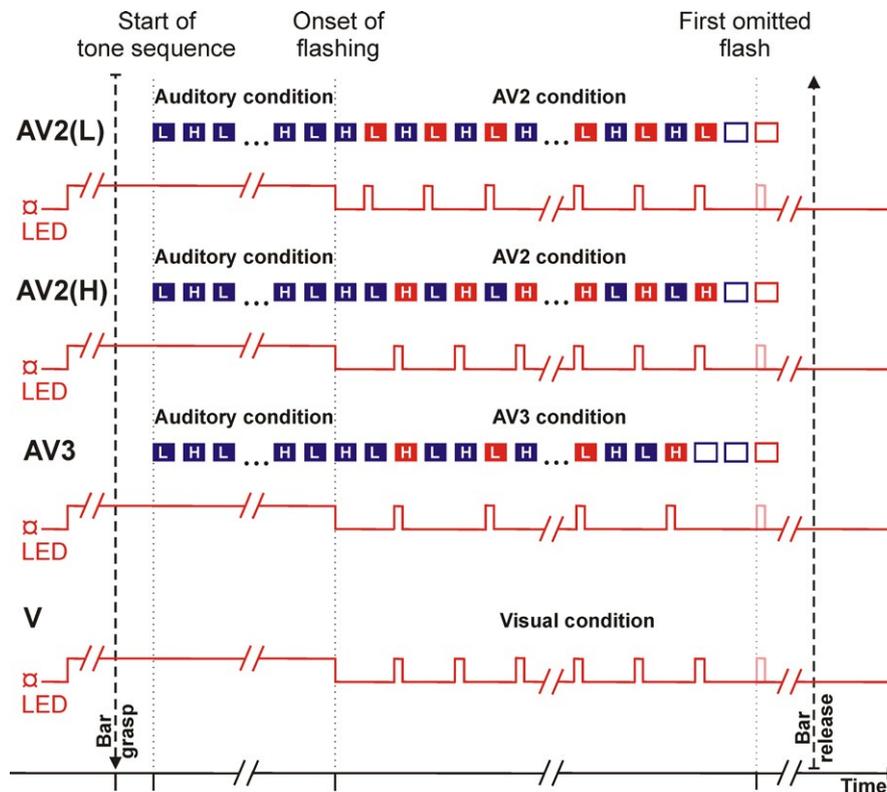
as long as required. All experiments were approved by the Animal Care and Ethics Authority of the State of Saxony-Anhalt (Landesverwaltungsamt Halle) and conformed to EU Directive (2010/63/EU) on the protection of animals used for scientific purposes.

## 2.2 | Apparatus

Experiments were carried out in a double-walled soundproof room (IAC 1202-A). For all behavioral training and testing, a monkey squatted in a primate chair. Its front compartment accommodated a green light-emitting diode for visual stimulation (LED, 2 visual degrees diameter), a touch bar, and a water-spout (see also Brosch, Selezneva, Bucks, & Scheich, 2004). For auditory stimulation two loudspeakers (Canton Karat 720) were placed at a distance of ~1 m on the left and the right side of a monkey.

A computer was used to generate visual and auditory stimuli and to monitor and control the behavioral procedure. Auditory stimuli were generated with the aid of an array processor (Tucker-Davis Technologies, Gainesville), interfaced with the computer, at a sampling rate of 100 kHz. The signal was D/A converted, amplified (Pioneer, A202) and presented to a monkey through the loudspeakers.

We used a multichannel drive equipped with up to five microelectrodes (Thomas Recording) to record the action potentials fired by a small group of neurons (multiunit activity) from different sites, mostly in upper layers, of the left primary auditory cortex. Following preamplification, the signals from each electrode were amplified and filtered at 0.5–5 kHz with a PGMA-64 (Thomas Recordings), and recorded with an A/D data acquisition system (Neuralynx) at a sampling rate of 42 kHz. For technical reasons, local field potentials were not available in this study.



**FIGURE 1** Schematic representation of the behavioral task and the stimuli used to modify the perceptual organization of tone sequences. After illumination of the LED light (indicated by the step on the lower red trace), the monkeys had to grasp a touch bar. After holding the bar for some time, the LED started to be periodically turned off and on (onset of flashing). This period constituted the visual condition. The monkeys had to release the bar when the periodic flashing of the LED had stopped, that is, after the first omitted flash (indicated by the gray shading). In some trials (shown in the upper parts of the figure) auditory stimuli (filled rectangles) were presented in combination with the visual stimuli. They consisted of a series of pure tones that alternated periodically between a low (L) and high (H) frequency. They started during the steady-state phase of visual stimulation (and constituted the auditory condition) and continued during the subsequent flashing phase in which either every second tone (constituting the AV2 condition) or every third tone (constituting the AV3 condition) was synchronized with the on-phase of the LED. The two AV2 conditions differed in whether the low or the high tone was synchronized with a flash. In most trials, tone presentation stopped before the flashing was stopped (indicated by the open rectangles). In the AV2 condition, there was one omitted tone before the first omitted flash. In the AV3 condition, there were two omitted tones before the first omitted flash

## 2.3 | Stimuli

To measure the frequency sensitivity and best frequency of neurons, we presented the monkeys with a random sequence of 400 pure tones at 40 different frequencies (from 200 Hz to 15.6 kHz in equal logarithmical steps). The tones had a duration of 100 ms and were presented every 1,000 ms. The sound pressure level was ~70 dB SPL.

For the behavioral task, visual stimuli were presented either alone or in combination with auditory stimuli (Figure 1). The auditory stimuli were selected following earlier testing such that each of the three monkeys participating in the current study was presented with perceptually ambiguous tone sequences. For the visual stimulation, the LED was initially turned on for maximally 10 s. This was followed by a period of 2.52–6 s during which the LED was periodically turned off and on 9–25 times. The on-phase of the LED had a duration of 35 ms and the off-phase had a fixed duration during a trial. In monkey C, the LED was turned on every 280 or 420 ms. In monkey W and E, the LED was turned on every 240 or 360 ms. In audiovisual trials, the auditory stimulation started about 1–3 s after the LED had been turned on for the first time in a trial. It lasted 3.48 and 9.6 s and consisted of 25–80 pure tones that alternated between a high and a low frequency. Each tone had a duration of 80 ms (including 5-ms rise/fall time) and was presented every 140 ms in monkey C and every 120 ms in monkey W and monkey E, respectively. The intensity of the tone bursts was ~60 dB SPL. Six to twenty-four of the tones were presented during the initial on-phase of the LED, and this period constituted the auditory condition. Another 18–49 tones were presented during the subsequent period in which the LED was periodically flashed for 2.52–6 s. This period constituted the audiovisual condition. Two types of audiovisual conditions were used that differed in the way of how flashes and tones were synchronized. In the AV2 condition, the flashes were presented with every second tone. In some trials, the flashes were synchronized with the high tones only, and in other trials, the synchronization was always with the low tones only. In the AV3 condition, the flashes were presented with every third tone, such that the flashes alternated between being synchronized with a high and a low tone. The flashes were turned on 9 ms before the corresponding tone to promote the perceptual integration of the two sequences (Vroomen & Keetels, 2010).

In Study 1, the same high and low tones were used during a given experimental session. Their frequency separation differed across sessions between 6 and 9 semitones in monkey C and 4 and 4.5 semitones in monkey W. In Study 2, four pairs of high and low tones (all centered on the same mean logarithmic frequency) were used during a given experimental session. The tone pairs differed in their frequency separations: 1.5, 4.5, 7.5, or 10.5 semitones. This enabled us to

compare the neuronal responses of a given site to perceptually ambiguous sequences with different ratios of 1-stream and 2-stream percepts. In both studies, the mean frequency varied from 400 Hz to 15 kHz across experimental sessions. The mean frequency was adjusted to the frequency sensitivity of the neurons simultaneously recorded from different electrodes in an experimental session so that the tones used during the behavioral task fell inside the sensitivity range and excited as many neurons as possible.

## 2.4 | Behavioral procedure

A trial started by turning on the LED for a period of up to 5 s. During this period the monkeys were required to make contact with the touch bar and hold it for a variable and unpredictable time. 1.34–5.23 s after bar touch, the LED started flashing periodically for 2.52–5.88 s. The monkeys' task was to release the touch bar not later than 1,000 ms after the flashing had been stopped, that is, after the first omitted flash (indicated by the last vertical dotted line in Figure 1). If the monkeys did so, the reaction time was recorded, the trial was scored correct and a water reward was administered. The monkeys could consume the reward during the following 5 s before the next trial started. Bar releases before or after the 1,000-ms reaction time window were scored incorrect responses. In 75%–90% of the trials, tones were also presented in synchrony with the flashes. In most of the audiovisual trials, the tone presentation ended before the flashing. In the AV2 condition, there was one omitted tone before the first omitted flash. In the AV3 condition, there were two omitted tones before the first omitted flash. In ~5% of the audiovisual trials, the tones ended after the flashes; these catch trials served to prevent the monkeys to respond to the termination of the tone sequence.

In Study 1, trials with visual stimuli only and trials with audiovisual stimuli were presented in a block-wise manner, each consisting of 20–50 trials. Within the block of AV2 stimuli, all flashes were synchronized either with the high or the low tone. The order of the blocks was counterbalanced across the sessions. In Study 2, the type of stimulation varied randomly from trials to trial, alongside with the frequency separation of the high and the low tone. This random design was used to mitigate differences between blocks with different types of stimulation which resulted from changes in neuronal excitability across conditions.

## 2.5 | Data analysis

For each multiunit recording and for each stimulus condition we computed a peri stimulus time histogram (PSTH) from the neuronal firing with a bin size of 20 ms, which was triggered on the offset of the initial LED on-period. Only correct trials were considered for data analyses.

To find out whether the firing of a multiunit was modulated by the periodic stimulation with flashes or with tones, we tested for the presence of temporal firing patterns that were synchronized to the stimulation periodicity. To this end, we computed the amplitude spectrum from the PSTH of specific time windows of interest (e.g., during the visual condition). With the aid of a bootstrap analysis, we assessed whether the amplitude of a target frequency was significantly ( $p < 0.001$ ) greater than the corresponding value of the amplitude spectrum of surrogate PSTHs. For the latter, we generated 10,000 simulated spike trains that had the same mean firing rate as the original spike train but with randomly distributed interspike intervals (as in Brosch et al., 2015). In the visual condition, the target frequencies were 2.78 and 4.17 Hz in monkey W and monkey E (corresponding to flash periods of 360 and 240 ms), respectively, and 2.38 and 3.57 Hz in monkey C (corresponding to flash periods of 420 and 280 ms). In the auditory and audiovisual conditions, the target frequencies were 4.17 and 8.33 Hz in monkey W and monkey E (corresponding to the 240-ms interval between the onsets of two tones with the same frequency and the 120-ms interval between the onsets of two tones with the different frequency). In monkey C, the temporal intervals of the tones corresponded to the target frequencies of 3.57 and 7.14 Hz.

The PSTH was also used to quantify the response of a multiunit to selected tones by computing the spike rate in the time interval from 0 to 60 ms after tone onset. This spike rate was normalized by dividing it through the spike rate during the visual condition.

Following the approach of Fishman et al. (2004), we examined the relationship between neuronal responses and auditory streaming by calculating the ratio of the responses to the high and low tones for specific stimulus conditions. Since none of the tone frequencies was at the best frequency of the neurons under investigation we first determined which of the two tones yielded the stronger response and then divided it by the weaker response. For the auditory condition, the initial four tones in the sequence were omitted because these responses were generally larger than those to the later tones. The response ratios of a given multiunit obtained in different stimulus conditions were compared using paired *t*-tests.

### 3 | RESULTS

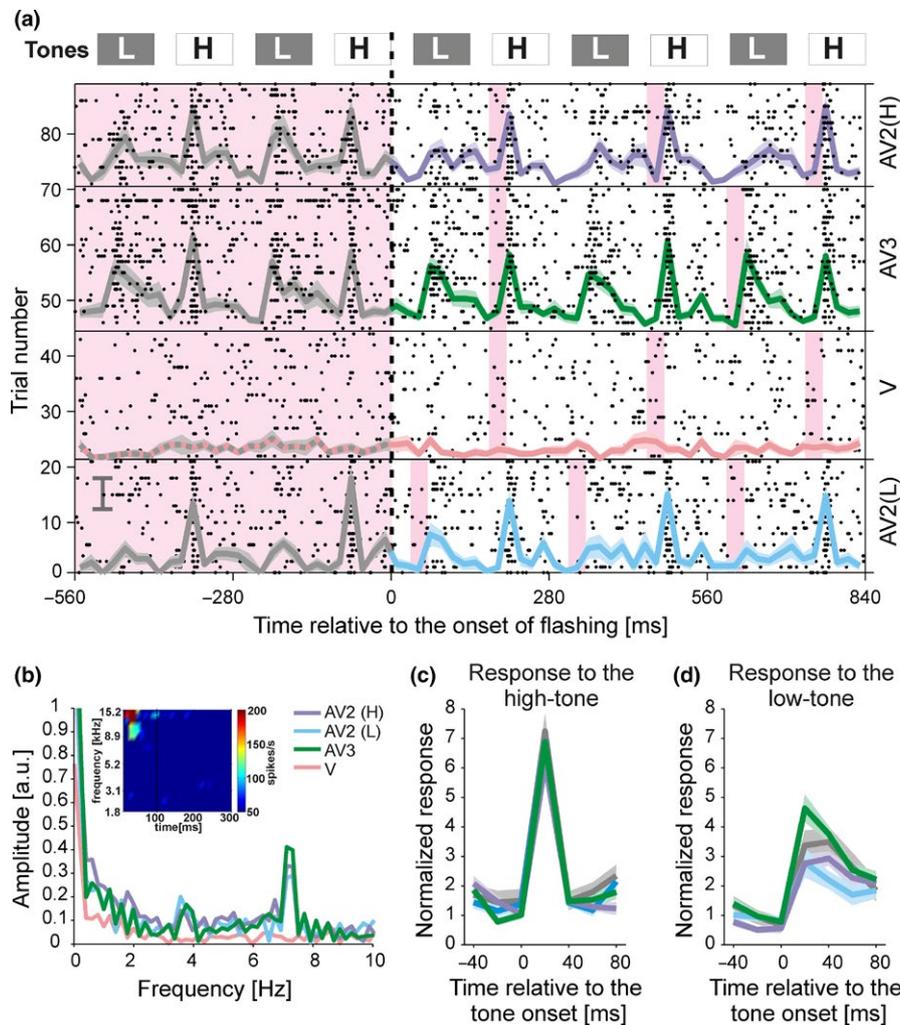
The present report is based on multiunit activity recorded from 598 sites in the left primary auditory cortex of three monkeys. In Study 1, 46 of 181 multiunits (30 from monkey C and 16 from monkey W) were used for detailed analyses because their firing was sufficiently stable during individual experimental sessions and because the monkeys performed a sufficient number of trials in these experimental sessions that allowed us to compare the stimulus conditions. All 46

multiunits were recorded during 19 experimental sessions. In Study 2, detailed analyses were performed on 158 out of 417 multiunits (73 from monkey E and 85 from monkey W), which were recorded during 68 experimental sessions. In all experiments, the monkeys were presented with repeating light flashes and were required to release a touch bar after the flashing had stopped (Figure 1). The flashes were presented alone, or they were synchronized with tones. In the latter, they were synchronized either with every second or with every third tone. Both studies tested the hypothesis that, compared to the stimulation with auditory stimuli alone, response ratios are smaller for AV2 stimuli and larger for AV3 stimuli. Study 1 and Study 2 differed in how the type of stimulation varied across trials and in the number of frequency separations that were tested during an experimental session.

#### 3.1 | Study 1

Monkey C performed a total of 1,535 trials in 10 experimental sessions during which 30 multiunits were recorded. He responded correctly to the first omitted flash in >94.4%, with little difference between the stimulus conditions. Monkey W performed a total of 2,376 trials in the nine experimental sessions considered here. He responded correctly in >94.0% of the trials, also with little difference between the different types of stimulation. Despite the very high scores, both monkeys clearly attended to the auditory stimuli and noticed the termination of the tone sequence because their reaction times differed between trials with different types of stimulation. Reaction times were longest in trials in which only visual stimuli were presented, of medium duration when AV2 stimuli were presented and of shortest duration when AV3 stimuli were presented. Pairwise *t*-tests (each  $p < 0.001$ ) revealed that in monkey C, average reaction times decreased significantly from 630 ( $\pm 15$ , *SE*) ms in trials with visual stimuli alone, to 528 ( $\pm 11$ ) ms in trials with AV2 stimuli, and to 475 ( $\pm 10$ ) ms in trials with AV3 stimuli. In monkey W, average reaction times were 392 ( $\pm 10$ ) ms for trials with visual stimuli alone, 370 ( $\pm 12$ ) ms in trials with AV2 stimuli, and 335 ( $\pm 12$ ) ms in trials with AV3 stimuli.

Figure 2 shows the neuronal firing of an example multiunit from the auditory cortex while monkey C was presented with different types of stimuli. Initially, the monkey performed a block of 21 trials with AV2 stimuli in which the flashes were synchronized with the low-frequency tones only. This block was followed by three blocks with the other stimuli, that is, with visual stimuli only, with AV3 stimuli, and with other AV2 stimuli. The latter block differed from the earlier AV2 block in that the flashes were synchronized with the high-frequency tones only. Inspection of the dot rastergram, together with the four overlaid peri stimulus time histograms (PSTHs), indicates that the firing was modulated by the stimulation only when tones were presented but not when



**FIGURE 2** Neuronal firing of an example multiunit cluster recorded from the auditory cortex of monkey C in different stimulus conditions. (a) Dot rastergrams showing the firing patterns during a 1,300-ms period relative to the onset of flashing. In the rastergram the dots in each line represent the occurrence of a spike relative to the onset of flashing in a given trial. The trials are chronologically ordered and organized into four blocks which differ by the rate of flashing and the presence of auditory stimuli. The first block shows the firing observed during 21 trials in which AV2 stimuli were presented and the low-frequency tones (8,239 Hz) were synchronized with flashes. Overlaid on the dot rastergram is a peri stimulus time histogram (PSTH, scale bar corresponds 50 spikes/s), which shows the variation of firing rate averaged across the 21 trials. The shaded curves reflect the standard errors of the mean for each of the PSTHs. The following block shows 24 trials in which only visual stimuli were presented. The third block shows 26 trials in which AV3 stimuli were presented. The last block shows 18 trials in which also AV2 stimuli were presented but in which the high-frequency tones (13,723 Hz) were synchronized with flashes. Note that the periodic increases in firing rate are synchronized with auditory stimuli. (b) Amplitude spectrum of the PSTHs, indicating that the firing of the multiunit was significantly synchronized with the periodic tone stimulation in the AV3 (green) and the two AV2 (light blue and pale purple) conditions, but not in the visual condition (pale red). The inset shows the frequency response area of the multiunit (see Brosch, Schulz, & Scheich, 1998 for details). (c,d) Normalized responses of the multiunit to high- and low-frequency tones in different AV conditions

only flashes were presented (Panel A). This was quantitatively confirmed by computing the amplitude spectrum from the PSTHs (Figure 2b). A bootstrap analysis (see Methods) revealed that during auditory stimulation, but not during visual stimulation, the firing was significantly modulated at a rate of 7.14 Hz or of 3.57 Hz, that is, at rates that are commensurate with the temporal interval of 140 ms between the high- and low-frequency tones; or with the interval of 280 ms between two tones of the same frequency. The absence of firing in auditory cortex synchronized with visual stimulation

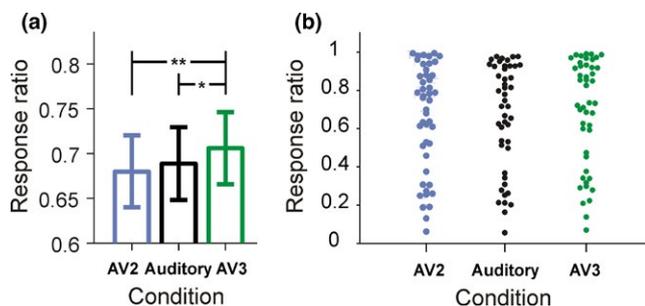
was confirmed in the other 45 multiunits of Study 1. In contrast, all 46 multiunits clearly exhibited firing synchronized with auditory stimulation.

To address the question whether the auditory streaming percept was reflected in the neuronal firing of auditory cortex, we assessed how individual multiunits responded to the high- and low-frequency tones in different stimulus conditions. Figure 2c,d show the corresponding PSTHs of the example multiunit under consideration. They were calculated by averaging across all tones presented in the different conditions,

except for the auditory condition, in which the initial four tones were not considered in order to minimize the influence of adaptation. Comparison of the different PSTHs revealed that the responses to the high-frequency tone were quite similar in the different stimulus conditions (panel C), whereas the responses to the low-frequency tone differed across the conditions (panel D). Here, relative to the auditory condition, the response to the low-frequency tone was enhanced in the AV3 condition, while it was suppressed in the two AV2 conditions. To calculate the ratio of the responses of this multiunit to the two alternating tones in the different stimulus conditions, we obtained the strength of the responses to the tones by computing the firing rate in the time intervals from 0 to 60 ms after tone onset. This revealed that, relative to the response ratio of 0.71 in the auditory condition, the response ratio was decreased to 0.66 in the two AV2 conditions but increased to 0.98 in the AV3 condition.

Similar differences in the response ratios between the stimulus conditions were seen in the other 45 multiunits of Study 1 (Figure 3). Generally, response ratios were smallest in the AV2 condition, of medium level in the auditory condition and largest in the AV3 condition. Pairwise *t*-tests revealed a statistically significant difference between the auditory and the AV3 condition ( $p = 0.0079$ ) and between the two AV conditions ( $p = 0.0402$ ). Similar results were obtained when response ratios were calculated in individual monkeys.

We could exclude that the differences in response ratios in the auditory and audiovisual conditions resulted from different responses to tones being synchronized with the light flashes and tones not being synchronized with the light flashes. This was seen both in the AV3 condition and in the AV2 condition (*t*-tests, all  $p > 0.6913$ , results not shown). Similar results were found for later time intervals.



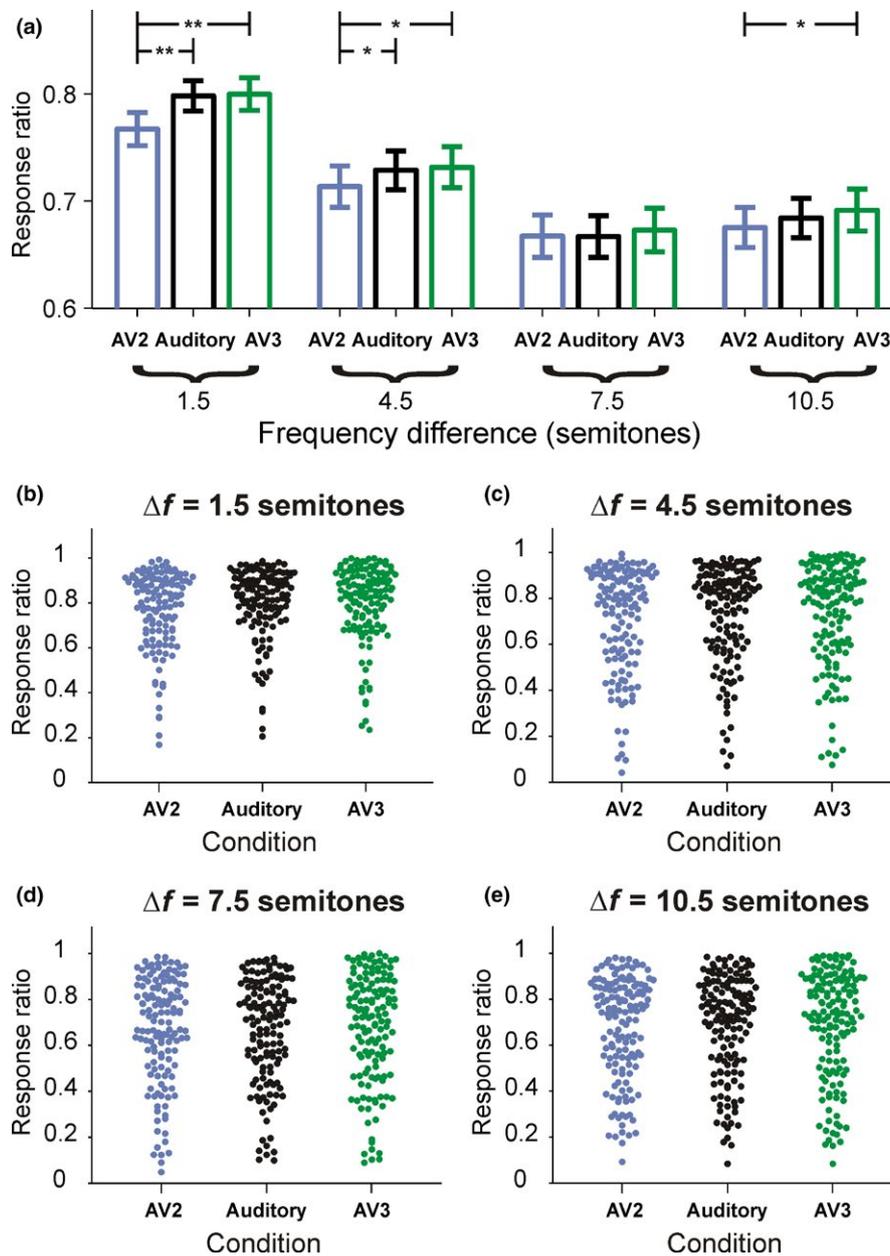
**FIGURE 3** Ratios of tone responses in auditory cortex differ between the three stimulus conditions used in Study 1. (a) Mean response ratios of 46 multiunits in the auditory (black), AV2 (blue) and AV3 (green) condition. The whiskers represent standard error of the mean. Asterisks indicate results of *t*-tests (\*\*: $p < 0.01$ , \*: $p < 0.05$ ). (b) Beeswarm plots showing the distributions of the response ratios for each of the three conditions

### 3.2 | Study 2

Monkey E performed a total of 11,138 trials in the 27 experimental sessions during which the firing of 79 multiunits available for Study 2 were recorded. He responded correctly in >94.4% of the trials for each of the three types of stimulation and each of the four frequency separations. Monkey W performed a total of 7,295 trials in the 41 experimental sessions considered here. He responded correctly in >90.0% of the trials for each of the three types of stimulation and each of the four frequency separations. As in Study 1, both monkeys had reaction times that were significantly shorter in trials in which audiovisual stimuli were presented than in trials in which only visual stimuli were presented (pairwise *t*-tests, each  $p < 0.0055$ ). In monkey E, average reaction times were 551 ( $\pm 6$ ) ms for trials with visual stimuli only, 533 ( $\pm 7$ ) ms for trials with AV2 stimuli, and 483 ( $\pm 10$ ) ms for trials with AV3 stimuli. In monkey W, average reaction times were 456 ( $\pm 11$ ) ms for trials with visual stimuli only, 390 ( $\pm 9$ ) ms for trials with AV2 stimuli, and 290 ( $\pm 8$ ) ms for trials with AV3 stimuli.

Consistent with Study 1, all 158 multiunits in auditory cortex used in Study 2 exhibited neuronal firing that was significantly modulated at rates that were commensurate with the temporal intervals between the tones. This was seen both during the auditory and the audiovisual conditions. In contrast, no multiunit exhibited firing synchronized with visual stimulation. We also noted that in all audiovisual conditions, there was no difference between the responses to tones that were synchronized with the light flashes and those that were not synchronized with the light flashes (*t*-tests, all  $p > 0.4437$ , results not shown). Extending the results of Study 1 and replicating those of previous reports (e.g., Fishman et al., 2001, 2004) we found that response ratio depended on the frequency separations of the tones used in the auditory condition. Generally, response ratio was greater when the frequency separation of the stimuli was smaller (Figure 4A, black bars).

Response ratios varied between the auditory and audiovisual conditions (Figure 4). As in Study 1, this was found for tone sequences that are perceptually ambiguous, that is, for the intermediate frequency separations of 4.5 and 7.5 semitones (even though effects were significant only for the smaller frequency separation [ $p = 0.0294$  for the AV2 condition vs. the AV3 condition and  $p = 0.0477$  for the AV2 condition vs. the auditory condition]). Significant effects of the type of stimulation were also found for tone sequences with little perceptual ambiguity. Effects were largest for sequences with the small frequency separation of 1.5 semitones (which are mostly perceived as a single auditory stream [ $p < 0.0001$  for the AV2 condition vs. the AV3 condition and  $p = 0.0002$  for the AV2 condition vs. the auditory condition]). Here the AV2 condition, which promotes stream segregation, resulted



**FIGURE 4** Ratios of tone responses in auditory cortex for different stimulus conditions and for different frequency separations obtained in Study 2. (a) Mean response ratios of 158 multiunits in the auditory, AV2 and AV3 condition, calculated for four frequency separations. Conventions as in Figure 3. (b–e) Beeswarm plots showing the distributions of the response ratios in the different conditions

in the largest change of response ratio. Effects were weaker but still significant for tone sequences with the large frequency separation of 10.5 semitones ( $p = 0.0346$ ). This tone sequence is mostly perceived as two auditory streams.

## 4 | DISCUSSION

This study finds that spiking activity in the primary auditory cortex of awake behaving monkeys is related to auditory streaming. The ratio of spike responses to repeating tone pairs was increased and thus responses were more similar in

the AV3 condition, which promoted stream integration. This ratio was decreased and thus responses were more dissimilar in the AV2 condition, which promoted stream segregation. The current results thus support and extend previous results which have suggested a relationship between neuronal activity in primary auditory cortex and the perceptual organization of tone sequences (e.g., Fishman et al., 2001). Most importantly, our experimental approach improves over the previous approaches because response ratios were observed in monkeys that attended to the auditory stimuli and were obtained with the same auditory stimuli and not, as in the previous studies, with different auditory stimuli.

Here we modified auditory streaming by synchronizing tone sequences in different ways with flash sequences (Selezneva et al., 2012). Stream segregation was promoted in the AV2 condition, in which every second tone was synchronized with a flash. Stream integration was promoted in the AV3 condition, in which every third tone was synchronized with a flash. Thus, our approach potentially has a confound similar to that of all previous studies on neurophysiological correlates of auditory streaming in mammals, in which different auditory percepts were induced with different (auditory) stimulations (Fishman et al., 2001; Kanwal, Medvedev, & Micheyl, 2003; Micheyl, Tian, Carlyon, & Rauschecker, 2005; Scholes, Palmer, & Sumner, 2015). Indeed, in contrast to the traditional view of primary auditory cortex as a unisensory brain structure (Brosch & Scheich, 2005), there is converging evidence in different species that some auditory cortical neurons respond to visual stimuli (Bizley, Nodal, Bajo, Nelken, & King, 2007; Brosch et al., 2005, 2015; Li et al., 2014), or that their auditory responses are modulated by concurrent visual stimuli (Bizley et al., 2007; Falchier, Clavagnier, Barone, & Kennedy, 2002; Kayser, Petkov, & Logothetis, 2008; Schroeder & Foxe, 2002; Schroeder & Lakatos, 2009). Despite these observations we can rule out that the variations in response ratio observed here resulted merely from comparing neuronal responses to different sensory stimuli. There was not even a single site in our sample of 215 sites in which the flashes modulated firing rates, neither in the visual condition nor in the different AV conditions. Visual modulations of tone-evoked responses were absent during the initial 60 ms after tone onset. Consistent with the slower processing speed of the visual system previous studies also did not report visual response modulations in this time window (Bizley et al., 2007; Kayser et al., 2008). Visual modulations were also absent in time windows later than 60 ms after tone onset, which corresponded to the longest tone onset interval tested here. Therefore, we conclude that individual flashes did not directly affect tone-evoked responses and, consequently, that response ratios did not change because responses to auditory stimuli were compared to response ratios to audiovisual stimuli. Instead we consider that in our experiments the visual stimulation as a whole had some “higher”-order top-down effect on the grouping of the tone sequence, and which is related to the match of the perceptual organizations of the tone and flash sequences. Similar “higher”-order interactions between auditory and visual stimuli that are largely independent of the exact temporal relationship of individual stimuli have also been reported for pairs of audiovisual stimuli (Apthorp, Alais, & Boenke, 2013; Rahne, Böckmann, von Specht, & Sussman, 2007; Shams, Kamitani, & Shimojo, 2002). It is also possible that the temporal coherence of the neuronal responses in the auditory and visual system contributes to

the differential effects of the two AV conditions on auditory streaming (Elhilali, Ma, Micheyl, Oxenham, & Shamma, 2009).

Evidence that neuronal activity in the auditory cortex and subcortical parts of the auditory system is related to auditory streaming has also been obtained in studies on humans with different types of brain signals, different behavioral paradigms, different ways of assessing and modifying auditory perception, and different tone sequences (Deike, Gaschler-Markefski, Brechmann, & Scheich, 2004; Deike, Scheich, & Brechmann, 2010; Gutschalk et al., 2005; Snyder, Alain, & Picton, 2006; Hill, Bishop, Yadav, & Miller, 2011; Yamagishi, Otsuka, Furukawa, & Kashino, 2016; but see Cusack, 2005). These studies converge in a way that auditory cortex is more strongly activated, or that neural responses to specific tones are enhanced, in trials in which subjects hear two streams as compared to trials in which subjects hear a single auditory stream. Even though not explicitly addressed, these findings suggest that the ratio of responses to the alternating tones is higher when subjects hear a single auditory stream than when subjects hear two streams. This would be consistent with the present findings that the ratio of responses to alternating tones is related to auditory streaming (also see Fishman et al., 2001). We note that a relationship between auditory streaming and neuronal activity was also found in the forebrain of birds (Itatani & Klump, 2014). However, although an objective task was used to assess auditory streaming, no neuronal correlate of the actual percept and the ensuing behavioral decision was present.

Although statistically significant relationships between neuronal responses in auditory cortex and auditory perception were obtained in two separate studies and in three monkeys the effect sizes were relatively small. Response ratios differed by maximally 4% when the AV2 condition was compared to the AV3 condition, and the differences were not statistically significant for all comparisons. We also noted in our experiments that the variations of response ratio that were related to frequency differences were larger than those related to differences in perception.

The small effect sizes that were related to auditory streaming could be a result of methodological limitations of this study. Here monkeys performed a visual task in which the auditory stimuli “only” facilitated their performance but did not trigger their behavioral responses (Selezneva et al., 2012). Thus the auditory modality was at most moderately involved in the monkeys’ behavioral decisions, which might obscure the degree to which activity in primary auditory cortex is related to perception. Effects may also be obscured by the choice of the auditory stimuli. Previous studies set one of the repeating tones to the best frequency of the auditory cortex neurons under consideration and then varied the frequency of the other tone to determine response ratios (Fishman, Kim, & Steinschneider, 2017; Fishman et al., 2001; Kanwal et al.,

2003; Scholes et al., 2015). In contrast, in most neurons of our study, the tones were not at the best frequency; mostly one of the two tones was above and the other was below the best frequency and both tones had a similar separation from the best frequency. Therefore, in this study response ratios were generally closer to unity than those in, for example, Fishman's study (Fishman et al., 2001), such that the different stimuli, particularly the AV3 stimuli, were less potent in changing response ratios than if response ratios had been tested with one tone set to the best frequency. We did not test whether analysis of the temporal coherence of the neuronal responses would reveal different effects of the two AV condition on auditory streaming (Elhilali et al., 2009). For technical reasons local field potentials were not available in this study. This may be considered a major limitation of the study because we could not address the question, for example, how the visual stimuli influence the phase of slow oscillations in A1, how slow oscillations are related to multisensory integrations, and how slow oscillations correspond to momentary fluctuations of neuronal excitability (Schroeder & Lakatos, 2009).

It is therefore likely that the responses to the alternating tones in primary auditory cortex are only partially related to auditory streaming and that auditory streaming is more related to activity in brain regions downstream of primary auditory cortex. This view would be consistent with earlier findings from our (Brosch et al., 2015; Selezneva, Scheich, & Brosch, 2006) and other groups (Beitel, Schreiner, Cheung, Wang, & Merzenich, 2003; David, Fritz, & Shamma, 2012; Otazu, Tai, Yang, & Zador, 2009; Weinberger, 2005) that sound-evoked activity in primary auditory cortex is, at most, modestly related to cognitive, procedural, and motivational aspects of auditory tasks (Tsunada & Cohen, 2014). It is therefore possible that sound-evoked activity in primary auditory cortex reflects the strength of the binding of the sequential tones rather than their perceptual category (one stream or two streams). These findings would also be consistent with the idea of Sussman, Bregman, and Lee (2014) that stimulus-driven processes of grouping do not fully create one organization or the other in perception, but pass "suggestions" about possible groupings to higher processes, with the strength proportional to the weight of evidence favoring each grouping. These considerations do not exclude that later or sustained activity in primary auditory cortex is more strongly related to the animals' choices and other aspects of auditory tasks (e.g., Brosch et al., 2005; Brosch, Selezneva, & Scheich, 2011; Selezneva, Oshurkova, Scheich, & Brosch, 2017).

Our results are broadly consistent with studies which have addressed the related question how activity in early visual cortex is related to the perception of visual stimuli (Sterzer, Kleinschmidt, & Rees, 2009). Several electrophysiological studies concluded that neural activity in the primary visual cortex had little or no relationship to how animals perceive visual stimuli, including ambiguous

stimuli that can result in binocular rivalry. In contrast, measurements of BOLD signals and magnetic fields in humans have repeatedly demonstrated a close linkage between signals in early, including primary visual cortex and the evolving spatiotemporal dynamics of rivalry perception. Moreover, transcranial magnetic stimulation over early visual cortex can induce perceptual alternations of rivaling stimuli. A reason for the discrepancies between different types of neural signals may be that neuronal firing reflects the results of local processing and thus the output of a cortical module (Maier et al., 2008). Field potentials, in turn, reflect synaptic and thus more the properties of neuronal inputs to a cortical module.

In contrast to early visual cortex, experiments in humans and monkeys revealed strong correlations between visual perception and neuronal activity in functionally specialized extrastriate cortex as well as in parietal and frontal regions. Therefore, it was proposed that frontoparietal activations reflect top-down processes that initiate a reorganization of activity in early visual cortex during perceptual reversals of ambiguous stimuli. Alternatively, and in accordance with the traditional view of multistability as a result of neural activity fluctuations in visual cortex, frontoparietal activations could merely reflect the feed-forward communication of salient neural events from visual cortex to higher order areas, similar to external stimulus changes.

This view is consistent with the conclusion of Kashino and colleagues (Yamagishi et al., 2016) that the formation and selection of auditory streams involves a widely distributed network of cortical and subcortical parts of the auditory system, including the thalamo-cortical loop. In part, this conclusion was based on their previous finding that the activity of the medial geniculate body occurred earlier than that of the auditory cortex during switching from nondominant to dominant percepts, while the activity of the auditory cortex occurred earlier than that of the medial geniculate body during switching from dominant to non-dominant percepts (Kondo & Kashino, 2009). Further experiments are required to clarify how the firing of single neurons in different parts of the auditory system is related to auditory streaming.

## ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB TR 31, He 1721/10-1, He 1721/10-2). We thank Prof. H. Scheich for valuable discussions, Vighneshvel Thirupathi for language editing, and Cornelia Bucks for help during experiments.

## CONFLICT OF INTEREST

We have no competing interests.

## DATA ACCESSIBILITY

All data generated or analyzed during this study are included in this published article. For additional information, please contact the corresponding author.

## AUTHORS' CONTRIBUTIONS

ES conceived of the study, designed the study, carried out the experiments and data analysis, interpreted the results, and drafted the manuscript; MB conceived of the study, designed the study, interpreted the results, and drafted the manuscript; AG carried out the experiments and participated in the interpretation of the results; EB interpreted the results and drafted the manuscript. All authors gave final approval for publication.

## ORCID

Elena Selezneva  <http://orcid.org/0000-0002-7495-5278>

## REFERENCES

- Apthorp, D., Alais, D., & Boenke, L. T. (2013). Flash illusions induced by visual, auditory, and audiovisual stimuli. *Journal of Vision, 13*, 1–15.
- Beitel, R., Schreiner, C., Cheung, S., Wang, X., & Merzenich, M. (2003). Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 11070–11075. <https://doi.org/10.1073/pnas.1334187100>
- Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex, 17*, 2172–2189. <https://doi.org/10.1093/cercor/bhl128>
- Bregman, A. S. (1990). *Auditory scene analysis. In The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brosch, M., & Scheich, H. (2005). Non-acoustic influence on neural activity in auditory cortex. In R. König, P. Heil, E. Budinger, & H. Scheich (Eds.), *Auditory cortex: Towards a synthesis of human and animal research* (pp. 127–143). Mahwah, NJ: Lawrence Erlbaum Associates.
- Brosch, M., & Scheich, H. (2008). Tone-sequence analysis in the auditory cortex of awake macaque monkeys. *Experimental Brain Research, 184*, 349–361. <https://doi.org/10.1007/s00221-007-1109-7>
- Brosch, M., Schulz, A., & Scheich, H. (1998). Neuronal mechanisms of auditory backward recognition masking in macaque auditory cortex. *NeuroReport, 9*, 2551–2555. <https://doi.org/10.1097/00001756-199808030-00023>
- Brosch, M., Selezneva, E., Bucks, C., & Scheich, H. (2004). Macaque monkeys discriminate pitch relationships. *Cognition, 91*, 259–272. <https://doi.org/10.1016/j.cognition.2003.09.005>
- Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *Journal of Neuroscience, 25*, 6797–6806. <https://doi.org/10.1523/JNEUROSCI.1571-05.2005>
- Brosch, M., Selezneva, E., & Scheich, H. (2011). Formation of associations in auditory cortex by slow changes of tonic firing. *Hearing Research, 271*, 66–73. <https://doi.org/10.1016/j.heares.2010.05.003>
- Brosch, M., Selezneva, E., & Scheich, H. (2015). Neuronal activity in primate auditory cortex during the performance of audiovisual tasks. *European Journal of Neuroscience, 41*, 603–614. <https://doi.org/10.1111/ejn.12841>
- Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience, 17*, 641–651. <https://doi.org/10.1162/0898929053467541>
- David, S. V., Fritz, J. B., & Shamma, S. A. (2012). Stimulus valence and task-relevance control rapid plasticity in primary auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 2144–2149. <https://doi.org/10.1073/pnas.1117717109>
- Deike, S., Gaschler-Markefski, B., Brechmann, A., & Scheich, H. (2004). Auditory stream segregation relying on timbre involves left auditory cortex. *NeuroReport, 15*, 1511–1514. <https://doi.org/10.1097/01.wnr.0000132919.12990.34>
- Deike, S., Scheich, H., & Brechmann, A. (2010). Active stream segregation specifically involves the left human auditory cortex. *Hearing Research, 265*, 30–37. <https://doi.org/10.1016/j.heares.2010.03.005>
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. A. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron, 61*, 317–329. <https://doi.org/10.1016/j.neuron.2008.12.005>
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience, 22*, 5749–5759. <https://doi.org/10.1523/JNEUROSCI.22-13-05749.2002>
- Fishman, Y. I., Arezzo, J. C., & Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America, 116*, 1656–1670. <https://doi.org/10.1121/1.1778903>
- Fishman, Y. I., Kim, M., & Steinschneider, M. (2017). A crucial test of the population separation model of auditory stream segregation in macaque primary auditory cortex. *Journal of Neuroscience, 37*, 10645–10655.
- Fishman, Y. I., Reser, D. H., Arezzo, J. C., & Steinschneider, M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hearing Research, 151*, 167–187. [https://doi.org/10.1016/S0378-5955\(00\)00224-0](https://doi.org/10.1016/S0378-5955(00)00224-0)
- Gutschalk, A., Micheyl, C., Melcher, J. R., Rupp, A., Scherg, M., & Oxenham, A. J. (2005). Neuromagnetic correlates of streaming in human auditory cortex. *Journal of Neuroscience, 25*, 5382–5388. <https://doi.org/10.1523/JNEUROSCI.0347-05.2005>
- Hill, K. T., Bishop, C. W., Yadav, D., & Miller, L. M. (2011). Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. *BMC Neuroscience, 12*, 85. <https://doi.org/10.1186/1471-2202-12-85>
- Itatani, N., & Klump, G. M. (2014). Neural correlates of auditory streaming in an objective behavioral task. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 10738–10743. <https://doi.org/10.1073/pnas.1321487111>
- Kanwal, J. S., Medvedev, A. V., & Micheyl, C. (2003). Neurodynamics for auditory stream segregation: Tracking sounds in the mustached bat's natural environment. *Network, 14*, 413–435. [https://doi.org/10.1088/0954-898X\\_14\\_3\\_303](https://doi.org/10.1088/0954-898X_14_3_303)

- Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, *18*, 1560–1574. <https://doi.org/10.1093/cercor/bhm187>
- Kondo, H. M., & Kashino, M. (2009). Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. *Journal of Neuroscience*, *29*, 12695–12701. <https://doi.org/10.1523/JNEUROSCI.1549-09.2009>
- Li, X., Yu, K., Zhang, Z., Sun, W., Yang, Z., Feng, J., ... He, J. (2014). Cholecystokinin from the entorhinal cortex enables neural plasticity in the auditory cortex. *Cell Research*, *24*, 307–330. <https://doi.org/10.1038/cr.2013.164>
- Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F. Q., & Leopold, D. A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nature Neuroscience*, *11*, 1193–1200. <https://doi.org/10.1038/nn.2173>
- Micheyl, C., Tian, B., Carlyon, B. P., & Rauschecker, J. P. (2005). Perceptual organization of sound sequences in the auditory cortex of awake macaques. *Neuron*, *48*, 139–148. <https://doi.org/10.1016/j.neuron.2005.08.039>
- Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United with Acustica: The Journal of the European Acoustics Association*, *88*, 320–333.
- Otazu, G. H., Tai, L. H., Yang, Y., & Zador, A. M. (2009). Engaging in an auditory task suppresses responses in auditory cortex. *Nature Neuroscience*, *12*, 646–654. <https://doi.org/10.1038/nn.2306>
- Rahne, T., Böckmann, M., von Specht, H., & Sussman, E. S. (2007). Visual cues can modulate integration and segregation of objects in auditory scene analysis. *Brain Research*, *1144*, 127–135. <https://doi.org/10.1016/j.brainres.2007.01.074>
- Rahne, T., Deike, S., Selezneva, E., Brosch, M., König, R., Scheich, H., ... Brechmann, A. (2008). A multilevel and cross-modal approach towards neuronal mechanisms of auditory streaming. *Brain Research*, *1220*, 118–131. <https://doi.org/10.1016/j.brainres.2007.08.011>
- Scholes, C., Palmer, A. R., & Sumner, C. J. (2015). Stream segregation in the anesthetized auditory cortex. *Hearing Research*, *328*, 48–58. <https://doi.org/10.1016/j.heares.2015.07.004>
- Schroeder, C. E., & Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Research. Cognitive Brain Research*, *14*, 187–198. [https://doi.org/10.1016/S0926-6410\(02\)00073-3](https://doi.org/10.1016/S0926-6410(02)00073-3)
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*, 9–18. <https://doi.org/10.1016/j.tins.2008.09.012>
- Selezneva, E., Gorkin, A., Mylius, J., Noesselt, T., Scheich, H., & Brosch, M. (2012). Reaction times reflect subjective auditory perception of tone sequences in macaque monkeys. *Hearing Research*, *294*, 133–142. <https://doi.org/10.1016/j.heares.2012.08.014>
- Selezneva, E., Oshurkova, E., Scheich, H., & Brosch, M. (2017). Category-specific neuronal activity in left and right auditory cortex and in medial geniculate body of monkeys. *PLoS ONE*, *12*, e0186556. <https://doi.org/10.1371/journal.pone.0186556>
- Selezneva, E., Scheich, H., & Brosch, M. (2006). Dual time scales for categorical decision making in auditory cortex. *Current Biology*, *16*, 2428–2433. <https://doi.org/10.1016/j.cub.2006.10.027>
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Brain Research. Cognitive Brain Research*, *14*, 147–152. [https://doi.org/10.1016/S0926-6410\(02\)00069-1](https://doi.org/10.1016/S0926-6410(02)00069-1)
- Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*, 1–13. <https://doi.org/10.1162/089892906775250021>
- Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, *13*, 310–318. <https://doi.org/10.1016/j.tics.2009.04.006>
- Sussman, E. S., Bregman, A. S., & Lee, W. W. (2014). Effects of task-switching on neural representations of ambiguous sound input. *Neuropsychologia*, *64*, 218–229. <https://doi.org/10.1016/j.neuropsychologia.2014.09.039>
- Tsunada, J., & Cohen, Y. E. (2014). Neural mechanisms of auditory categorization: From across brain areas to within local microcircuits. *Frontiers in Neuroscience*, *8*, 161.
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception and Psychophysics*, *72*, 871–884. <https://doi.org/10.3758/APP.72.4.871>
- Weinberger, N. (2005). Reconceptualizing the primary auditory cortex: Learning, memory and specific plasticity. In R. König, P. Heil, E. Budinger, & H. Scheich (Eds.), *Auditory cortex: Towards a synthesis of human and animal research* (pp. 465–491). Mahwah, NJ: Lawrence Erlbaum Associates.
- Yamagishi, S., Otsuka, S., Furukawa, S., & Kashino, M. (2016). Subcortical correlates of auditory perceptual organization in humans. *Hearing Research*, *339*, 104–111. <https://doi.org/10.1016/j.heares.2016.06.016>

**How to cite this article:** Selezneva E, Gorkin A, Budinger E, Brosch M. Neuronal correlates of auditory streaming in the auditory cortex of behaving monkeys. *Eur J Neurosci*. 2018;00:1–12. <https://doi.org/10.1111/ejn.14098>