

28 Primary Auditory Cortical Responses while Attending to Different Streams

PINGBO YIN¹, LING MA², MOUNYA ELHILALI^{1,3}, JONATHAN FRITZ¹,
AND SHIHAB SHAMMA^{1,2,3}

1 Introduction

Auditory streaming is a fundamental perceptual component of auditory scene analysis. It manifests itself in the everyday ability of humans and animals to parse complex acoustic information arising from multiple sound sources into meaningful auditory “streams”. While seemingly effortless, the neural mechanisms underlying auditory streaming remain a mystery largely because experiments in non-human species has been hampered by the difficulty of assessing the subjective perception of streaming without relying on introspection and language. Here we overcome this difficulty through the use of specially designed stimuli and psychoacoustic tasks to induce, manipulate, and at the same time *objectively* assess streaming in animals. While the basic structure of the stimuli and tasks are identical to those used with humans in Micheyl, Oxenham, and Shamma (this volume) they are slightly adapted so that physiological data can be collected *simultaneously* with behavior in animals.

2 Background

There has recently been growing interest in the neural basis of auditory stream segregation using animal models ranging from awake monkeys to birds (Fishman et al. 2004, Bee and Klump 2004). Using the classic ABAB... repeating two-tone paradigm (Bregman 1990), these studies have inferred that neural responses to the tones become more segregated reflecting the well-known streamed percept of the tones. For instance, Fishman et al. (2004) recorded multi-unit and local field potentials while setting the frequency of the A tone to correspond to the BF of the cortical site contacted. The frequency of the B tone was set either below or above the BF such that the response elicited by the

¹Institute for Systems Research, University of Maryland, College Park MD USA 20742

²Bioengineering Program, University of Maryland, College Park, MD USA 20742

³Electrical and Computer Engineering Department, University of Maryland, College Park, MD USA 20742, sas@eng.umd.edu

B tones was approximately half that elicited by the A tones. The results showed that when the tone repetition rate was increased from 2.5 to 20 Hz, the responses of cortical sites with a BF at the A-tone frequency displayed less and less activation to the B tones. Similarly, changing the frequency separation, presentation rate, tone duration, and other parameters elicited response changes consistent with perceptual trends observed in human listeners (Fishman et al. 2004; Bee et al. 2004). Finally, single-unit recordings of responses in the primary auditory cortex of awake-monkeys to 10-s repeating (ABA_) tone sequences, have demonstrated that responses change gradually following the onset of the sequence in a way that was consistent with the phenomenon of “build-up of stream segregation” (Micheyl et al. 2005b).

Despite these encouraging results, two criticisms can be made of all neural studies of streaming so far. First, it could be argued that the neural response patterns putatively associated with one- and two-stream percepts could merely be reflections of the change in stimulus parameters, rather than providing the neural basis for streaming. Stronger neural evidence for streaming would be obtained if one could demonstrate parallel changes in neural responses and percepts in the absence of concomitant changes in the physical stimulus. Second, no previous studies provided behavioral tests of the percepts that the animals were experiencing with the stimuli used in the experiments since all involved *non-behaving* animals. This may be particularly important for streaming-related issues, as attention seems to be involved in the formation or modulation of auditory streams (e.g., Carlyon et al. 2001, Cusack 2005; Sussman 2005). The experiments described below aimed to address these two objections by combining behavior with physiological experiments.

3 Behavioral Tasks

The perceptual phenomena investigated here are similar to those already described in Micheyl et al. (this volume). The stimuli have been adapted for animal experimentation by (1) creating tasks with objective performance criteria, similar to those already used with human subjects by Micheyl et al. (2005a), and (2) selecting parameter ranges appropriate for the ferret and the isolated cortical units. Ferrets were trained on the two behavioral tasks illustrated in Fig. 1. The two tasks share an identical initial sequence of random tone maskers (A-tones) and an embedded sequence of B-tones. This entire portion of the stimulus is referred to as *reference*. The two tasks differ in the final three bursts that the animals must attend to, which are designated as *targets*. In task 1, the *ferret* detects that the A maskers become a repeated set of tones (B-tone remains unchanged throughout). By contrast, in task 2, the target is a 1/4 octave change in the frequency of the B tone (B'). The hypothesis is that animals must now switch attention from the broadly distributed masker tones (“A”) to the spectrally narrow B-tone. With the addition of this task, we

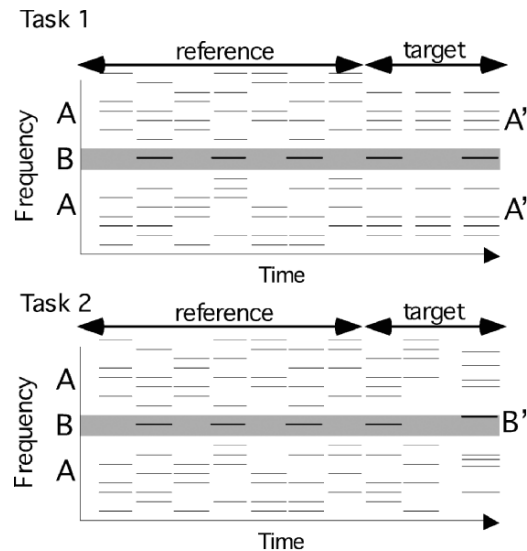


Fig. 1 Two behavioral tasks for ferrets: Task 1 – detect stationary A (A’); Task 2 – detect a change in B (B’)

can compare the neural correlates of the perceptual state of the animals in three different conditions: (1) naïve, (2) attending to the stream of B-tones, (3) attending to the broader spectrum (or the “stream” of the A-maskers).

The behavioral training paradigm was the conditioned avoidance procedure, which involved a combination of positive and negative reinforcements (Heffner and Heffner 1995; Fritz et al. 2003). Specifically, two ferrets readily learned to lick a waterspout during the *reference* signals, and then to immediately withdraw upon hearing the targets. They achieved adequate performance levels (*Discrimination Rate* > 0.65) in both tasks at various parameter settings, the two most important of which are the width of the protection zone (< 2 ERBs; shaded region in Fig. 1) and the level of B-tones relative to the A-maskers (< 20 dB). Increasing these parameters further separates the A and B streams, facilitating the detection of *targets*.

4 Physiological Responses

Responses of single-units in the primary auditory cortex were measured or contrasted while animals were in each of three conditions: *naïve*, *tasks 1 and 2*. In particular, because of the broadband and random structure of the A-tones, it was possible to estimate the tuning curves of the isolated cells with the reverse-correlation method (deCharms et al. 1998, and Fig. 2), using only the *reference* portion of the stimuli (which was identical across all conditions).

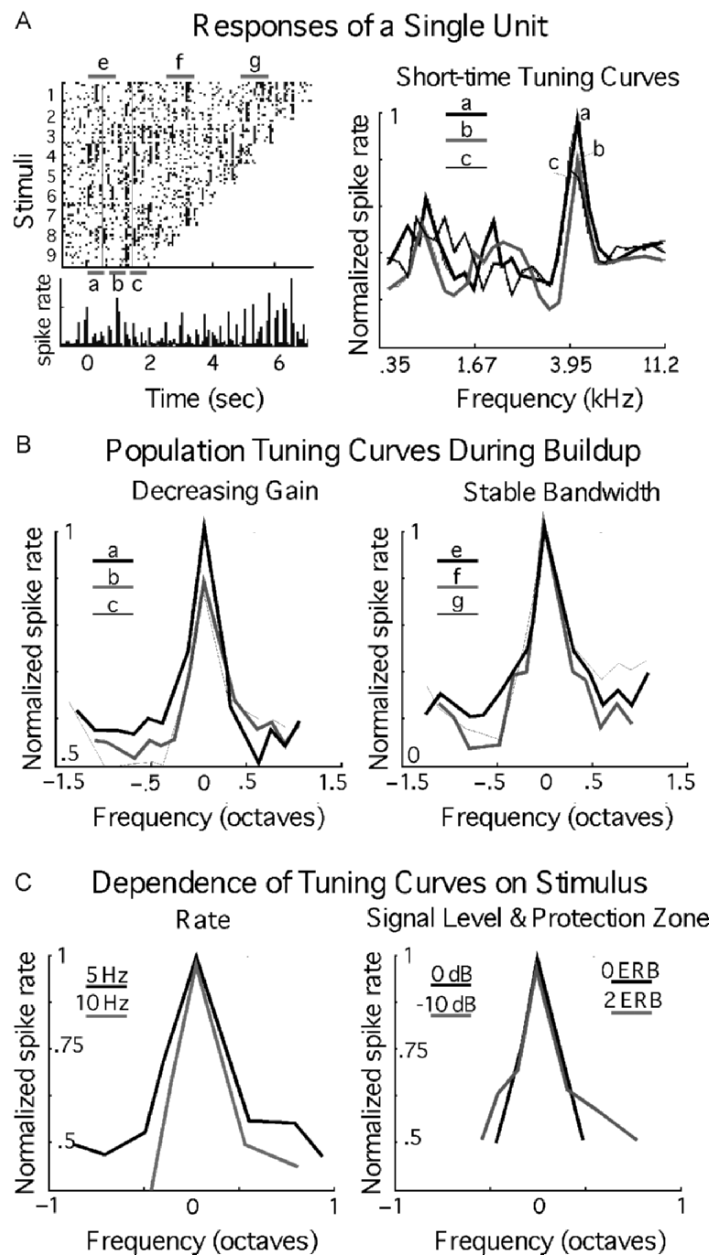


Fig. 2 Correlates of streaming in the naïve animal: A raster, post-stimulus histogram, and short-time tuning curves of a single unit responding to the tone-in-masker stimulus (see text for details). Short-time tuning curves were computed over the intervals *a-b-c* and *e-f-g* marked below and above the histogram, respectively; B the gain of the tuning curves *decreases* during the buildup, whereas its bandwidth remains *unchanged*; C tuning curve bandwidths decrease with increasing rates, but remain approximately unchanged with respect to other stimulus parameters

4.1 Responses in the Naïve Ferret

As a benchmark for the behavioral neurophysiology measurements, we recorded responses in 23 single units in AI of a naïve ferret while passively listening to the stimuli of Fig. 1. In particular, we extracted the gain and bandwidth of the tuning curves for different stimulus parameters (presentation rate, size of protection region, relative B-tone level), and as they varied during the early “build-up” period following the onset of the stimulus. Figure 2A (left panel) illustrates the raster and PST histograms of responses from a single unit to ten repetitions of nine different sequences of different lengths. Tuning curves are computed by reverse-correlating the responses with the spectrogram of the stimulus (deCharms et al. 1998) over a series of short-time intervals following the onset of the sequences (marked at bottom of raster as a-b-c in Fig. 2A; right panel). In Fig. 2B,C we illustrate the total *population* tuning curves computed by centering and then averaging the tuning curves from 23 single units.

During the buildup period, the data reveal two changes: (1) The *gain* of the tuning curves declines by about 10%, first rapidly, but then much more slowly or not at all past 1 s (Fig. 2B; left panel); (2) The *bandwidths* of the tuning curves remain essentially unchanged throughout (Fig. 2B, right panel). This suggests that decreasing responses to off-BF tones observed previously (Fishman et al. 2001) is due to a change in the tuning curve gain and *not* its bandwidth.

We also examined changes in the population tuning curves under the following stimulus conditions: (1) *Presentation rate* (5 vs 10 Hz): there was a significant decrease in the average bandwidth at faster rates of about 25% (Fig. 2C, left panel). This finding is consistent with the decrease of responses to the off-BF tone at higher rates observed in previous studies (e.g., Fishman et al. 2001), and which was attributed to “forward suppression” or “adaptation” of responses to tone frequencies remote from the BF by preceding tones at the BF. (2) *Signal-to-masker levels* (0 vs 10 dB) and *Protection zone widths* (0 vs 2 ERB): increasing the B-tone level or protection zone surrounding it facilitates its perceptual streaming (see psychoacoustics above), and hence one might expect a narrowing of the tuning curves with both. However, there were no significant changes in tuning curve bandwidths with either manipulation (Fig. 2C; right panel). There are no comparable physiological data in the literature regarding the signal-to-masker level manipulation. The protection zone width, however, is analogous to the frequency separation between the A and B tones in the classic two-tone streaming experiments. Therefore, the absence of tuning curve bandwidth changes here suggests that, decreasing responses to the off-BF tones with increasing ΔF may simply reflect the finite bandwidth of the tuning curves and not a change in tuning related to the perceptual state.

4.2 Responses in Behaving Ferrets

We explored how responses adapted when the focus of the animal's attention changed from the spectrally-broad A stream in task 1 to the spectrally-narrow B-tone stream in task 2. Specifically, we hypothesized that the tuning curves may sharpen during the onset of streaming in task 2 to reflect this shift in focus.

Responses of 15 single-units were measured from two animals while they engaged in the two tasks of Fig. 1. Recall that tuning curves were only estimated and compared based on responses during the *reference* portion of the stimuli (which was identical in both tasks). As in the naïve animal, tuning curve *gains* declined by about 10% during the initial 1–2 s (not shown). However, *bandwidth* of the tuning curves in most cells (60%) changed depending on the behavioral task (as discussed in detail below). An example from one single-unit is shown in Fig. 3. This cell was tuned at a BF of 2.7 kHz. The frequency of the B-tone in the stimulus was set about 1/2 octave below the BF (at 1.6 kHz; arrowheads in Fig. 3B), surrounded by a protection zone of 3 ERB on either side. We also measured the spectro-temporal receptive field (STRF) of this cell (left panel of Fig. 3D) *before* any behavioral tasks commenced using the methods of Fritz et al. (2003). The STRF indicates a strong excitatory region near the BF of the cell at 2.7 kHz (white area), surrounded by a weak inhibitory field (darker). The range of frequencies of the A-maskers and B-tone used in the behavioral tasks are schematically depicted on the left of the STRF by the dashed and solid lines, respectively.

Response PSTH during the two tasks are shown in Fig. 3A. The cell responded during both stimulus intervals, labeled A (*only* A-maskers present) and A+B (A-maskers+B-tone). In task 1 (left panel), the responses during the A intervals were bigger than those during the intervening A+B intervals, presumably because the B-tone coincided with the weak inhibitory fields just outside of the excitatory tuning curve of the cell. The average tuning curves *during* this task (Fig. 3B) were relatively broad, with little evidence of the inhibitory sidebands. During task 2, responses to the A+B interval diminished considerably relative to their levels in task 1 (Fig. 3A; right panel). Evidently, the reason is a suppression of the tuning curve near the B-tone causing a substantial narrowing of the excitatory tuning curve (see arrows in Fig. 3B). This emergent inhibitory sideband apparently persisted since it could also be detected in the STRF measured *after* the completion of the task (Fig. 3D; right panel). In addition to the task-dependent changes in the overall (average) tuning curves that persisted after the end of the experiments (e.g., as in Fig. 3D), there were also rapid tuning curve changes that occurred *within* each stimulus trial and may correlate with the “build up” of streaming. For instance, in Fig. 3C, the short-time average tuning curves from a population of ten cells responding during task 1, demonstrate a rapid increasing in bandwidth during the initial response period (1–2 s following stimulus onset). So it is possible that there are two kinds of superimposed tuning curve and STRF changes: (1) persistent changes that reflect the overall effect of the

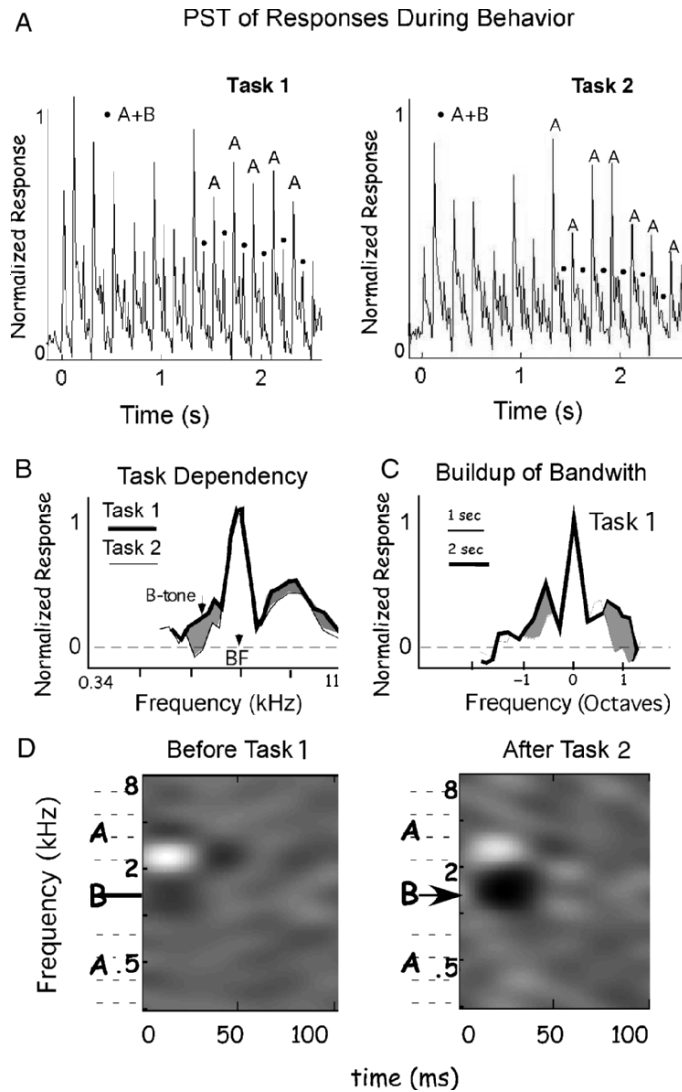


Fig. 3 Effects of behavior on responses and tuning curves: **A** PSTH of responses from a single unit during the two behavioral tasks, collected from all trials as shown in Fig. 2A. The primary difference between the *two panels* is the substantial *decrease* of responses during the A+B intervals (indicated by the “dots”) during task 2; **B** normalized tuning curves computed *only* from responses during the *reference* interval of the stimuli (Fig. 1). *Shaded region* highlights the sharpening of the curves during task 2. The *arrows* indicate the BF of the cell, and the frequency of the B-tone; **C** the rapid broadening of the tuning curves in the population of single units within 2 s following onset of the stimulus trials in task 1; **D** STRFs of the cell measured *before* and *after* the two tasks. The *dashed lines* represent the maskers. *Before* task 1, the lone excitatory region (*white*) is near 2.7 kHz. *After* task 2, an inhibitory field (*black*) emerges at the frequency of the B-tone

attentional state of the animal (e.g., attending to stream A or B throughout the block of trials enhances the representation of the attended stream relative to the other), and (2) rapid changes within a trial reflecting more automatic processes (e.g., correlates of the buildup of streaming).

5 Summary and Discussion

The results described here suggest that attentive behavior alters two aspects of the neural correlates of streaming: (1) those associated with the build-up and formation of the streams, and (2) others more global and that persist *after* the behavior. In the *first*, stable tuning curves and highly phasic responses following the onset of tone sequences in the naïve give way to narrowing bandwidths and less precipitously decreasing gains in the behaving animal. We conjecture that these changes facilitate the formation of two streams in the attentive animal. *Second*, attention to a foreground (*target*) sequence against a background (*reference*) induces *additional* tuning curve or STRF changes, with *target* frequencies enhanced and *reference* frequencies depressed. Unlike the first set of changes that start anew at the onset of each trial, these latter changes are not reset every trial, but instead may persist *after* the behavior. They are very similar to the “rapid plasticity” effects observed in STRFs of animals during and after similar *target/reference* discrimination behaviors (Fritz et al. 2005). Such plasticity, in fact, is not related to streaming per se since it occurs even when the *target/reference* stimuli are not “streams”, e.g., if they are presented at slow rates as in Fritz et al. (2005). Therefore, it is clear that “attention” induces complex intertwined changes in the response properties, with different time-courses and longevity. Perhaps, attending to a “streamed” *reference* simply amplifies the effects already dictated by the structure of the task. To clarify these factors, it will be necessary in the future to contrast the effects of *streaming* on the tuning curves in behavioral tasks where “streams” play both the role of *reference* and/or *target*.

Acknowledgments. This research is funded by NIH grants R01DC007657 and R01DC005779.

References

- Bregman A (1990) Auditory scene analysis: perceptual organization of sound. MIT Press
 Bee M, Klump G (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. *J Neurophysiol* 92:1088–1104
 Carlyon R, Cusack R, Foxton J, Robertson I (2001) Effects of attention and unilateral neglect on auditory stream segregation, *J Exp Psychol: Hum Percept Perform* 27:115

- Cusack R (2005) The intraparietal sulcus and perceptual organization. *J Cogn Neurosci* 17:641–651
- deCharms R, Blake D, Merzenich M (1998) Optimizing sound features for cortical neurons, *Science* 280:1439–1443
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear Res* 151:167–187
- Fishman Y, Arezzo J, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *J Acoust Soc Am* 116:1656–1670
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex, *Nat Neurosci* 6:1216–1223
- Fritz J, Elhilali M, Shamma S (2005) Active listening: task-dependent plasticity of receptive fields in primary auditory cortex, *Hear Res* 206:159–176
- Heffner HE, Heffner RS (1995) Conditioned avoidance. In: Klump G, Dooling R et al. (eds) *Methods in comparative psychoacoustics*, Birkhauser, Basel, pp 79–94
- Micheyl C, Tian B, Carlyon R, Rauschecker J (2005b) Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48(1):139–148
- Micheyl C, Carlyon R, Cusack R, Moore B (2005a) Performance measures of auditory organization. In: Pressnitzer D, de Cheveigné A, McAdams S, Collet L (eds) *Auditory signal processing: physiology, psychoacoustics, and models*. Springer, Berlin Heidelberg New York
- Sussman E (2005) Integration and segregation in auditory scene analysis. *JASA* 117:1285