# Neural adaptation to tone sequences in the songbird forebrain: Patterns, determinants, and relation to the build-up of auditory streaming

**Mark A. Bee[†],[*]**,
Department of Ecology, Evolution, and Behavior, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St. Paul, MN 55108, USA

**Christophe Micheyl[†]**,
Department of Psychology, University of Minnesota, Minneapolis, N218 Elliott Hall, 75 East River Road, Minneapolis, MN 55455, USA

**Andrew J. Oxenham**, and
Department of Psychology, University of Minnesota, Minneapolis, N218 Elliott Hall, 75 East River Road, Minneapolis, MN 55455, USA

**Georg M. Klump**
AG Zoophysiologie & Verhalten, Fakultät V, Institut für Biologie und Umweltwissenschaften, Carl von Ossietzky Universität Oldenburg, 26111, Oldenburg, Germany

## Abstract

Neural responses to tones in the mammalian primary auditory cortex (A1) exhibit adaptation over the course of several seconds. Important questions remain about the taxonomic distribution of multi-second adaptation and its possible roles in hearing. It has been hypothesized that neural adaptation could explain the gradual "build-up" of auditory stream segregation. We investigated the influence of several stimulus-related factors on neural adaptation in the avian homologue of mammalian A1 (field L2) in starlings (*Sturnus vulgaris*). We presented awake birds with sequences of repeated triplets of two interleaved tones (ABA–ABA–…) in which we varied the frequency separation between the A and B tones (ΔF), the stimulus onset asynchrony (SOA, time from tone onset to onset within a triplet), and tone duration (TD). We found that SOA generally had larger effects on adaptation compared with ΔF and TD over the parameter range tested. Using a simple model, we show how time-dependent changes in neural responses can be transformed into neurometric functions that make testable predictions about the dependence of the build-up of stream segregation on various spectral and temporal stimulus properties.

### Keywords

Adaptation; Auditory stream segregation; Build-up; Field L; Starling; *Sturnus vulgaris*

---

[*]mbee@umn.edu, Phone: ++1-612-624-6749, Fax: ++1-612-624-6667.
[†]The first two authors contributed equally to this study and author order was determined by flipping a coin.

## Introduction

Many ecologically relevant sounds, such as those associated with acoustic signaling and locomotion, comprise sequences of distinct acoustic elements that unfold over time. In addition, natural acoustic scenes comprise concurrent sound sequences produced by multiple sources. The auditory system must parse these overlapping sound sequences into separate ongoing percepts, or "auditory streams," that correspond to separate sources and that can be selectively attended (Bregman, 1990; Carlyon, 2004). For humans, two familiar examples of stream segregation occur when we listen to specific instruments in a musical ensemble (Hartmann and Johnson, 1991) or to a single person speaking in a crowded social gathering, such as a cocktail party (Brokx and Nooteboom, 1982). The natural behaviors of many nonhuman animals indicate that they also regularly encounter and solve problems of auditory stream segregation in their daily lives (reviewed in Bee and Micheyl, 2008; Hulse, 2002).

Psychophysical studies of stream segregation in humans have traditionally used simplified stimuli composed of repeating sequences of two alternating tones (A and B) that differ in frequency or some other acoustic property (e.g., ABA–…; Fig. 1; reviewed in Moore and Gockel, 2002). These stimulus sequences evoke two rather different percepts, depending on an interaction between the frequency separation (ΔF) between the A and B tones and the tone repetition rate (Bregman, 1990;van Noorden, 1975). When ΔF is small (e.g., 1–10%), most listeners perceive the stimulus as a single stream of "galloping" tones alternating in frequency regardless of tone repetition rate (Fig. 1a). In contrast, when ΔF is large, and provided the tone repetition rate is not too slow (e.g., > 2–3 tones/s), the percept is that of two separate streams, each containing only tones of the same frequency (Fig. 1b). Psychophysical studies of goldfish (Fay, 1998,2000), songbirds (MacDougall-Shackleton et al., 1998), and monkeys (Izumi, 2001,2002) suggest that a wide taxonomic diversity of animals also experiences auditory stream segregation when listening to simple acoustic sequences similar to those used to study streaming in humans (reviewed in Fay, 2008).

Neurophysiological studies of human subjects have identified putative neural correlates of streaming in auditory cortex (Cusack, 2005; Gutschalk et al., 2007; Kondo and Kashino, 2009; Micheyl et al., 2007; Snyder et al., 2006, [2009]; Sussman et al., 1999, 2007; Wilson et al., 2007; Winkler et al., 2005). Electrophysiological studies of songbirds and mammals have used simple sequences of two interleaved tones to elucidate the contribution to stream segregation of well-known bottom-up aspects of auditory processing (reviewed in Fay, 2008). For example, recordings of single- and multi-unit responses to tone sequences in the primary auditory cortex (A1) of monkeys (Fishman et al., 2001, 2004; Micheyl et al., 2005), and its evolutionary homologue in songbirds (field L2; Bee and Klump, 2004, 2005), have identified important roles for neural frequency selectivity and forward suppression in accounting for the above-described effects of ΔF and tone repetition rate on the segregation of pure-tone sequences. Recent single-unit recordings in ferret A1 suggest that temporal incoherence between the responses of neurons tuned to different frequencies contributes to the formation of separate streams (Elhilali et al., 2009). It is worth pointing out that some of the neural phenomena that have been identified in A1 and L2 may ultimately originate in the auditory periphery, as indicated by results from a recent study of stream segregation at the level of the cochlear nucleus in Guinea pigs (Pressnitzer et al., 2008).

An important feature of auditory streaming in humans is that the perception of two separate streams is not instantaneous, but instead is cumulative and grows over the first several seconds (e.g., 5 – 10 s) following the onset of stimulation (Anstis and Saida, 1985; Bregman, 1978; Carlyon et al., 2001; Cusack et al., 2004; Micheyl et al., 2005). This

dynamic feature of stream segregation is known as "build-up." Recent studies of mammalian A1 have tested the hypothesis that the build-up effect stems from long-term (i.e., multi-second) adaptation of neural responses (Micheyl et al., 2005; Pressnitzer et al., 2008; Snyder et al., 2006). Specifically, Micheyl et al. (2005) developed a model describing how changes over time in neural responses in A1 to repeated, interleaved tones could be related to proportions of "two streams" responses measured psychophysically. According to the model, spike counts evoked by A and B tones in units with a characteristic frequency (CF) corresponding to either of these frequencies are compared to a pre-determined threshold. If the count exceeds the threshold, the corresponding tone is declared "detected" by the considered unit, otherwise, the tone is declared "undetected." A "one stream" response occurs when neurons with a CF corresponding to the frequency of the A tone produce above-threshold responses to both the A and B tones. By comparison, a "two streams" response is produced whenever neurons with a CF corresponding to A fail to respond above threshold to the B tone. This relatively simple model could account quantitatively for some features of the perceptual build-up, including its dependence on $\Delta F$. Thus far, however, both the perceptual build-up and the ability of neural adaptation to explain it have been studied only in mammals and under a limited range of stimulus parameters that focused on $\Delta F$ but did not systematically manipulate various temporal parameters also known to influence streaming (Anstis and Saida, 1985; Beauvois, 1998; Micheyl et al., 2005; van Noorden, 1975).

Here, we report results from a study of neural adaptation in the auditory forebrain (field L2) of the European starling (*Sturnus vulgaris*). We had three primary objectives. First, we tested the hypothesis that neurons in field L2 exhibit multi-second adaptation by examining patterns of change in neural responses to sequences of repeated triplets of two interleaved tones (ABA–ABA–...). Second, we examined how several stimulus-related factors determined the patterns of neural adaptation by orthogonally varying $\Delta F$ (Fig. 1c), stimulus onset asynchrony (SOA), defined as the onset of one tone to the onset of the next tone within a triplet (Fig. 1c), and tone duration (TD, Fig. 1c). Our factorial experimental design allowed us to characterize the separate and inter-dependent influences of these temporal and spectral stimulus parameters on the characteristics of neural adaptation. Finally, we used the signal detection model of Micheyl et al. (2005) to generate neurometric functions from which we derived qualitative predictions concerning the perceptual build-up of stream segregation.

As songbirds, starlings represent a good animal model for investigating the neural mechanisms for auditory stream segregation for a number of reasons. As in humans, starling social behavior involves acoustically signaling in large groups (e.g., dawn choruses and communal roosts). Like human speech, starling songs comprise long sequences of distinct acoustic elements, and the temporal structure of these elements has important biological functions (Eens et al., 1989, 1991; Gentner, 2008; Gentner and Hulse, 1998). In addition, starlings are adept at perceptually analyzing complex acoustic scenes and are able to segregate attended songs from sound mixtures that include other starlings' songs, other species' songs, and the sounds of a dawn chorus (Hulse et al., 1997; Wisniewski and Hulse, 1997). Finally, although it is presently unknown whether starlings (or any other nonhuman animals) experience the perceptual build-up of auditory streaming, previous work using the ABA– stimulus paradigm has shown that starlings can segregate interleaved tone sequences into separate auditory streams (MacDougall-Shackleton et al., 1998).

# Materials and methods

## Subjects

Four wild-caught, adult starlings (2 males, 2 females; 71.2 – 94.3 g) were used as subjects in this experiment. The care and treatment of the animals were in accordance with the procedures of animal experimentation approved by the Bezirksregierung Weser-Ems. The neurophysiological data collected for this experiment represent a portion of the data collected during a large study that was designed to investigate the mechanisms of stream segregation in the starling forebrain. Results unrelated to multi-second adaptation and its possible contribution to the build-up effect have been published elsewhere (Bee and Klump, 2004, 2005), and we refer readers to those studies for additional details on equipment, methodology, and experimental design that are not reported here.

## Surgery, electrophysiology, and sound generation

We performed surgery under general anesthesia (Isoflurane: 5% for induction, 1.5–2.5% for maintenance). Anesthetized animals were fixed in a stereotaxic holder that allowed us to implant recording electrodes into the field L2 of the right hemisphere. It has been suggested that field L2 is the homologue of layer IV of the mammalian primary auditory cortex (Jarvis et al., 2005). We implanted four recording electrodes (3.6 to 12.1 MΩ; 1 kHz a/c in 0.9% NaCl) that had been fixed to a small head-mounted microdrive that could be used to lower the electrodes into the brain. Two indifferent electrodes were implanted in the left rostral hemisphere of the brain. Electrophysiological recordings began 3–9 days following surgery.

Experimental recordings were made inside a radio-shielded sound chamber (IAC402A). Multi-unit neural activity was recorded via radio telemetry from awake birds using a head-mounted FM radio transmitter that could be fitted into a socket mounted on the bird's head. The transmitter's signal was received by a dipole antenna inside the test chamber and then demodulated by an FM tuner located outside the chamber, bandpass filtered (600–4500 Hz), and stored on a hard drive for later analyses. At the beginning of a recording session, we attached the transmitter and temporarily restrained the bird while we lowered the electrodes until a site was found at which auditory-evoked activity was elicited in response to a series of test tones. Once a suitable recording site was found, the bird was released into a test cage in the chamber and given *ad libitum* access to food and water. We were unable to sort our multi-unit recordings into reliable single unit data. Thus, we currently lack data on the degree of similarity in the properties of adaptation exhibited by the neurons composing multi-unit clusters in starling field L2. We note, however, that previous studies suggest that the cells recorded in multi-unit clusters in this area of the starling forebrain exhibit similar responses regarding other properties, such as tuning in the spectral and temporal domains (Itatani and Klump, 2009; Nieder and Klump, 1999).

We generated acoustic stimuli (16-bit, 44.1 kHz) using custom-designed software running on a Linux workstation that allowed for the synchronous playback of acoustic stimuli and recording of neural responses. The output of the computer soundcard was amplified and presented through a speaker mounted from the ceiling of the sound chamber approximately 70 cm above the position of a starling sitting in\ the test cage. The frequency response inside the test cage was flat (± 4 dB) over the range of frequencies used in this study. For each multi-unit recording site, we generated a frequency tuning curve prior to starting an experimental trial using procedures detailed in Bee & Klump (2004). We determined the recording site's CF as the frequency with the lowest threshold. At the completion of all recordings from a particular bird, it was sacrificed and standard histological methods were used to confirm that electrodes were implanted in field L2 (Bee and Klump, 2004; Nieder

and Klump, 1999). In total, we made multi-unit recordings from 46 recording sites in field L2 from the four birds.

## Experimental design

The acoustic stimuli consisted of repeating sequences of pure-tone triplets, ABA– ABA– ABA–…, where A and B represent tones of (usually) different frequencies, and the dash stands for a silent gap. Three parameters were varied across stimulus sequences (Fig. 1c): the frequency difference ($\Delta$F) between the A and B tones, the stimulus onset asynchrony (SOA), and the tone duration (TD).

We fixed the frequency of A tones at the recording site's CF and varied the frequency of the B tone away from that of the CF over a 1-octave range along a semitone scale by a $\Delta$F value of 0, 2, 4, 6, 8, 10, or 12 semitones. The frequency of the B tone was constant within a given stimulus sequence. (Note that in the 0-semitone condition, the stimulus comprised sequences of AAA– triplets, but for convenience we refer to this as an ABA– triplet in which $\Delta$F = 0 semitones.) For recording sites with CFs above about 1 kHz and below about 3 kHz, the direction of $\Delta$ imposed on the B tones relative to the CF was determined randomly; for CFs below 1 kHz or above 3 kHz, the frequency of the B tones was increased or decreased, respectively, to ensure that the frequencies of the B tones remained well within the starling's hearing range (Dooling et al., 1986; Klump et al., 2000). The TD within a triplet was set at 25 ms, 40 ms, or 100 ms. Each tone was gated on and off with 5-ms Gaussian ramps. The SOA was set to 100%, 200%, 400%, or 800% of TD. The duration of the silent gap between consecutive triplets was equal to that of twice the SOA minus the TD.

The seven $\Delta$Fs, three TDs, and four SOAs were tested in all factorial combinations, resulting in 84 different stimulus sequences. For each stimulus sequence, the ABA triplet and the following silent inter-triplet interval were repeated 29 times in sequence. Across the factorial combinations of SOA and TD, these stimulus sequences ranged between about 3s and 93 s in duration. For those triplets in which responses to one or more tones were contaminated by movement artifacts (see Bee and Klump, 2004, for details), the spike count(s) for the affected tone(s) were estimated using linear interpolation of the spike counts in response to the same tone(s) in the two adjacent triplets. Responses to stimuli that had more than nine triplets affected by movement artifacts were rejected and these stimuli were automatically repeated. At each recording site, we presented the stimulus sequences in a different randomized order at 70 dB SPL (re. 20 µPa, fast RMS, C-weighted) as recorded at the approximate position of a bird's head while sitting in the test cage. There is close agreement between a multi-unit site's CF and the best frequency of neurons stimulated with pure tones at the suprathreshold intensities at which our triplet sequences were presented (Bee and Klump, 2004; Nieder and Klump, 1999). A silent interval of at least 7 s separated consecutive stimulus sequences to limit adaptation across sequences. Silent intervals shorter than 7 s reset the cumulative build-up of stream segregation in human perceptual experiments (Bregman, 1978).

## Data analysis

Spike counts were computed within 20-ms windows that started between 11 and 14 ms after physical tone onset according to the measured response latency of the recording site (see Bee and Klump, 2004, for details). Neurons in starling field L2 exhibit 'primary-like' responses to pure tones, in which the discharge rate during the first 25 ms following response onset is significantly higher than the subsequent sustained excitation (Nieder and Klump, 1999). Therefore, the 20-ms time window over which spike counts were quantified focused our investigation on the phasic onset responses to tones. We chose the 20-ms

analysis window because it represented the longest duration common to all three tone durations (25, 40, and 100 ms) that also excluded the 5-ms off ramp for the 25-ms TD. This was important, because using an analysis window that was specific to each TD could introduce a potential confound based on the covariation between the duration of the analysis window and the duration of the tones. This confound could, in turn, make it difficult or impossible to examine the effects of TD on adaptation that were independent of simple differences in the duration of the analysis window. For example, increases in the magnitude of forward suppression that might occur with longer-duration tones could become more difficult to detect if tone duration and the duration of the analysis window increased concomitantly. Nevertheless, because field L2 neurons also have a sustained response following the phasic onset, we also conducted parallel analyses using spike-count windows that spanned the entire duration of a tone at each specified level of TD. The results of these TD-specific analyses are fully described in the Supplementary Material, and we highlight TD-related differences from analyses based on 20-ms spike-count windows in the Results.

The spike-count data as a function of time were fitted with a single time-constant exponential-decay function (Harris and Dallos, 1979; Huang, 1981; Kiang et al., 1965; Smith, 1979; Smith and Zwislocki, 1975) defined by the following equation:

$$C(t) = C_0 - De^{-Rt}. \tag{1}$$

In this equation, $C(t)$ denotes the spike count at time $t$ in seconds relative to the offset of the spike-count window for the tone of the considered type (leading A, B, or trailing A) in the first triplet of the sequence; $C_0$ denotes the "initial spike count," or spike count at $t = 0$; $D$ denotes the dynamic range of adaptation, defined as the difference between the initial and asymptotic (predicted) spike counts; and $R$ represents the adaptation rate, which is inversely related to the time constant of adaptation, $\tau$. The variables $C_0$, $D$, and $R$ were treated as free parameters in the fitting procedure. Fits were performed using a least-square fitting procedure based on the Nelder-Mead algorithm (Matlab, The MathWorks, Natick MA). The resulting best-fitting parameter values were analyzed using repeated-measures analyses of variance (rmANOVA) computed using Statistica 7.1 (StatSoft, 2006). We used a significance criterion of $\alpha = 0.05$ and we report partial $\eta^2$ as a measure of effect size (i.e., an estimate of the extent to which the null hypothesis of "no effect" is false). Partial $\eta^2$ can vary from 0 to 1 and corresponds to the proportion of the combined effect and error variance that can be attributed to the effect, and thus represents a non-additive "variance-accounted-for" measure of effect size. The interpretation of partial $\eta^2$ values is similar to that of the more familiar coefficient of determination ($r^2$).

In order to generate neurometric functions, we transformed the spike-count data into estimated proportions of "two-streams" responses using a model similar to that described in Micheyl et al. (2005). First, the mean spike-count data (across all 46 sites) in response to the B tone were fitted with exponential-decay functions of time, as given by Equation 1. Second, the proportion of two-streams responses was estimated as

$$P(t) = B(k, N, p(t)) \tag{2}$$

where $B(k, N, p(t))$ denotes the cumulative probability density function of a binomial distribution with parameter $N$ (number of 1-ms bins in the 20-ms spike-count window) fixed at 20, and parameter $p(t)$ (probability of a spike occurring in each 1-ms bin) set to $C(t)/20$, evaluated at $k$. The original model by Micheyl et al. (2005) used empirically derived spike-

count distributions, which were estimated by pooling neural responses from a single A1 unit across multiple presentation of the same stimulus sequence. Due to the large number of stimulus conditions that were tested in the current study, each stimulus sequence was only presented once at each recording site. As a result, empirical spike-count distributions could not be estimated for each unit. Consequently, a binomial distribution was assumed. This distribution reflects the simplifying assumption that the spike-generation process can be approximated as a renewal process with a constant-event probability and that the probabilities of events within 1-ms bins are independent. As in the model of Micheyl et al. (2005), it is assumed that a two-stream response is chosen whenever the value of the decision variable is less than the decision "criterion" (here, $k$); otherwise, the response is one stream. The value of $k$ was set to equal two times the average spontaneous rate of field L2 neurons (Nieder and Klump, 1999).

## Results

### General patterns of neural responses

Spike counts decreased over time as a function of repeated triplet presentations. Figures 2 and 3 illustrate how spike counts changed over time for the 25-ms and 100-ms TDs, the 100% and 800% SOAs, and the 2- and 10-semitone ΔFs. For many conditions, the spike counts evoked by the A and B tones decreased rapidly at first, and then more slowly over the course of several seconds. Some adaptation was even evident over a time course of about 80 s in responses to A tones in the longest tone sequences used (Fig. 3b, d; TD = 100 ms and SOA = 800%). The magnitudes of the decreases from the initial spike count were generally larger at shorter SOAs (Figs 2a, 2c, 3a, 3c) than at longer SOAs (Figs 2b, 2d, 3b, 3d). This trend is most evident by comparing the solid and dashed lines in each plot of Figures 2 and 3. In addition, for the B tone response (middle column in Figs 2 & 3), there was usually more adaptation at smaller ΔFs (Figs 2a, 2b, 3a, 3b) than at larger ΔFs (Figs 2c, 2d, 3c, 3d). To investigate these general trends in greater detail, we performed ANOVAs comparing the effects of ΔF, SOA, and TD on the parameters of the exponential-decay functions fitted to the spike count data from each recording site.

### Analyses of factors influencing adaptation

**Initial spike counts ($C_0$)**—Initial spike counts ($C_0$) in response to the first ABA triplet in a stimulus were most strongly affected by a tone's position within a triplet (leading A, B, or trailing A), followed by differences in SOA and ΔF (Table 1; Fig. 4a–c). Compared to the initial spike counts elicited by both of the flanking A tones, those in response to the middle B tones were generally lower and inversely related to ΔF (see Fig. 4b,c). This can be explained simply by frequency selectivity: as ΔF increased, the frequency of the B tones moved away from the CF, whereas the A tones always remained at the CF. Initial spike counts elicited by B tones and trailing A tones increased with SOA (see Fig. 4a,c). In contrast, there were negligible effects of SOA and ΔF on the initial spike counts in response to the leading A tone of each triplet (Fig. 4a,c).

Compared to the effects of ΔF, SOA, and tone position, differences in TD generally had relatively smaller effects on initial spike counts (Table 1; Fig. 4a–c). This is not surprising, because in these analyses, the spike-count window was fixed at 20 ms regardless of TD. In the Supplementary Material, however, we show that when the analysis window equaled the tone duration, initial spike counts increased significantly as a function of TD; this effect was expected because more of the sustained portion of the primary-like response was included in the spike-count window.

The stimulus-dependent patterns in initial spike counts based on 20-ms analysis windows were clearly reflected in the results of a 3 Tone × 7 ΔF × 4 SOA × 3 TD rmANOVA. All four main effects were significant (Table 1), with the effect of Tone associated with the largest effect size, followed by SOA and ΔF, with TD having the smallest effect size. The two-way interaction of ΔF × Tone had the largest effect size of any interaction (Table 1, $\eta^2$ = 0.75), followed by the SOA × Tone interaction (Table 1, $\eta^2$ = 0.45). While several other interactions were also statistically significant, they were generally associated with much smaller effect sizes (Table 1, $\eta^2 \leq 0.06$). Similar overall patterns of results were described previously for spike counts that were averaged across all triplets in a stimulus sequence, and these patterns were attributed to the effects of frequency selectivity and forward suppression (Bee and Klump, 2004,2005). We do not discuss initial spike counts further and refer readers instead to our detailed discussions of the effects of frequency selectivity and forward suppression in our earlier studies. In the present study, we were primarily concerned with the changes in spike counts that occurred over the course of repeating ABA triplets.

**Dynamic range of adaptation (D)—**The dynamic range of adaptation (*D*) corresponds to the magnitude of the change in the fitted spike counts between the first and the last triplet in the sequence of 29 repeated triplets. Three general patterns emerged from an analysis of the dynamic range of adaptation. First, the dynamic range decreased markedly as SOA increased (Fig. 4d). That is, more adaptation occurred at shorter SOAs (and hence, faster tone repetition rates). These SOA-dependent differences in dynamic range did not depend heavily on, or vary consistently with, other manipulated stimulus properties (Fig. 4d,f). Second, the dynamic ranges in response to the middle B tone, but not the leading or trailing A tones, decreased somewhat as ΔF became larger (Fig. 4e,f). This was expected because, as ΔF increased, the frequency of the B tone moved farther from the CF, and the magnitude of the responses to the B tones approached the spontaneous rate. Finally, in these analyses using a fixed spike-count window of 20 ms, the effects of TD on dynamic range were relatively small; across some conditions, there was a slight trend for dynamic range to be smaller at the shortest (25 ms) TD (Fig. 4d,e). As shown in the Supplementary Material, using spike-count windows equal to TD revealed larger and direct relationships between TD and dynamic range. This difference between the two types of analyses can be explained by considering that increasing the duration of spike-count windows from 20 ms to TD resulted in an approximately proportional overall increase in spike counts as a function of TD. Since dynamic range was measured in both analyses as the *arithmetic* difference between initial and asymptotic spike counts, a proportional overall increase in spike counts resulted in a larger dynamic range using TD-specific analysis windows.

The overall patterns of changes in dynamic range described above for the 20-ms analysis windows were reflected in the outcome of a 3 Tone × 7 ΔF × 4 SOA × 3 TD rmANOVA. There were significant main effects of Tone, TD, SOA, and ΔF, and significant two-way interactions of TD × Tone, ΔF × Tone, and SOA × TD (Table 1). No other interactions were significant. The main effects of Tone (Table 1, $\eta^2$ = 0.51) and SOA (Table 1, $\eta^2$ = 0.41) were associated with the largest effects sizes, followed by the main effect of TD (Table 1, $\eta^2$ = 0.17). All other factors in the ANOVA model, including several statistically significant ones, were associated with relatively smaller effect sizes (Table 1, $\eta^2 \leq 0.10$).

**Rate of adaptation (R)—**The rate of adaptation (*R*) characterizes how rapidly spike counts decreased across repeated triplets, and it is inversely related to the time constant of adaptation. The most striking effect on the rate of adaptation was that of SOA (Fig. 4g,i). Adaptation occurred more rapidly at shorter SOAs (i.e., at faster tone repetition rates). Stimulus-dependent differences in the rates of adaptation were largely independent of ΔF (Fig. 4h,i) and were not strongly influenced by tone position within a triplet (Fig. 4g–i).

Finally, in these analyses using fixed (20 ms) spike-count windows, the rates of adaptation increased at shorter TDs (Fig. 4h,i), and there was a tendency for the effects of SOA to be more pronounced at the 25-ms and 40-ms TDs compared with the 100-ms TD (Fig. 4g). As shown in the Supplementary Material, the stimulus-dependent differences in adaptation rate were similar in separate analyses based on spike counts computed over the entire TD.

The rmANOVA for the rate of adaptation yielded significant main effects of Tone, SOA, and TD, but not $\Delta F$ (Table 1). The effect size associated with Tone (Table 1, $\eta^2 = 0.08$) was considerably smaller than those associated with the main effects of SOA (Table 1, $\eta^2 = 0.76$) and TD (Table 1, $\eta^2 = 0.66$). The only significant interaction was that between SOA and TD (Table 1, $\eta^2 = 0.11$); the magnitude of the SOA-dependent differences in adaptation rates was smaller at longer TDs (Fig. 4g). All other interactions were non-significant and associated with small effect sizes (Table 1, $\eta^2 \leq 0.04$).

### Neurometric functions

To explore how the observed patterns of neural adaptation might be related to the build-up of stream segregation, the spike counts elicited by the B tones were transformed into predicted proportions of "two streams" responses (i.e., neurometric functions) using the model of Micheyl et al. (2005) as described above. The neurometric functions are shown in Figures 5a and 5b. The different colors within each panel represent different $\Delta F$s; to avoid clutter, only $\Delta F$s of 2, 6, and 10 semitones are shown. A general trend, which can be seen in all panels, was an increase in the proportions of "two streams" responses with increasing $\Delta F$. This effect, which can be attributed to frequency selectivity, is consistent with earlier findings in which neurometric functions for the build-up of streaming have been computed (Micheyl et al., 2005; Pressnitzer et al., 2008). Figure 5 also illustrates how differences in SOA (Fig. 5a) and TD (Fig. 5b) influenced the shapes of neurometric functions. Consider first the effects of SOA. Each panel in Figure 5a shows neurometric functions for one TD, from 25 ms at the top to 100 ms at the bottom, with differences in SOA at each TD indicated by different line styles. Both the rate and the asymptotic level of the build-up were higher at shorter SOAs (i.e., faster tone rates); this is best seen by comparing the different line styles within each color group. The effects of SOA were most apparent at intermediate and large $\Delta F$s (green and red curves); at the 2-semitone $\Delta F$ (blue curves), the proportion of "two streams" responses was low, even at the shortest SOA. The effects of SOA can be most easily contrasted with those of TD by comparing the relative spread among the different line styles in Figures 5a (SOA) and 5b (TD). In the latter, each panel shows neurometric functions for one SOA, from 100% at the top to 800% at the bottom, with differences in TD at each SOA indicated by different line styles. Notice that the differences among the lines within each color group are much less pronounced as a function of TD (Fig. 5b) than as a function of SOA (Fig. 5a).

## Discussion

Adaptation of neural responses to ongoing or repeating stimulation is a common feature of sensory systems, where it provides a mechanism for adjusting the limited dynamic range of neurons to changing stimulus statistics (Baccus and Meister, 2002; Brenner et al., 2000; Carandini and Ferster, 1997; Dean et al., 2005; Fairhall et al., 2001; Müller et al., 1999). One of the main findings of the present study was that neurons in the avian auditory forebrain exhibit multi-second adaptation in response to tone sequences. In most of the conditions tested in this study, initial decreases in spike counts occurred relatively rapidly over the first 1–3 s of sequence onset. However, in some conditions more prolonged decreases were observed, which could sometimes span tens of seconds.

Historically, adaptation has most often been identified with a rapid, short-term change in a neuron's responsiveness, with time constants of less than 100 ms (Abbas, 1984; Boettcher et al., 1990; Chimento and Schreiner, 1990, 1991; Eggermont and Spoor, 1973a, [b]; Huang and Buchwald, 1980; Møller, 1976; Smith, 1977, 1979; Smith and Zwislocki, 1975; Westerman and Smith, 1984; Yates et al., 1985). Early reports of long-term (multi-second) adaptation remained rare (Kiang et al., 1965). In a study of North American bullfrogs, Megela and Capranica (1983) reported multi-second adaptation in the torus semicircularis (inferior colliculus) and thalamus, but not in auditory nerve fibers. More recently, studies of cats, rats, mice, and Guinea pigs, have reported multi-second adaptation to occur at many levels of the mammalian auditory system, including the auditory nerve (Javel, 1996), cochlear nucleus (Pressnitzer et al., 2008), inferior colliculus (Malmierca et al., 2009; Perez-Gonzalez et al., 2005), thalamus (Anderson et al., 2009), and primary auditory cortex (Asari and Zador, 2009; Ulanovsky et al., 2004; Ulanovsky et al., 2003; von der Behrens et al., 2009). The present results extend these findings by demonstrating that multi-second adaptation to repeating tones also occurs in the avian forebrain. This taxonomic diversity suggests that multi-second neural adaptation could be a common feature of vertebrate auditory processing.

One of the hypothesized functions of multi-second adaptation in the auditory system is to attenuate or "filter out" neural representations that correspond to repetitive elements, so that responses to "novel" sound events are comparatively enhanced (Jaaskelainen et al., 2004; Ulanovsky et al., 2003, 2004). For example, Ulanovsky et al. (2003) found that neurons in cat A1 responded less strongly to a tone when it occurred frequently within a sequence compared to when the same tone had a lower probability of occurrence. This has been referred to as "stimulus-specific adaptation" (SSA), and it has been suggested to provide a single-unit basis for the mismatch negativity (MMN) component of the long-latency evoked potentials (Jaaskelainen et al., 2004; Nelken and Ulanovsky, 2007; Ulanovsky et al., 2003). None of our results are inconsistent with the findings of these earlier studies of SSA. However, the influence of SSA cannot explain the patterns of results that we observed to occur as functions of ΔF, SOA, and TD, as the relative frequencies of occurrence of the A and B tones in our stimuli were constant across all stimulus conditions.

A second hypothesized role of multi-second neural adaptation in the auditory system relates to the formation of auditory streams (Micheyl et al., 2005; Pressnitzer et al., 2008). Specifically, it has been shown that the properties of adaptation to repeating ABA sequences in mammalian A1 and cochlear nucleus are consistent with those of the build-up of stream segregation (Micheyl et al., 2005; Pressnitzer et al., 2008). To date, however, these studies have focused on the influence of ΔF, and they have not measured the influence of specific temporal parameters on multi-second neural adaptation. Three such parameters can be distinguished, including SOA and TD, which together determine the third, inter-stimulus interval (ISI). In the present study, we varied SOA and TD orthogonally and did not systematically vary ISI. Therefore, we discuss the following conclusions in terms of the parameters we actually manipulated with the understanding that SOA and ISI are directly related and TD and ISI are inversely related.

One of our key findings was that SOA was a major determinant of both the dynamic range and rate of neural adaptation. The dynamic range of adaptation decreased most markedly as a function of increasing SOA for all three tone types (leading A, B, trailing A). The rate of adaptation also decreased as a function of increasing SOA for all three tone types. In other words, there was more adaptation and it occurred more rapidly at shorter SOAs, which correspond both to faster tone repetition rates and shorter ISIs. Compared to SOA, TD and ΔF had small effects on neural adaptation, at least over the range of values tested here.

Consistent with these observations, when the spike counts were transformed into neurometric functions, we found that the predicted build-up of "two streams" responses decreased as SOA increased but was affected to a lesser extent by differences in TD. It will be important to test these general predictions in future psychophysical studies. In addition, future studies of the perceptual build-up and its underlying neural mechanisms should consider varying SOA, TD, and ISI orthogonally in different tests (e.g., two at a time *sensu* Bregman et al. 2000) in an effort to tease apart the relative contributions of each specific time interval.

It is interesting to note that the stimulus conditions under which multi-second adaptation effects were most prominent were those that involved relatively fast tone repetition rates. This suggests that multi-second adaptation could be related to incomplete recovery from either short-term adaptation (Abbas, 1984; Boettcher et al., 1990; Chimento and Schreiner, 1990, 1991; Huang and Buchwald, 1980) or forward suppression (Bee and Klump, 2004, 2005; Fishman et al., 2004; Fishman et al., 2001; Wehr and Zador, 2005). In vivo whole-cell recordings suggest an involvement of synaptic depression in forward suppression beyond 50–100 ms (Wehr and Zador, 2005). Moreover, intracellular recordings from mammalian auditory cortex suggest that synaptic depression plays a role in multi-second adaptation (Asari and Zador, 2009). To the extent that synaptic depression accumulates over the course of several seconds, this could explain the influence of SOA on multi-second adaptation observed here.

Uncovering the neural mechanisms underlying the perceptual organization of acoustic scenes remains an important goal in auditory neuroscience (Carlyon, 2004; Snyder and Alain, 2007). Our results lend additional support to the hypothesis that the build-up of stream segregation could be related to multi-second adaptation (Micheyl et al., 2005; Pressnitzer et al., 2008). This conclusion is consistent with the broader notion that low level, "bottom-up" processes, such as frequency filtering, forward suppression, and long-term adaptation (Bee and Klump, 2004, 2005; Fishman et al., 2001, 2004; Micheyl et al., 2005; Pressnitzer et al., 2008), contribute to the perceptual organization of sound sequences (Bregman, 1990; Carlyon, 2004). One limitation of this study (and all previous studies of the neural basis of streaming in animal models) relates to a lack of behavioral data on the build-up effect in nonhuman species. Therefore, an important next step will be to determine the extent to which starlings and other animals experience the perceptual build-up of auditory streaming.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Abbreviations

| | |
|---|---|
| A1 | Mammalian primary auditory cortex |
| CF | Characteristic frequency |
| ΔF | Frequency separation |
| SOA | Stimulus onset asynchrony |
| SSA | Stimulus-specific adaptation |
| TD | Tone duration |

## Acknowledgments

## References

Abbas PJ. Recovering from long-term and short-term adaptation of the whole nerve action potential. Journal of the Acoustical Society of America 1984;75:1541–1547. [PubMed: 6736416]

Anderson LA, Christianson GB, Linden JF. Stimulus-pecific adaptation cccurs in the auditory thalamus. Journal of Neuroscience 2009;29:7359–7363. [PubMed: 19494157]

Anstis S, Saida S. Adaptation to auditory streaming of frequency-modulated tones. J Exp Psychol-Hum Percept Perform 1985;11:257–271.

Asari H, Zador AM. Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. J Neurophysiol 2009;102:2638–2656. [PubMed: 19675288]

Baccus SA, Meister M. Fast and slow contrast adaptation in retinal circuitry. Neuron 2002;36:909–919. [PubMed: 12467594]

Beauvois MW. The effect of tone duration on auditory stream formation. Perception & Psychophysics 1998;60:852–861. [PubMed: 9682609]

Bee MA, Klump GM. Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. J Neurophysiol 2004;92:1088–1104. [PubMed: 15044521]

Bee MA, Klump GM. Auditory stream segregation in the songbird forebrain: Effects of time intervals on responses to interleaved tone sequences. Brain, Behavior and Evolution 2005;66:197–214.

Bee MA, Micheyl C. The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? Journal of Comparative Psychology 2008;122:235–251. [PubMed: 18729652]

Boettcher FA, Salvi RJ, Saunders SS. Recovery from short-term adaptation in single neurons in the cochlear nucleus. Hearing Research 1990;48:125–144. [PubMed: 2249955]

Bregman AS. Auditory streaming is cumulative. J Exp Psychol Hum Percept Perform 1978;4:380–387. [PubMed: 681887]

Bregman, AS. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press; 1990.

Brenner N, Bialek W, van Steveninck RD. Adaptive rescaling maximizes information transmission. Neuron 2000;26:695–702. [PubMed: 10896164]

Brokx JPL, Nooteboom SG. Intonation and the perceptual separation of simultaneous voices. Journal of Phonetics 1982;10:23–36.

Carandini M, Ferster D. A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. Science 1997;276:949–952. [PubMed: 9139658]

Carlyon RP. How the brain separates sounds. Trends Cogn Sci 2004;8:465–471. [PubMed: 15450511]

Carlyon RP, Cusack R, Foxton JM, Robertson IH. Effects of attention and unilateral neglect on auditory stream segregation. J Exp Psychol-Hum Percept Perform 2001;27:115–127. [PubMed: 11248927]

Chimento TC, Schreiner CE. Time course of adaptation and recovery from adaptation in the cat auditory-nerve neurophonic. Journal of the Acoustical Society of America 1990;88:857–864. [PubMed: 2212311]

Chimento TC, Schreiner CE. Adaptation and recovery from adaptation in single fiber responses of the cat auditory nerve. Journal of the Acoustical Society of America 1991;90:263–273. [PubMed: 1652600]

Cusack R. The intraparietal sulcus and perceptual organization. Journal of Cognitive Neuroscience 2005;17:641–651. [PubMed: 15829084]

Cusack R, Deeks J, Aikman G, Carlyon RP. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. J Exp Psychol-Hum Percept Perform 2004;30:643–656. [PubMed: 15301615]

Dean I, Harper NS, McAlpine D. Neural population coding of sound level adapts to stimulus statistics. Nature Neuroscience 2005;8:1684–1689.

Dooling RJ, Okanoya K, Downing J, Hulse S. Hearing in the starling (*Sturnus vulgaris*): Absolute thresholds and critical ratios. Bulletin of the Psychonomic Society 1986;24:462–464.

Eens M, Pinxten R, Verheyen RF. Temporal and sequential organization of song bouts in the starling. Ardea 1989;77:75–86.

Eens M, Pinxten R, Verheyen RF. Organization of song in the European starling: Species-specificity and individual differences. Belgian Journal of Zoology 1991;121:257–278.

Eggermont JJ, Spoor A. Cochlear adaptation in guinea pigs: A quantitative description. Audiology 1973a;12:193–220. [PubMed: 4722021]

Eggermont JJ, Spoor A. Masking of action potentials in the guinea pig cochlea and its relation to adaptation. Audiology 1973b;12:221–241. [PubMed: 4722022]

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA. Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 2009;61:317–329. [PubMed: 19186172]

Fairhall AL, Lewen GD, Bialek W, van Steveninck RRD. Efficiency and ambiguity in an adaptive neural code. Nature 2001;412:787–792. [PubMed: 11518957]

Fay RR. Auditory stream segregation in goldfish (*Carassius auratus*). Hearing Research 1998;120:69–76. [PubMed: 9667432]

Fay RR. Spectral contrasts underlying auditory stream segregation in goldfish (*Carassius auratus*). Jaro 2000;1:120–128. [PubMed: 11545140]

Fay, RR. Sound source perception and stream segregation in nonhuman vetebrate animals. In: Yost, WA.; Popper, AN.; Fay, RR., editors. Auditory Perception of Sound Sources. New York: Springer; 2008. p. 307-323.

Fishman YI, Arezzo JC, Steinschneider M. Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. Journal of the Acoustical Society of America 2004;116:1656–1670. [PubMed: 15478432]

Fishman YI, Reser DH, Arezzo JC, Steinschneider M. Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hearing Research 2001;151:167–187. [PubMed: 11124464]

Gentner TQ. Temporal scales of auditory objects underlying birdsong vocal recognition. Journal of the Acoustical Society of America 2008;124:1350–1359. [PubMed: 18681620]

Gentner TQ, Hulse SH. Perceptual mechanisms for individual vocal recognition in European starlings, Sturnus vulgaris. Animal Behaviour 1998;56:579–594. [PubMed: 9784206]

Gutschalk A, Oxenham AJ, Micheyl C, Wilson C, Melcher JR. Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. Journal of Neuroscience 2007;27:13074–13081. [PubMed: 18045901]

Harris DM, Dallos P. Forward masking of auditory-nerve fiber responses. J Neurophysiol 1979;42:1083–1107. [PubMed: 479921]

Hartmann WM, Johnson D. Stream segregation and peripheral channeling. Music Percept 1991;9:155–184.

Huang C, Buchwald JS. Changes of acoustic nerve and cochlear nucleus evoked-potentials due to repetitive stimulation. Electroen Clin Neuro 1980;49:15–22.

Huang CM. Time constants of acoustic adaptation. Electroen Clin Neuro 1981;52:394–399.

Hulse SH. Auditory scene analysis in animal communication. Advances in the Study of Behavior 2002;31:163–200.

Hulse SH, MacDougall-Shackleton SA, Wisniewski AB. Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (*Sturnus vulgaris*). Journal of Comparative Psychology 1997;111:3–13. [PubMed: 9090135]

Itatani N, Klump GM. Auditory streaming of amplitude-modulated sounds in the songbird forebrain. J Neurophysiol 2009;101:3212–3225. [PubMed: 19357341]

Izumi A. Auditory sequence discrimination in Japanese monkeys: Effect of frequency proximity on perceiving auditory stream. Psychologia 2001;44:17–23.

Izumi A. Auditory stream segregation in Japanese monkeys. Cognition 2002;82:B113–B122. [PubMed: 11747868]

Jaaskelainen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levanen S, Lin FH, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW. Human posterior auditory cortex gates novel sounds to consciousness. Proceedings of the National Academy of Sciences of the United States of America 2004;101:6809–6814. [PubMed: 15096618]

Jarvis E, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, Striedter G, Wild JM, Ball GF, Dugas-Ford J, Durand SE, Hough GE, Husband S, Kubikova L, Lee DW, Mello CV, Powers A, Siang C, Smulders TV, Wada K, White SA, Yamamoto K, Yu J, Reiner A, Butler AB. Avian brains and a new understanding of vertebrate brain evolution. Nature Reviews Neuroscience 2005;6:151–159.

Javel E. Long-term adaptation in cat auditory-nerve fiber responses. Journal of the Acoustical Society of America 1996;99:1040–1052. [PubMed: 8609287]

Kiang, NYS.; Watanabe, T.; Thomas, EC.; Clark, EF. Discharge patterns of single fibers in the cat auditory nerve. Cambridge, MA: M.I.T. Press; 1965.

Klump, GM.; Langemann, U.; Gleich, O. The European starling as a model for undestanding perceptual mechanisms. In: Manley, GA.; Fastl, H.; Kössl, M.; Oeckinghaus, H.; Klump, GM., editors. Auditory Worlds: Sensory Analysis and Perception in Animals and Man. Weinheim, Germany: Wiley-VCH; 2000.

Kondo HM, Kashino M. Involvement of the Thalamocortical Loop in the Spontaneous Switching of Percepts in Auditory Streaming. Journal of Neuroscience 2009;29:12695–12701. [PubMed: 19812344]

MacDougall-Shackleton SA, Hulse SH, Gentner TQ, White W. Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences. Journal of the Acoustical Society of America 1998;103:3581–3587. [PubMed: 9637040]

Malmierca MS, Cristaudo S, Perez-Gonzalez D, Covey E. Stimulus-cpecific adaptation in the inferior colliculus of the anesthetized rat. Journal of Neuroscience 2009;29:5483–5493. [PubMed: 19403816]

Megela AL, Capranica RR. A Neural and behavioral study of auditory habituation in the bullfrog, *Rana catesbeiana*. Journal of Comparative Physiology 1983;151:423–434.

Micheyl C, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Courtenay Wilson E. The role of auditory cortex in the formation of auditory streams. Hearing Research 2007;229:116–131. [PubMed: 17307315]

Micheyl C, Tian B, Carlyon RP, Rauschecker JP. Perceptual organization of tone sequences in the auditory cortex of awake Macaques. Neuron 2005;48:139–148. [PubMed: 16202714]

Møller AR. Dynamic properties of primary auditory fibers compared with cells in cochlear nucleus. Acta Physiologica Scandinavica 1976;98:157–167. [PubMed: 983725]

Moore BCJ, Gockel H. Factors influencing sequential stream segregation. Acta Acustica United with Acustica 2002;88:320–333.

Müller JR, Metha AB, Krauskopf J, Lennie P. Rapid adaptation in visual cortex to the structure of images. Science 1999;285:1405–1408. [PubMed: 10464100]

Nelken I, Ulanovsky N. Mismatch negativity and stimulus-specific adaptation in animal models. J Psychophysiol 2007;21:214–223.

Nieder A, Klump GM. Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. Hearing Research 1999;127:41–54. [PubMed: 9925015]

Perez-Gonzalez D, Malmierca MS, Covey E. Novelty detector neurons in the mammalian auditory midbrain. European Journal of Neuroscience 2005;22:2879–2885. [PubMed: 16324123]

Pressnitzer D, Sayles M, Micheyl C, Winter IM. Perceptual organization of sound begins in the auditory periphery. Current Biology 2008;18:1124–1128. [PubMed: 18656355]

Smith RL. Short-term adaptation in single auditory-nerve fibers: Some post-stimulatory effects. J Neurophysiol 1977;40:1098–1112. [PubMed: 903799]

Smith RL. Adaptation, saturation, and physiological masking in single auditory-nerve fibers. Journal of the Acoustical Society of America 1979;65:166–178. [PubMed: 422812]

Smith RL, Zwislocki JJ. Short-term adaptation and incremental responses of single auditory-nerve fibers. Biol Cybern 1975;17:169–182. [PubMed: 1125344]

Snyder JS, Alain C. Toward a neurophysiological theory of auditory stream segregation. Psychological Bulletin 2007;133:780–799. [PubMed: 17723030]

Snyder JS, Alain C, Picton TW. Effects of attention on neuroelectric correlates of auditory stream segregation. Journal of Cognitive Neuroscience 2006;18:1–13. [PubMed: 16417678]

Snyder JS, Holder WT, Weintraub DM, Carter OL, Alain C. Effects of prior stimulus and prior perception on neural correlates of auditory stream segregation. Psychophysiology 2009;46:1208–1215. [PubMed: 19674396]

StatSoft. STATISTICA (data analysis software system). 2006 version 7.1. www.statsoft.com.

Sussman E, Ritter W, Vaughan HG. An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 1999;36:22–34. [PubMed: 10098377]

Sussman ES, Horvath J, Winkler I, Orr M. The role of attention in the formation of auditory streams. Perception & Psychophysics 2007;69:136–152. [PubMed: 17515223]

Ulanovsky N, Las L, Farkas D, Nelken I. Multiple time scales of adaptation in auditory cortex neurons. Journal of Neuroscience 2004;24:10440–10453. [PubMed: 15548659]

Ulanovsky N, Las L, Nelken I. Processing of low-probability sounds by cortical neurons. Nature Neuroscience 2003;6:391–398.

van Noorden LPAS. Temporal Coherence in the Perception of Tone Sequences. Eindhoven University of Technology. 1975

von der Behrens W, Bauerle P, Kossl M, Gaese BH. Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. Journal of Neuroscience 2009;29:13837–13849. [PubMed: 19889995]

Wehr M, Zador AM. Synaptic mechanisms of forward suppression in rat auditory cortex. Neuron 2005;47:437–445. [PubMed: 16055066]

Westerman LA, Smith RL. Rapid and short-term adaptation in auditory-nerve responses. Hearing Research 1984;15:249–260. [PubMed: 6501113]

Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ. Cortical fMRI activation to sequences of tones alternating in frequency: Relationship to perceived rate and streaming. J Neurophysiol 2007;97:2230–2238. [PubMed: 17202231]

Winkler I, Takegata R, Sussman E. Event-related brain potentials reveal multiple stages in the perceptual organization of sound. Cognitive Brain Research 2005;25:291–299. [PubMed: 16005616]

Wisniewski AB, Hulse SH. Auditory scene analysis in European starlings (*Sturnus vulgaris*): Discrimination of song segments, their segregation from multiple and reversed conspecific songs, and evidence for conspecific song categorization. Journal of Comparative Psychology 1997;111:337–350.

Yates GK, Robertson D, Johnstone BM. Very rapid adaptation in the guinea pig auditory nerve. Hearing Research 1985;17:1–12. [PubMed: 3997676]
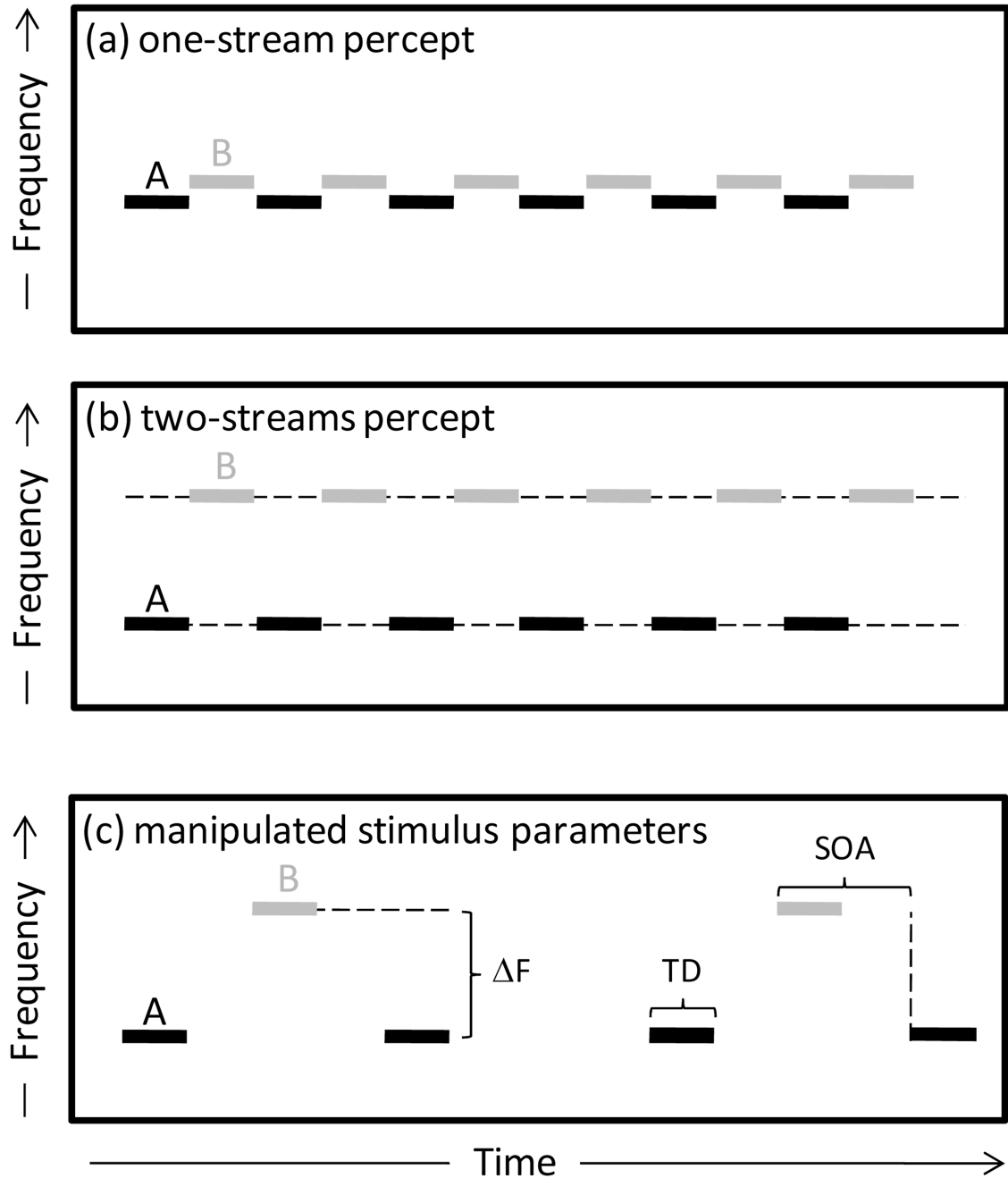
**Fig. 1.**
Auditory stream segregation. (a–b) Schematic diagram illustrating how a sequence of two interleaved tones differing only in frequency (A and B) can be perceived as either one stream (a) or two streams (b) depending on the degree of frequency separation (ΔF) between them. (c) Depiction of the three stimulus parameters manipulated in this study, including the frequency separation between the A and B tones (ΔF, in semitones), the stimulus onset asynchrony (SOA), defined as the time between the onset of two consecutive tones within a triplet, and the tone duration (TD).
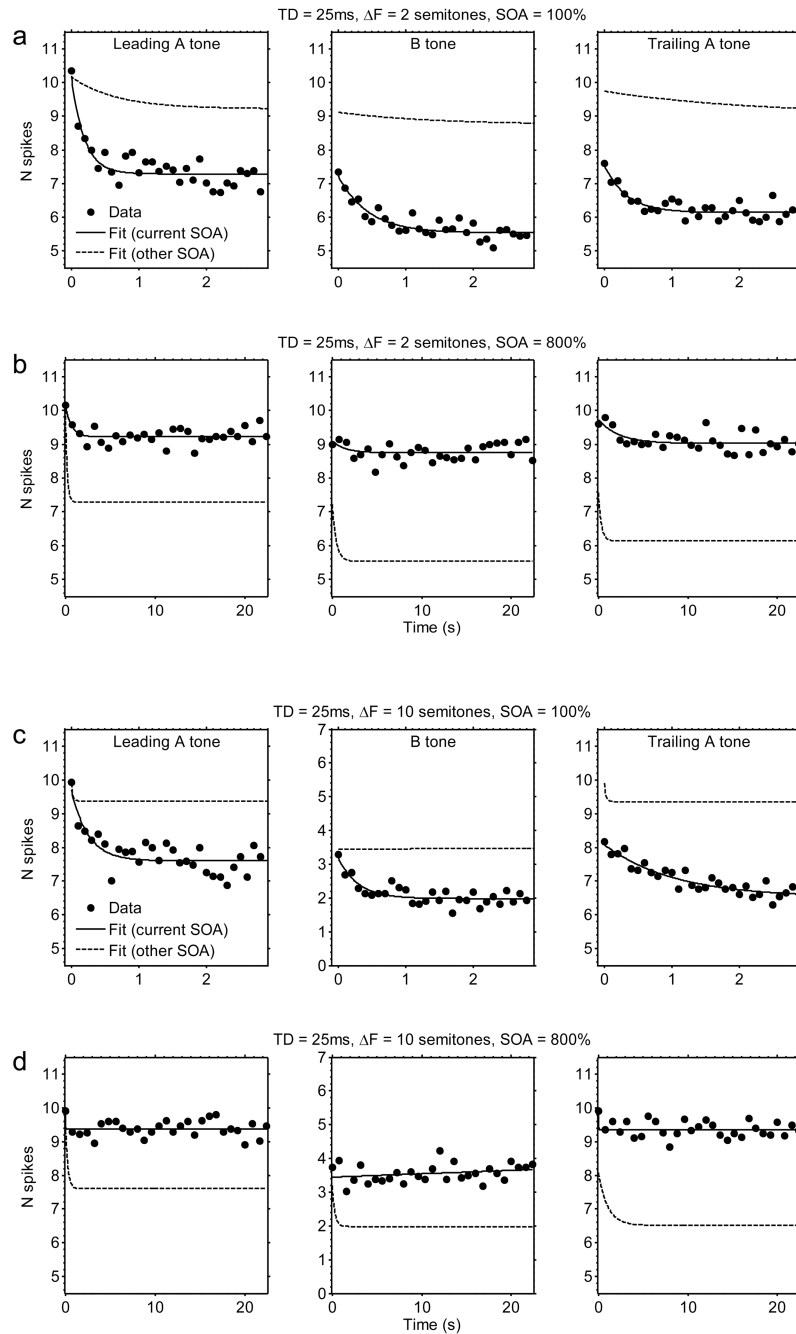
**Fig. 2.**

Actual and fitted spike-count data as a function of time. The data depicted here are for the following stimulus conditions: (a) TD = 25 ms, SOA = 100%, $\Delta$F = 2 semitones; (b) TD = 25 ms, SOA = 800%, $\Delta$F = 2 semitones; (c) TD = 25 ms, SOA = 100%, $\Delta$F = 10 semitones; (d) TD = 25 ms, SOA = 800%, $\Delta$F = 10 semitones. The points show mean spike counts averaged across all sites (N = 46). The solid line shows the best-fitting exponential-decay curve. Note that the spike counts evoked by the three tone types (leading A tone, B tone, and trailing A tone) are shown side-by-side within each row, although in reality these tones were

temporally interleaved within each triplet; this display format facilitates comparisons among spike counts evoked by the different tones. The time is expressed in s relative to the offset of the spike-count window corresponding to the first tone of the considered type in the sequence—so that all functions start at 0 s. For pairs of plots corresponding to conditions testing the same TD and $\Delta F$ but different SOAs (100% versus 800%; e.g., pair a and b, or pair c and d), the exponential-decay curve fitted to the data shown in each plot (solid line) is re-plotted as a dashed line in the adjacent paired plot showing data for the other level of SOA. This re-plotting was done to facilitate comparisons across panels corresponding to different SOAs, which have different time scales. In addition, note that plots of the actual data points are equivalent to those that result when the data are plotted as a function of triplet number.
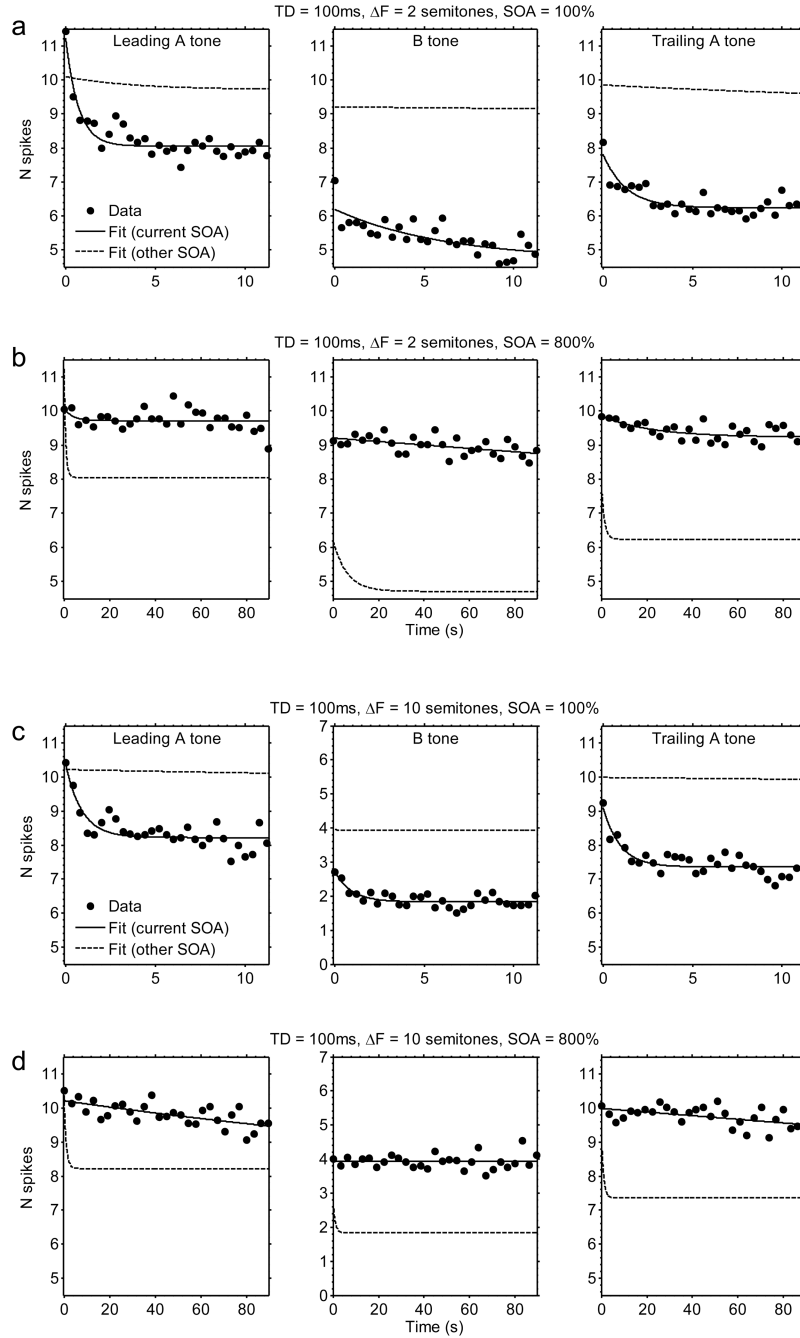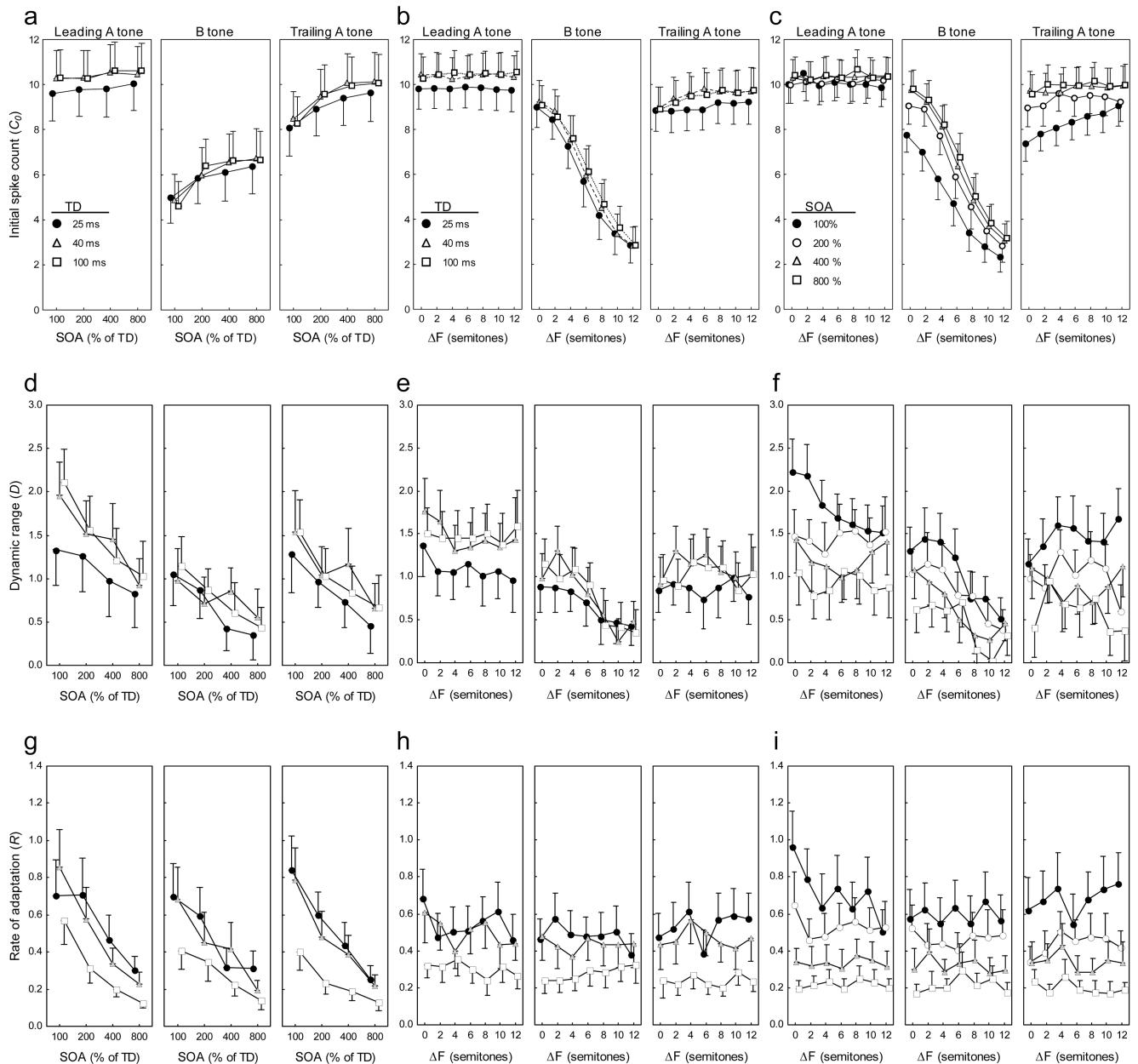
**Fig. 3.**
Actual and fitted spike-count data as a function of time. The data depicted here are for the following stimulus conditions: (a) TD = 100 ms, SOA = 100%, ΔF = 2; (b) TD = 100 ms, SOA = 800%, ΔF = 2; (a) TD = 100 ms, SOA = 100%, ΔF = 10; (b) TD = 100 ms, SOA = 800%, ΔF = 10. All other details as in Fig. 2.

**Fig. 4.**
Factors affecting responses to ABA- triplets. Shown are the mean (± s.e.m.; N = 46) fitted values for (a–c) initial spike counts ($C_0$), (d–f) dynamic range ($D$), (g–i) and the rate of adaptation ($R$). The left column of plots shows values as functions of stimulus onset asynchrony (SOA), with tone duration (TD) as the parameter. The middle column of plots depicts values as functions of frequency separation (ΔF), with TD as the parameter. The right column of plots depicts values as functions of ΔF, with SOA as the parameter. Model parameters are based on a using a 20-ms analysis window to quantify spike counts across triplets in the various conditions.
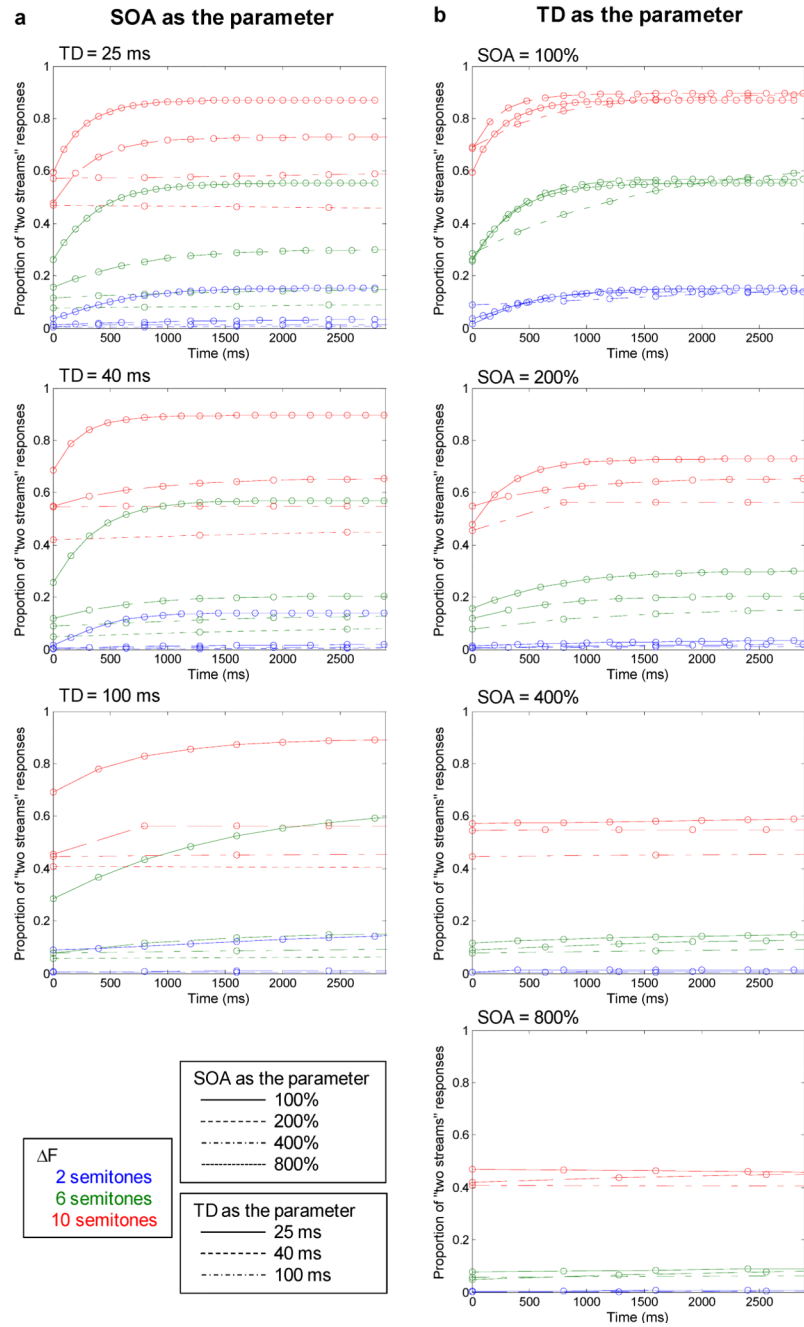
**Fig. 5.**
Neurometric functions based on neural responses to B tones in avian field L2. Depicted here is the probability of two-streams responses as a function of time, with parameters ΔF (indicated by different colors) and either (a) SOA or (b) TD (indicated by differences in line styles).

**Table 1**

Results of a 3 tone position (Tone, leading A, B, trailing A) × 7 frequency separation ($\Delta$F) × 4 stimulus onset asynchrony (SOA) × 3 tone duration (TD) repeated measures ANOVA for the fitted values of initial spike counts ($C_0$), dynamic range ($D$), and the rate of adaptation ($R$).

| Source | df | Initial Spike Counts ($C_0$) | | | Dynamic range ($D$) | | | Rate of adaptation ($R$) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $F$ | $P$ | $\eta^2$ | $F$ | $P$ | $\eta^2$ | $F$ | $P$ | $\eta^2$ |
| Tone | 2, 90 | 178.6 | <0.0001 | 0.80 | 46.9 | <0.0001 | 0.51 | 4.1 | 0.0197 | 0.08 |
| $\Delta$F | 6, 270 | 103.8 | <0.0001 | 0.70 | 5.2 | 0.0002 | 0.10 | 0.5 | 0.7805 | 0.01 |
| SOA | 3, 135 | 127.1 | <0.0001 | 0.74 | 31.4 | <0.0001 | 0.41 | 143.9 | <0.0001 | 0.76 |
| TD | 2, 90 | 47.5 | <0.0001 | 0.51 | 9.5 | 0.0004 | 0.17 | 88.4 | <0.0001 | 0.66 |
| $\Delta$F × Tone | 12, 540 | 132.7 | <0.0001 | 0.75 | 4.9 | <0.0001 | 0.10 | 1.5 | 0.1371 | 0.03 |
| SOA × Tone | 6, 270 | 37.3 | <0.0001 | 0.45 | 1.1 | 0.3401 | 0.02 | 1.7 | 0.1521 | 0.04 |
| SOA × $\Delta$F | 18, 810 | 3.0 | 0.0005 | 0.06 | 1.3 | 0.2173 | 0.03 | 0.6 | 0.8604 | 0.01 |
| SOA × TD | 6, 270 | 2.9 | 0.0145 | 0.06 | 2.3 | 0.0456 | 0.05 | 5.3 | 0.0002 | 0.11 |
| TD × $\Delta$F | 12, 540 | 1.1 | 0.3358 | 0.02 | 1.0 | 0.4675 | 0.02 | 1.0 | 0.4237 | 0.02 |
| TD × Tone | 4, 180 | 5.7 | 0.0003 | 0.11 | 2.8 | 0.0308 | 0.06 | 1.3 | 0.2814 | 0.03 |
| SOA × Tone × $\Delta$F | 36, 1620 | 3.1 | <0.0001 | 0.06 | 1.5 | 0.0800 | 0.03 | 0.9 | 0.6283 | 0.02 |
| SOA × Tone × TD | 12, 540 | 2.6 | 0.0052 | 0.05 | 1.2 | 0.2700 | 0.03 | 1.7 | 0.1043 | 0.04 |
| SOA × TD × $\Delta$F | 36, 1620 | 1.1 | 0.3552 | 0.02 | 0.7 | 0.7733 | 0.02 | 0.8 | 0.6475 | 0.02 |
| TD × $\Delta$F × Tone | 24, 1080 | 1.0 | 0.4776 | 0.02 | 0.6 | 0.8432 | 0.01 | 0.9 | 0.5429 | 0.02 |
| SOA × TD × $\Delta$F × Tone | 72, 3240 | 0.9 | 0.6334 | 0.02 | 1.0 | 0.4126 | 0.02 | 0.9 | 0.6155 | 0.02 |