

Behavioral Measures of Auditory Streaming in Ferrets (*Mustela putorius*)

Ling Ma
University of Maryland

Christophe Micheyl
University of Minnesota

Pingbo Yin
University of Maryland

Andrew J. Oxenham
University of Minnesota

Shihab A. Shamma
University of Maryland

An important aspect of the analysis of auditory “scenes” relates to the perceptual organization of sound sequences into auditory “streams.” In this study, we adapted two auditory perception tasks, used in recent human psychophysical studies, to obtain behavioral measures of auditory streaming in ferrets (*Mustela putorius*). One task involved the detection of shifts in the frequency of tones within an alternating tone sequence. The other task involved the detection of a stream of regularly repeating target tones embedded within a randomly varying multitone background. In both tasks, performance was measured as a function of various stimulus parameters, which previous psychophysical studies in humans have shown to influence auditory streaming. Ferret performance in the two tasks was found to vary as a function of these parameters in a way that is qualitatively consistent with the human data. These results suggest that auditory streaming occurs in ferrets, and that the two tasks described here may provide a valuable tool in future behavioral and neurophysiological studies of the phenomenon.

Keywords: ferrets, auditory perception, streaming, frequency discrimination, detection

Humans and many other animal species are faced with the problem that the environments they inhabit often contain multiple sound sources. The sounds emanating from these sources mingle before reaching the listener’s ears, resulting in potentially complex acoustic “scenes.” The listener’s brain must analyze these complex acoustic scenes in order to detect, identify, and track sounds of interest or importance, such as those coming from a mate, predator, or prey. This is known among auditory researchers as the “cocktail party” problem (Cherry, 1953) or, more generally, the “auditory scene analysis” problem (Bregman, 1990). One important aspect of the auditory system’s solution to this problem relates to the formation of auditory “streams.” An *auditory stream* refers to sound elements, or groups of sounds, that are usually associated with an individual sound source and are perceived by the listener

as a coherent entity. The sound of an oboe in the orchestra, a conspecific song in a bird chorus, the voice of a speaker in a crowd, and the light footfalls of a predator in the savanna are all examples of auditory streams. The “auditory streaming” phenomenon can be demonstrated using sounds with very simple spectral and temporal characteristics, namely, sequences of tones that alternate between two frequencies (A and B) in a repeating ABAB or ABA–ABA pattern, where A and B denote tones of (usually) different frequencies and the hyphen represents a silent gap. Such sequences have been found to evoke two dramatically different percepts, depending on spectral and temporal stimulus parameters (Bregman & Campbell, 1971; Miller & Heise, 1950; van Noorden, 1975). When the tones are close in frequency, most listeners report hearing a single coherent stream of tones with an alternating pitch; this percept is referred to as *stream integration*. In contrast, when the tones are more widely spaced in frequency and occur in relatively quick succession, the stimulus sequence “splits” perceptually into two streams, as if produced by two separate sound sources; this is referred to as *stream segregation* (an audio demonstration of the phenomenon can be found on <http://www.tc.umn.edu/~cmicheyl/demos.html>). The formation of auditory streams has been the object of a large number of psychophysical studies over the past 50 years (for reviews, see Bregman, 1990; Carlyon & Gockel, 2008; Moore & Gockel, 2002). The neural basis of the phenomenon has also attracted considerable attention, inspiring studies with approaches ranging from single or multiunit recordings in macaques (Fishman, Reser, Arezzo, & Steinschneider, 2001; Micheyl, Tian, Carlyon, & Rauschecker, 2005), bats (Kanwal, Medvedev, & Micheyl, 2003), birds (Bee & Klump,

Ling Ma, Neural Systems Laboratory, Department of Bioengineering, University of Maryland; Christophe Micheyl and Andrew J. Oxenham, Auditory Perception and Cognition Laboratory, Department of Psychology, University of Minnesota; Pingbo Yin, Neural Systems Laboratory, Institute for Systems Research, University of Maryland; Shihab A. Shamma, Neural Systems Laboratory, Department of Electrical and Computer Engineering, Institute for Systems Research, University of Maryland.

This work was supported by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders, Grant R01 DC 007657.

Correspondence concerning this article should be addressed to Ling Ma, Department of Bioengineering, Room 2241, Jeong H. Kim Engineering Building (Bldg. #225), University of Maryland, College Park, MD 20742. E-mail: lma@umd.edu

2004; Itatani & Klump, 2009), guinea pigs (Pressnitzer, Sayles, Micheyl, & Winter, 2008), and ferrets (Elhilali, Ma, Micheyl, Oxenham, & Shamma, 2009) to electro- or magnetoencephalography and functional MRI in humans (e.g., Gutschalk, Oxenham, Micheyl, Wilson, & Melcher, 2007; Gutschalk et al., 2005; Snyder, Alain, & Picton, 2006; Sussman, Ritter, & Vaughan, 1999; Wilson, Melcher, Micheyl, Gutschalk, & Oxenham, 2007).

Although there exists a substantial body of experimental data on auditory streaming in humans, and neuroscientists are starting to explore the neural basis of this phenomenon in both humans and nonhuman animals, the evidence for auditory streaming in animals remains limited (for recent reviews, see Bee & Micheyl, 2008; Fay, 2008). Measuring auditory streaming in nonhuman species is not as easy as measuring it in humans, who can be asked directly what they perceive. This may explain why behavioral studies of auditory streaming in animals remain relatively few and far between. The earliest such study was performed by Hulse, MacDougall-Shackleton, and Wisniewski (1997). In this study, starlings were trained to discriminate 10-s excerpts of conspecific birdsongs, and subsequently tested for generalization with mixtures of two simultaneous birdsongs (conspecific plus heterospecific or conspecific plus natural noises or chorus). Performance with two simultaneous birdsongs was still relatively high (about 85% correct), and starlings readily generalized to mixtures of familiar songs in unfamiliar backgrounds, suggesting that they were able to segregate the target song from the background. This result was interpreted as evidence for auditory stream segregation in starlings. Further evidence that starlings experience stream segregation was obtained in an elegant study by MacDougall-Shackleton, Hulse, Gentner, and White (1998), using stimuli and a task that were perhaps less ecological but more comparable to those used in human psychoacoustical studies. In this study, starlings were conditioned using sequences of constant-frequency tones arranged temporally into triplets (i.e., groups of three tones separated by a silent gap), which, in human listeners, yield a “galloping” percept (van Noorden, 1975). The birds were later tested for generalization to sequences of triplets in which the middle tone had a different frequency from the two outer tones. The results showed decreasing generalization with increasing frequency separation between the middle and outer tones. This effect is consistent with the results of psychoacoustical studies of auditory streaming in humans, which indicate that as frequency separation increases, the middle and outer tones are increasingly likely to be heard as separate streams, and that when this happens, the galloping rhythm is no longer heard.

Fay (1998) provided evidence that auditory streaming is also present in a species in which the phenomenon is perhaps less expected to play an important role. He conditioned goldfish on a mixture of two trains of acoustic pulses, which differed in spectral content (low vs. high center frequencies) and repetition rate (19 pulses/s for the low-frequency pulses and 85 pulses/s for the higher frequency pulses). Later, the fish were tested for generalization using single (low or high center frequency) pulse trains over a range of rates (between 19 and 85 pulses/s). In the group tested with the low center frequency pulses, generalization decreased toward higher pulse rates; in the group tested with the high center frequency pulses, the converse was observed. This pattern of results is consistent with the hypothesis that, during the conditioning phase, the fish heard the low-frequency and high-frequency

tones as separate streams. In a subsequent study (Fay, 2000), the fish were conditioned using trains of pulses alternating between two center frequencies (a high frequency, 625 Hz, and a lower frequency drawn between 240 and 500 Hz) at an overall rate of 40 pulses/s (20 pulses/s at a given frequency). Subsequently, the fish were tested for generalization using only 625-Hz pulses presented at various rates (from 20 to 80 Hz). Generalization to rates near 20 pulses/s was stronger in the group conditioned with pulses at 240 and 625 Hz than in the other training groups. This outcome is consistent with the hypothesis that the mixture with the widest frequency spacing was heard as two separate streams, whereas the other mixtures were less easily segregated, because of the smaller frequency separation between the alternating tones.

Some evidence that auditory streaming is also present in species that are more closely related to humans has been provided by Izumi (2002). To test for auditory streaming in Japanese monkeys, Izumi used an approach inspired by psychoacoustical studies in which listeners had to recognize familiar melodies, the notes of which were played in alternation—the interleaved melodies paradigm (e.g., Dowling, 1968, 1973). One of the main findings of these studies was that listeners’ performance in the identification of such “interleaved melodies” usually increased with the mean frequency (or pitch) separation between the two melodies. This effect, which is well known to music composers, can be explained on the basis of the observation that frequency separation facilitates stream segregation (van Noorden, 1977). Izumi replaced the melodies by short sequences of tones, which were either rising or falling in frequency. The monkeys were first trained to discriminate such sequences presented in isolation. Then, the sequences were interleaved temporally with a sequence of unrelated tones. Performance in this interleaved condition improved as the mean frequency separation between the tones in the two interleaved sequences increased, consistent with the results of interleaved melodies studies in humans, which have been interpreted in terms of stream segregation (Bey & McAdams, 2002, 2003; Dowling, 1968, 1973).

The behavioral findings reviewed above suggest that both auditory streaming and frequency-selective attention are relatively basic auditory abilities, shared by various animal species. The current experiments were performed in the context of a broader research project, the ultimate goal of which is to investigate neural correlates of auditory streaming and selective attention in the auditory and prefrontal cortices of behaving ferrets. One of the major subgoals of this project involves devising behavioral tasks that can be used to manipulate—and at the same time, measure—auditory streaming and selective attention in ferrets. In particular, we were looking for behavioral tasks that could be used to encourage stream segregation and frequency-selective attention. The conditioning-generalization paradigms used by Fay (1998, 2008) and MacDougall-Shackleton et al. (1998) were “neutral”; the ferrets were not rewarded specifically for segregating (or for integrating) streams. From this point of view, these studies are comparable to human studies in which listeners are simply asked to report whether they hear a stimulus sequence as one stream or two streams and not encouraged by instructions, or task demands, to try to hear the sequence in a specific way (see van Noorden, 1975). Here, we were specifically interested in manipulating the attentional and perceptual state of the ferret to later measure the influence of such a manipulation on neural responses, compared

with passive or “neutral” listening to the same stimuli. A second important constraint in the design of our experiments stemmed from our long-term objective of characterizing the influence of behavior in the task on neural responses, as measured using, for example, “classic” frequency-response curves or spectrotemporal receptive fields. We reasoned that this objective, and the interpretation of the results in terms of sequential streaming and frequency-selective attention, would be facilitated by the use of stimuli with relatively simple and tightly controlled spectrotemporal characteristics, in contrast to the use of natural sounds (e.g., birdsongs) used by Hulse et al. (1997) and Wisniewski and Hulse (1997).

These considerations led us to adapt stimuli and tasks from two previous psychophysical studies, both of which involved performance-based measures of auditory streaming and selective attention. The stimuli and task that we used in our first experiment were adapted from an experiment in humans by Micheyl, Carlyon, Cusack, and Moore (2005), who found that thresholds for the discrimination of changes in the frequency of the last B tone in an ABA sequence were influenced by stimulus parameters known to control the stream segregation of pure tone sequences. Specifically, they found that thresholds increased (i.e., worsened) as the frequency separation between the A and B tones (ΔF_{AB}) decreased, and that they decreased (i.e., improved) as the tone presentation rate and overall length of the sequence increased. Given that large A–B separations, fast tone presentation rates, and long sequence lengths are all facilitating factors of stream segregation, this pattern of results is consistent with a beneficial influence of stream segregation on the ability to discriminate changes in the frequency of the B tones. A likely explanation for the influence of stream segregation on frequency-discrimination performance in this experiment is in terms of selective attention. When the A and B tones are heard as separate streams, attention can more easily be focused selectively on the B tones. This limits potential interference from the A tones in the processing of the pitch of the B tones (Micheyl & Carlyon, 1998). In particular, when the A and B tones are heard within a single stream, the pitch “jumps” between the A and B tones may interfere with the detection of the (usually) smaller frequency shift in the last B tone (Watson, Kelly, & Wroton, 1976); when the A and B tones are perceived as separate streams, the pitch jumps are no longer heard, and listeners can focus solely on the B tones. In Experiment 1, we adapted the stimuli and task used by Micheyl, Carlyon, et al. (2005) to measure stream segregation in ferrets. On the basis of our experience training ferrets in auditory-perception tasks (e.g., Atiani, Elhilali, David, Fritz, & Shamma, 2009; Fritz, Elhilali, David, & Shamma, 2007; Fritz, Shamma, Elhilali, & Klein, 2003; Kalluri, Depireux, & Shamma, 2008; Yin, Mishkin, Sutter, & Fritz, 2008), these animals can detect frequency differences, but they have difficulties making low- versus high-pitch judgments—an observation confirmed by recent results (Walker, Schnupp, Hart-Schnupp, King, & Bizley, 2009). Therefore, we changed the task from pitch-direction identification to simple pitch-change detection. Under the hypothesis that ferrets experience stream segregation, we predicted that their thresholds for the detection of a change in the frequency of B tones in ABAB . . . sequences should decrease (improve) with increasing A–B frequency separations.

The stimuli and task in our second experiment were inspired by studies of “informational masking” in humans. The expression

informational masking refers to masking effects that cannot be explained primarily in terms of peripheral interactions, and that do not depend critically on the energy ratio of the target and masker (for a recent review, see Kidd, Mason, Richards, Gallun, & Durlach, 2008). Informational-masking effects are especially large when the spectral characteristics of the masker vary randomly across presentations and the target and masker are easily confusable. However, these effects can be dramatically reduced by stimulus manipulations that promote the perceptual segregation of the signal and masker. For instance, the detection threshold for a target tone of fixed frequency can be elevated by 40 dB or more if the tone is presented synchronously with a multitone masker, the frequencies of which are drawn randomly on each trial (Neff & Green, 1987); this is the case even if the masker frequencies are not allowed to fall within the same critical band as the target. However, if the constant-frequency target tones repeat at a rate sufficiently fast for them to form a stream, which separates (“pops out”) from the randomly varying masker tones, they become easily detectable again (Kidd, Mason, Deliwala, Woods, & Colburn, 1994; Kidd, Mason, & Richards, 2003; Micheyl, Shamma, & Oxenham, 2007). In general, performance in the detection of the target tones improves as the width of the protected region and the repetition rate of the target tones increase (Kidd et al., 1994, 2003; Micheyl, Shamma, et al., 2007). Under the hypothesis that ferrets experience the effect, we predicted that these two trends would be observed in Experiment 2.

Method

Subjects

Two female ferrets (*Mustela putorius*) obtained from Marshall Farms (North Rose, NY) were used in these experiments. Both of them were young adults (about 2 years old), each about 780 g in weight. The ferrets were housed in pairs in a cage in facilities accredited by the American Association for Laboratory Animal Care and were maintained in a 12-hr artificial light cycle. They were only brought to the Neural Systems Lab during training and testing sessions. The ferrets had free access to dry food all the time, but water access was restricted to water reward during task performance (usually 5 days/week); otherwise, they had continuous access to water. Animal condition was carefully monitored on a daily basis, and weight was maintained above 80% of ad libitum weight. The care and use of animals in this study was consistent with National Institutes of Health guidelines. All procedures for behavioral testing of ferrets were approved by the institutional animal care and use committee of the University of Maryland, College Park.

Experimental Design

Two domestic ferrets were trained to perform two different tasks, performance in which has been previously found to be related to stream segregation in humans. The stimuli and behavioral paradigms are detailed below.

Experiment 1: Detection of a frequency shift within a stream. On each trial, a sequence of pure tones alternating between two frequencies (A and B) in a repeating ABAB . . . pattern, as illustrated in Figure 1, was presented. On 78% of the trials,

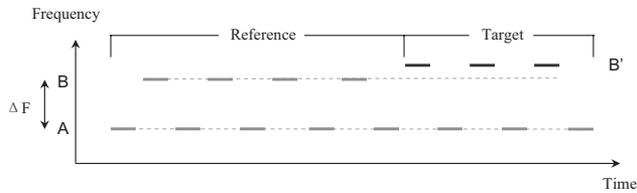


Figure 1. Spectrotemporal structure of the stimuli used in Experiment 1. This shows an example stimulus sequence on a trial containing target tones. The gray bars represent tones. The “reference” portion of the stimulus consisted of A and B tones alternating between two frequencies, A and B. The “target” tones had a higher frequency than the B tones, denoted as B’. Two stimulus parameters were varied: the frequency separation between the A and B tones, ΔF_{AB} , and the frequency separation between the B and B’ tones, $\Delta F_{BB'}$ (see text for additional details).

the B frequency changed to a higher frequency, B’. On remaining 22% of trials, the B frequency did not change; these trials are hereafter referred to as “shams.” The task of the ferret was to detect the frequency change, when that change was present. Given the findings of Micheyl, Carlyon, et al. (2005), we predicted that if ferrets experience streaming, then their performance in the detection of a change in the frequency of the B tones (from B to B’) should be higher when the A–B frequency separation is large (promoting stream segregation between A and B) than when it is small (making it difficult or impossible to hear the B tones stream as a separate entity).

The ferret was trained to lick a waterspout during the “reference” sequence of AB tones, and to stop licking on detecting the change from B to B’. Initially, the ferret was trained to detect a relatively large difference in the frequency of the B tones (from B to B’) in the absence of any A tones. Once performance in this condition reached an asymptote, the A tones with a frequency 19 semitones (ST) below that of the B tones were introduced. Initially, the level of the A tones was set 50 dB below that of the B tones, which were always presented at 70 dB sound pressure level (SPL). The level of the A tones was then raised progressively, over the course of several weeks, depending on the ferret’s performance. Eventually, the ferret was able to perform the task relatively well with the A tones at the same level as the B tones. At that point, data collection began. Overall, training took about 7 months. The actual test phase lasted 12 days (4 days for each ΔF_{AB}). During this test phase, the ferret performed at least 70 trials each day.

Detailed stimulus parameters were as follows. Each tone was 75 ms long, including 5-ms onset and offset cosine ramps. Consecutive tones were separated by a silent gap of 50 ms. Therefore, the repetition rate of the elementary AB pattern was equal to 4 Hz. The frequency of the B tone was fixed at 1500 Hz. The frequency of the A tone was constant within a block of trials, but varied across testing days in order to produce different frequency separations between the A and B tones, denoted here as ΔF_{AB} . Three ΔF_{AB} s were tested: small (6 ST), medium (9 ST), and large (12 ST, i.e., 1 octave). Although the smallest ΔF_{AB} (6 ST) used here is relatively large, and would be considered “intermediate” in humans, we found while training the ferret that the ferret could not do the task with smaller A–B separations; consequently, we decided to use larger separations.

The frequency separation between the B and B’ tones, denoted as $\Delta F_{BB'}$, varied randomly across trials within a test session in order to produce different levels of task difficulty, yielding different levels of performance. In all conditions in which the A tones were present, five values of $\Delta F_{BB'}$ were tested: 4%, 12%, 20%, 28%, and 36%. A larger number of $\Delta F_{BB'}$ s were tested during the initial training phase, in which the A tones were absent: 6%, 8%, 10%, 12%, 14%, 16%, 18%, 21%, and 27%.

The total trial duration varied randomly from 1.875 to 7.875 s, depending on the number of “reference” pairs (AB) presented before the introduction of the B’ tones. This number was selected randomly between four and 28 on each trial. The number of “target” pairs (AB’) was fixed at three. If the ferret stopped licking within 850 ms after the introduction of the first B’ tone in the sequence, this was counted as a hit; otherwise, the trial was categorized as a miss. If a stop-lick response was produced on a sham trial, it was counted as a false alarm; otherwise, the trial was counted as a correct rejection. False alarms had no consequence. Following the 850 ms after the introduction of the first B’ tone in the sequence, the spout became electrified, and hence the ferret received a mild shock if it continued licking afterward, and the trial was labeled a miss. Each trial included silent periods of 400 ms prestimulus, and 600 ms poststimulus.

Experiment 2: Detection of regularly repeating target tones in a random multitone background. This experiment was inspired by psychoacoustic experiments on informational masking (Kidd et al., 1994, 2003; Micheyl, Shamma, et al., 2007). An example spectrogram of the stimuli used in this experiment is shown in Figure 2. On each trial, a sequence consisting of multiple tone pips with random frequencies and random onset times (“maskers”) was presented. At some point in this random sequence, a regularly repeating sequence of constant-frequency tones (“targets”) was introduced. The task of the ferret was to detect the target sequence amid the randomly varying masker tones. The ferret was trained to withhold licking until the target was introduced and to start licking on detecting the target. If a lick response occurred within 150 to 1,050 ms after the onset of the first target

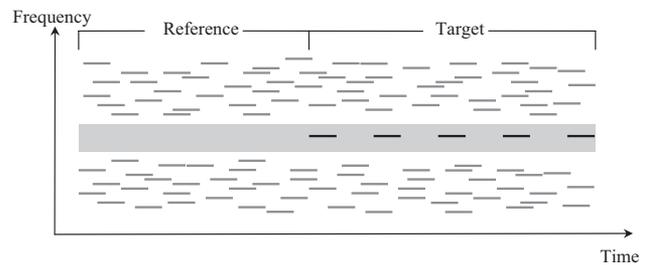


Figure 2. Schematic spectrogram of an example stimulus presented on a trial in Experiment 2. During the “reference” portion of the stimulus, only masker tones (gray bars) with random frequencies and onsets times were presented. During the “target” part, target tones (dark bars) repeating regularly at a constant frequency were introduced. The gray area around the target represents the protected zone (PZ), within which masker tones were not allowed to fall (see text for additional details).

tone,¹ it was counted as a hit and reinforced with 1/3 ml of water. Misses had no consequence. In this experiment, there were no “sham” trials; the target tones were presented on all trials. However, the start time of the target sequence varied randomly between 720 and 2,160 ms after the onset of the masker sequence. When the ferret produced a lick response before the onset of the target sequence, this was counted as a false alarm, the trial was aborted, and was followed by a short timeout.

Note that response contingencies are reversed here compared with Experiment 1, in that positive reinforcement is used (as opposed to negative reinforcement in Experiment 1). This control is specifically important for physiological experiments because auditory cortical responses and adaptation can depend critically on whether the “target stimuli” in the tasks were aversively or positively reinforced (David, Fritz, & Shamma, 2008). Consequently, we felt it was important to demonstrate in this study that ferrets could perform both forms of the streaming tasks so as to facilitate recordings from their auditory areas during such behaviors.

The stimulus details were as follows. Each tone pip (target or masker) was 70 ms long, including 5-ms onset and offset ramps. On each trial, five target tones were presented. Consecutive target tones were separated by a silent gap of 110 ms, yielding a repetition rate of about 5.6 Hz. Trial length varied randomly between 1.62 and 3.06 s across trials. These durations include the variable-length “reference” sequence (0.72 to 2.16 s) plus the fixed-duration “target” sequence. The masker tones occurred at an average rate of 89 tones/s. The masker tones were generated as follows. First, eight different masker-tone frequencies were selected at random for every 90 ms; then, the masker tones were shifted pseudorandomly in time in such a way that they were not synchronous with the target, except by chance. The masker tone frequencies were drawn at random from a fixed list of values spaced 1 ST (i.e., approximately 6%) apart, excluding a “protected zone” (PZ) around the frequency of the target tones. The half-width of the PZ determines the minimum allowed frequency separation between the target and the closest masker component on either side. Three half-widths were tested: small (6 ST), medium (10 ST), and large (14 ST). Masker frequencies were selected from within a 2-octave range on both sides of the PZ. The target frequency was roved daily from 3.5 to 4.1 kHz. PZ was varied randomly from day to day, and target intensity and trial lengths varied within a session. The masker tones were presented at 50 dB SPL (each). Target-tone levels of -4, 0, +4, +8, and +12 dB relative to the level of the masker tones were tested. These values were chosen to produce different hit rates, allowing a psychometric function to be traced.

In a separate set of conditions we studied the influence of target repetition rate on performance. Three rates were tested: 3.7, 5.6, and 11.1 target tones/s. These different rates were produced by varying the duration of the silent interval between consecutive target tones (from 20 to 200 ms).

For this task, the training phase spanned 14 months, including sporadic intermissions of a few weeks during which the ferret did not behave. Typically, the ferret was trained 5 days per week. Initially, the ferret was trained with a very wide PZ (16 ST). The width of the PZ was progressively reduced, week after week (or sometimes, month after month). When the ferret’s performance reached an asymptote, training stopped and testing proper started. During the test phase, the ferret performed a total of 24 sessions

using the 6- and 10-ST PZ half-widths (12 sessions for each of these two half-widths), at a pace of one session per day. For the 14-ST PZ half-width, the ferret performed 11 daily sessions.

Apparatus

Ferrets were tested in a custom-designed cage (8 × 15 × 9 in.) mounted inside a Sonex-foam lined and single-walled soundproof booth (Industrial Acoustics Corporation, Lodi, NJ). The stimuli were generated using Matlab (The MathWorks, Natick, MA). They were sampled at 40 kHz, played out at 16-bit resolution (NI-DAQ), amplified (Yamaha A520), and finally delivered through a speaker (Manger) mounted in the front of the cage, at approximately the same height above the testing cage as the metal spout that delivered the water reward. Lick responses were registered by a custom “touch” circuit.

Data Analysis

The behavioral data were analyzed using techniques from signal detection theory (Green & Swets, 1966). In particular, the responses of the ferrets were used to compute the area under the receiver operating characteristic (ROC). The area under the ROC provides an unbiased measure of performance, with 0.5 reflecting chance performance and 1 reflecting perfect performance (Green & Swets, 1966; Hanley & McNeil, 1982). In Experiment 2, ROCs were derived by varying the duration of the response window, defined as the time interval within which a start-lick event was registered. The rationale for this analysis is that longer response times correspond to more liberal placements of the internal decision criterion (Luce, 1986; Yin, Fritz, & Shamma, 2009). The range of possible occurrence times of the first target tone in the stimulus sequence was from 0.72 to 2.16 s. On trials on which the target tones occurred relatively early (i.e., between 720 ms and 1,620 ms), the response window started 150 ms after the onset of the first target, and a lick event occurring within the response window was counted as a hit. On trials on which the targets occurred relatively late (i.e., between 1.62 s and 2.16 s), the response window started 150 ms after the time at which the first target should have started had the target tones been early, and a lick occurring within the response window was counted as a false alarm. The duration of the response window was varied from 0 to 900 ms in 50-ms increments. Areas under the resulting ROCs were approximated using trapezoids. The advantage of this method is that it does not require specific assumptions regarding the underlying distributions.

In Experiment 1, we had to employ a different data-analysis technique to accommodate the different experimental design and different response contingencies. First, we measured the duration for which the ferret had made contact with the waterspout within a 400-ms “reference” epoch, which just preceded the introduction of the target tones. If this duration was less than 20 ms (5% of the reference-epoch duration), we considered this as an indication that the ferret was not ready for task performance, and data from the current trial were not included in subsequent analyses. In contrast,

¹ These parameters were chosen on the basis of the consideration that the quickest reaction time in ferrets is approximately 150 ms and that the target was 900 ms long.

trials on which the ferret licked the waterspout for at least 20 ms during the reference period were retained for further analysis. These trials were divided into two groups, depending on whether the ferret had come into contact with the waterspout during the time period within which shocks could be delivered if the ferret had not stopped licking. This “shock period” started 850 ms after the onset of the first target tone and lasted for 400 ms. If the ferret had made contact with the spout during the shock period, the trial was categorized as a “miss” or as a “correct rejection,” depending on whether or not target tones were presented on that trial. If the ferret had not made contact with the spout during the shock period, the trial was categorized as a “hit” or as a “false alarm,” depending on whether or not target tones were presented on that trial.

As a result, a single pair of hit and false-alarm rates was available for each condition. When the ROC contains a single point, approximation using trapezoids can lead to severe underestimation of the ROC area. Accordingly, in this experiment, we had to resort to parametric assumptions. Specifically, the ROC area was computed as the surface under a binormal curve passing through three points: (0,0), (1,1), and the point defined by the measured pair of hit and false-alarm rates.

Thresholds in both experiments were estimated by fitting the ROC area data as a function of the relevant stimulus parameter ($\Delta F_{BB'}$ in % for Experiment 1, relative target level in dB for Experiment 2) with a sigmoid function defined by the following equation,

$$Pc(\Delta) = 0.5 + \rho[1 + e^{-(\Delta-\theta)/\sigma}]^{-1} \quad (1)$$

where Pc is the proportion of correct responses; Δ denotes the value of the stimulus parameter ($\Delta F_{BB'}$ in % for Experiment 1, relative target level in dB for Experiment 2); ρ is the dynamic range of the psychometric function, which corresponds to the

difference between the guessing rate (0.5) and the lapse (i.e., miss) rate, λ ; θ is the threshold, defined as the frequency difference (in Experiment 1) or target level (in Experiment 2) at which $Pc = (0.5 + \lambda)/2$, that is, the midpoint between the guessing rate and the lapse rate; σ is a standard deviation parameter, which corresponds to the reciprocal of the slope of the psychometric function. For Experiment 1, a data point corresponding to $Pc(0) = 0.5$ was introduced to reflect the fact that when $\Delta F_{BB'}$ was equal to 0% (i.e., the B and B' tones had the same frequency, and there was no signal for the ferret to detect), performance should be at chance. In addition, the contribution of each mean data point to the overall fit was weighted by the inverse of the variance around the mean, and constraints were placed on the slope parameter to prevent unrealistically steep psychometric functions in the 9-ST ΔF_{AB} condition. For each condition, we computed 95% confidence intervals (CIs) around the threshold estimates using a statistical resampling-with-replacement technique (bootstrap with 1,000 draws) assuming binomial dispersion (Efron & Tibshirani, 1993).

Experiment 1

Results

The results of Experiment 1 are illustrated in Figure 3. Figure 3a shows an example of binormal ROC, determined as explained in the Data Analysis section. This ROC was obtained based on a pair of hit and false-alarm rates measured using a 12-ST ΔF_{AB} and a 36% $\Delta F_{BB'}$. In this example, the area under the ROC, which is shown in gray, was equal to 0.87, indicating good performance.

The ROC area was computed in a similar way for all other ΔF_{AB} and $\Delta F_{BB'}$ conditions. The resulting set of ROC areas is shown as data points in Figure 3b. These data were fitted using sigmoid

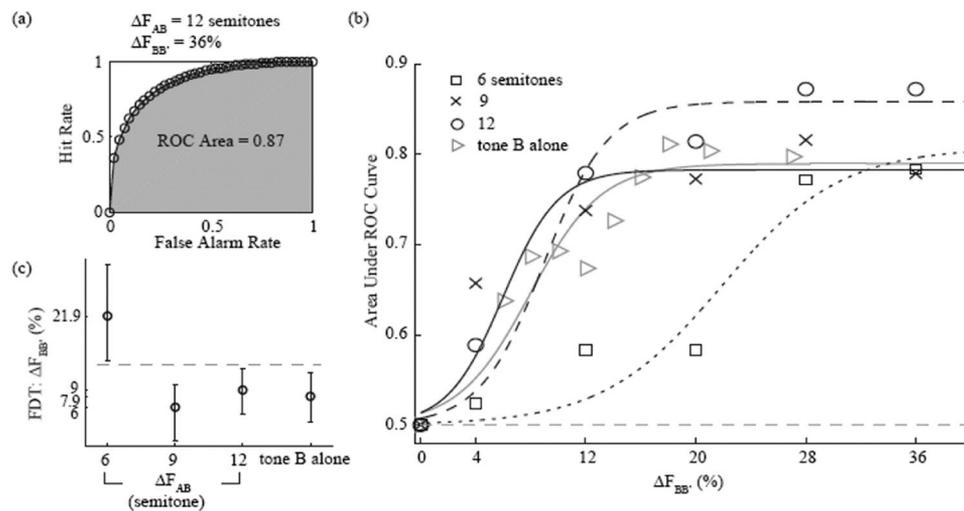


Figure 3. Performance measures in Experiment 1. (a) Example receiver operating characteristic (ROC) curve obtained using the technique described in the main text. This curve was computed based on data obtained at 12-ST ΔF_{AB} and a 36% $\Delta F_{BB'}$. (b) ROC area as a function of $\Delta F_{BB'}$ and best-fitting psychometric functions. ROC areas are shown as symbols; the best-fitting psychometric functions as lines. The different symbols and line styles indicate different ΔF_{AB} conditions (6, 9, and 12 ST). (c) Frequency discrimination thresholds (FDTs) estimated based on the psychometric functions shown in Panel b. The error bars indicate 95% CIs (bootstrap) around the mean FDTs. The dashed line indicates nonoverlapped over 95% CIs.

psychometric functions for each ΔF_{AB} condition separately, and the best-fitting functions were used to estimate a threshold (defined as the $\Delta F_{BB'}$ value corresponding to the midpoint between chance performance and the upper asymptote) for each ΔF_{AB} condition.

The resulting threshold estimates are plotted in Figure 3c, along with the 95% CIs (computed using bootstrap). It can be seen that thresholds were highest for the lowest ΔF_{AB} tested (6 ST) and substantially smaller ($p < .05$) for larger separations (9 and 12 ST). In fact, the thresholds measured with separations of 9 and 12 ST were not significantly larger ($p > .05$) than those measured in the baseline condition, which did not contain any A tones. Comparing the thresholds for each ΔF_{AB} condition with those for the baseline condition, the effect sizes (Glass's Δ) were 6.09 with 95% CI [3.91, 4.21] (6 ST vs. baseline condition), -0.83 with 95% CI $[-0.92, -0.74]$ (9 ST vs. baseline condition), and 0.48 with 95% CI [0.39, 0.57] (12 ST vs. baseline condition). The only apparent difference in results between the 9 and 12 ST conditions was that the asymptotic proportion of correct responses was somewhat higher in the latter condition (around 0.9) than in the other conditions. We have no explanation for this marginal observation. Although asymptotic proportions of correct responses below 1 are usually regarded as indicative of attentional "lapses" (Klein, 2001), there was no a priori reason to expect a lower lapse rate in the 12-ST condition than in the other conditions.

Discussion

The pattern of results illustrated in Figure 3c is qualitatively consistent with the psychophysical data obtained by Micheyl, Carlyon, et al. (2005) in human listeners. These authors found that frequency discrimination thresholds for target B tones inside repeating ABA sequences comparable to those used here improved as the frequency separation between the A and B tones increased. They explained this effect, and other effects of stimulus parameters (including rate and sequence length), in terms of stream segregation and selective attention. Specifically, they suggested that stimulus manipulations that promoted the perceptual segregation of the A and B tones into separate streams made it easier for listeners to attend selectively to the B tones and to ignore the irrelevant but potentially interfering pitch information conveyed by the A tones.

The presented finding of smaller thresholds for the detection of frequency changes in the B tones for larger A–B frequency separations is consistent with the hypothesis that stream segregation facilitates selective attention in ferrets and leads consequently to improved detection thresholds. However, this interpretation can be further supported by other factors that are known to modulate streaming such as tone presentation rate and sequence length. Unlike our previous experiments in humans (Micheyl, Carlyon, et al., 2005), we did not manipulate these parameters in the current experiments. Hence, it will be interesting to explore in the future the dependence of detection thresholds at a fixed ΔF_{AB} on presentation rate and sequence length, and whether this dependence is consistent with a beneficial influence of stream segregation.

Although the ferret data show trends that are qualitatively similar to those observed in psychophysical studies with human listeners, there are also several important differences between the ferret and human data. First, the frequency-discrimination thresholds that were measured in the ferret are substantially larger than

those that have been measured in humans typically. In traditional 2I-2AFC experiments, highly trained human listeners can achieve frequency-discrimination thresholds of 0.1–0.2% (roughly 1.5–3 Hz) at 1.5 kHz (e.g., Moore, 1973; Wier, Jesteadt, & Green, 1977; Micheyl, Delhommeau, Perrot, & Oxenham, 2006). In an ABA context, Micheyl, Carlyon, et al. (2005) measured thresholds of less than 1% using test frequencies in the vicinity of 1 kHz—at least, when the A–B frequency separation was sufficiently large for listeners to hear the A and B tones as separate streams. These values are substantially smaller than the 8% or more average thresholds measured here under comparable (though not identical) stimulus conditions, using different procedures. On the other hand, the thresholds that were measured in this study compare well with those obtained in an earlier study (Sinnott, Brown, & Brown, 1992) in the gerbil (9% at 1 kHz, 10% at 2 kHz). These thresholds are also not very far off from those measured in the rat, guinea pig, or chinchilla, where frequency discrimination thresholds ranging from 2% to 7% on average (with substantial intersubject variability) have been obtained using test frequencies between 2 and 5 kHz (Heffner, Heffner, & Masterton, 1971; Kelly, 1970; Sloan, Dodd, & Rennaker, 2009; Nelson & Kriester, 1978; Syka, Rybalko, Brožek, & Jilek, 1996; Talwar & Gerstein, 1998, 1999). Thus, even though ferrets are not rodents (they are carnivores, most closely related to weasels), their frequency discrimination thresholds appear to be similar to those of rodents, which are generally much larger than those measured in humans. In all other species in which frequency discrimination thresholds have been measured, to our knowledge, the results indicate that these thresholds are not quite as low as those measured in highly trained human listeners—although for cats, they can be as low as 0.85% at 1 kHz and 0.75% at 2 kHz (Elliott, Stein, & Harrison, 1960); for dog, roughly 0.9% at 1 and 2 kHz (Baru, 1967). This appears to be the case even for monkeys, in which frequency discrimination thresholds ranging between 1.6% and about 4% have been reported (Prosen, Moody, Sommers, & Stebbins, 1990; Sinnott & Brown, 1993; Sinnott, Petersen, & Hopp, 1985). It has been suggested that small frequency-discrimination thresholds below 4–5 kHz in humans reflect the use of temporal (i.e., phase-locking) information (Moore, 1973; Sek & Moore, 1995; Micheyl, Moore, & Carlyon, 1998), whereas monkeys and other animals may rely more heavily on tonotopic (i.e., rate-place) information (Prosen et al., 1990; Sinnott & Brown, 1993; Sinnott et al., 1985).

A second noteworthy difference between the ferret results and the human data is that, during the training phase, the ferret was found to be largely unable to perform consistently above chance when the A–B frequency separation was smaller than 6 ST; this is why smaller separations were not included in the design of Experiment 1. In contrast, in humans, thresholds could still be reliably measured for ΔF_{AB} s as small as 1 ST (Micheyl, Carlyon, et al., 2005). A possible explanation of this discrepancy is that, although human listeners almost certainly heard the tones as a single stream in these conditions, they could still perform the task above the chance level by comparing the frequency of the last B tone with that of a temporally adjacent A tone, or by sensing an overall increase or decrease in pitch between the last two triplets. The ferret was perhaps not able to adapt its listening strategy depending on ΔF_{AB} to take advantage of a different cue at small A–B separations than at larger ones. In this context, the observation that ferrets appear to need larger ΔF_{AB} s than humans to

perform reliably in the task could be due to larger frequency separations being needed to induce a percept of stream segregation in ferrets than in humans. For humans, the fission boundary, which corresponds to the smallest frequency separation below which listeners are no longer able to hear two separate streams (van Noorden, 1975), is approximately equal to 0.4 times the equivalent rectangular bandwidth (ERB) of auditory filters (Rose & Moore, 2000, 2005). At 1 kHz, the ERB for normal-hearing listeners is 132 Hz (about 13% of the center frequency; Moore, 2003), yielding a fission boundary of approximately 5% of the center frequency, or slightly less than 1 ST. The data of Micheyl, Carlyon, et al. (2005) indicate that the listeners in that study usually needed ΔF_{AB} 's larger than 1 ST to be able to discriminate changes in the frequency of the B tone relatively accurately. To the extent that the fission boundary for stream segregation scales with auditory-filter bandwidths across species, one should expect this boundary to be larger in ferrets than humans. Even though, to our knowledge, auditory-filter bandwidths have not been measured in ferrets, the various other mammalian species in which they have been measured behaviorally, which include the cat (Pickles, 1979; Nienhuys & Clark, 1979), chinchilla (Seaton & Trahiotis, 1975), and macaque monkey (Gourevitch, 1970; for a review, see Fay, 1988), indicate somewhat larger bandwidths than in humans (Shera, Guinan, & Oxenham, 2002).

Another factor, which may explain why ferrets require larger A–B frequency separations than humans, relates to frequency-selective attention bandwidths. Frequency-selective attention is likely to be critically involved for successful selective processing of changes in frequency in the presence of extraneous tones—in the present case, temporally adjacent (A) tones between the target (B or B') tones. In humans, frequency-selective attention has traditionally been measured using the “probe-signal” method (Greenberg & Larkin, 1968). In these experiments, the listener detects a tone close to its masked threshold in noise, and on a small proportion of randomly selected trials, the signal is presented at another frequency (probe). The results of these experiments reveal that, as distance between the probe frequency and the signal frequency increases, the percentage of correct detections decreases, forming an inverted V-shaped selective-attention curve (Dai, Scharf, & Buus, 1991; Greenberg & Larkin, 1968). The width of the curve provides an indication of the bandwidth of the frequency-selective attention filter in the listener. In humans, this width is closely related to the bandwidth of auditory filters (Moore, Hafter, & Glasberg, 1996). To the extent that a similar relationship exists in ferrets, and that auditory-filter bandwidths are wider in ferrets than in humans, this could explain why ferrets need larger A–B separations to successfully detect changes in the frequency of specific tones in a stimulus sequence that contains other (irrelevant) frequencies and frequency changes.

Experiment 2

Results

Figure 4a illustrates how the hit and false-alarm rates measured in this experiment increased with the duration of the response window. In this particular example, the PZ half-width was 14 ST, and the target tones were 12 dB above the masker tones. Both the hit rate and the false-alarm rate tended to increase with the dura-

tion of the response window. However, the hit rate increased more steeply than the false-alarm rate, indicating that in this condition, the ferret could reliably detect the target tones.

These pairs of hit and false-alarm rates were used to construct the ROC area shown in Figure 4b. In this example, the ROC area was equal to 0.81. ROC areas were computed in this way for each PZ width and target-level condition. The resulting ROC areas are plotted as a function of target level in Figure 4c. As expected, ROC area increased with target level. Overall, performance was lower at the smallest PZ half-width (6 ST) than at larger half-widths (10 and 14 ST). This effect is illustrated in Figure 4d, which shows how thresholds (estimated on the basis of the psychometric function fits as explained in the Data Analysis section) improved as the PZ width increased. Comparing the thresholds among each PZ condition, the effect sizes (Hedges' g) were 4.41 with 95% CI [4.25, 4.57] (6 vs. 10 ST condition), 3.03 with 95% CI [2.90, 3.16] (6 vs. 14 ST condition), and 0.62 with 95% CI [0.53, 0.71] (14 vs. 10 ST condition).

Figure 5 illustrates how the ROC area varied over the time course of the stimulus sequence, from 150 ms after the onset of the first target tone until 150 ms after the offset of the last target tone. The different panels correspond to different target levels (relative to the masker), from low (left) to high (right). Within each panel, the different curves correspond to the different PZ half-widths that were tested. The different data points within each curve correspond to ROC areas based on pairs of hit and false-alarm rates computed using increasing response-window durations (in 50-ms increments). The ROC area generally increased over time following the introduction of the target tones, $F(18, 3040) = 100.17, p < .001, \eta_p^2 = .37$. This effect became more marked as the PZ became wider, $F(36, 3040) = 2.52, p < .001, \eta_p^2 = .03$, and as the level of the target tone was raised from 4 dB below to 12 dB above the masker level, $F(72, 3040) = 1.83, p < .001, \eta_p^2 = .04$.

Figure 6 illustrates the influence of the target repetition rate on detection performance. These data were obtained using a target level 4 dB below the masker level. The different line styles indicate different PZ widths. As can be seen, the ROC area was larger at the largest (14 ST) PZ half-width than at the two smaller widths, $F(2, 27) = 8.58, p < .01, \eta_p^2 = .39$ and it increased with target repetition rate, $F(2, 27) = 4.35, p < .05, \eta_p^2 = .25$.

Discussion

The effects illustrated in Figure 4 are qualitatively consistent with earlier results in the human psychoacoustics literature, which show improvements in thresholds (Richards & Tang, 2006) or d' (Micheyl, Shamma, et al., 2007) in a task involving the detection of regularly repeating target tones among randomly varying masker tones as the width of the PZ around the target tones increases. However, there are some noteworthy differences between the ferret data and human data. For instance, Richards and Tang (2006) observed threshold improvements of 10 dB or more as they increased the half-width of the PZ around 1 kHz target tones from 20 to 350 Hz. These values correspond to half-widths of approximately 0.34 and 5.2 ST, respectively. These values are substantially smaller than those used in the current study (6–14 ST), indicating that the ferret needed substantially larger PZ widths than human listeners to detect the target tones.

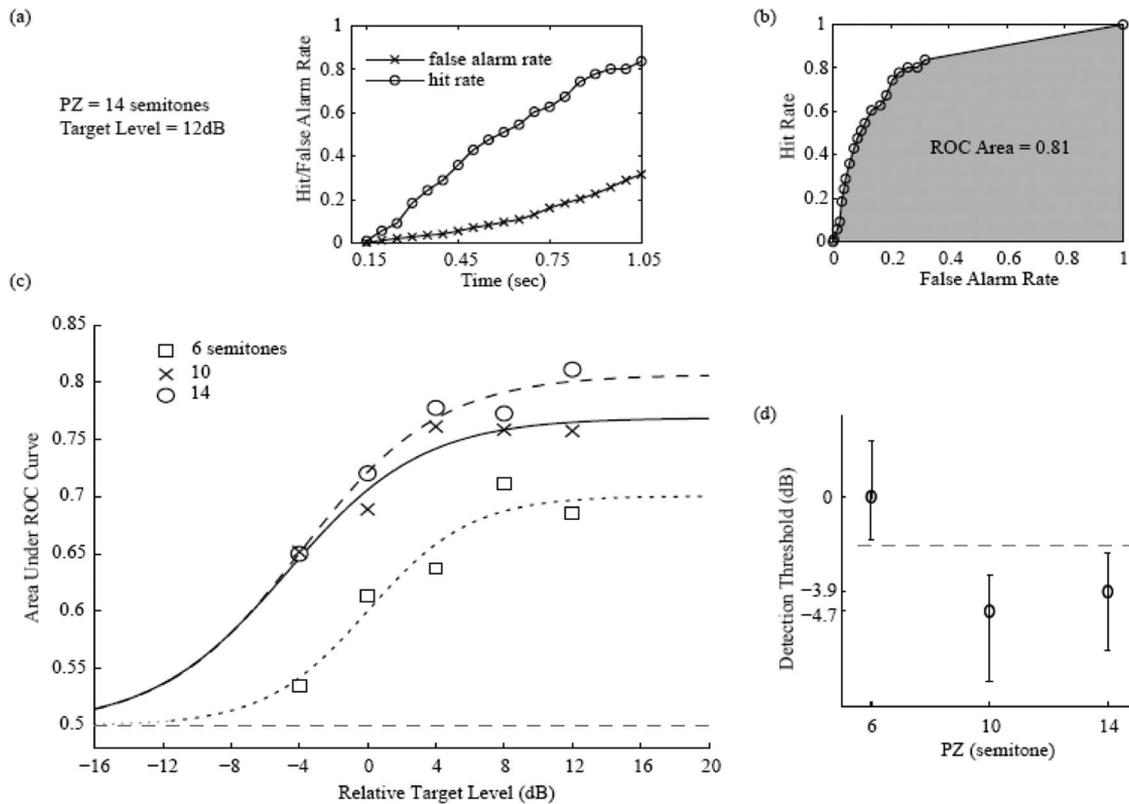


Figure 4. Performance measures in Experiment 2. (a) Example series of hit and false-alarm rates generated by varying the response window from 150 to 1,050 ms after the target onset. These example data were obtained using a protected zone (PZ) half-width of 14 ST and a relative target level of +12 dB (relative to the masker level). (b) Example receiver operating characteristic (ROC) curve obtained by plotting the series of hit rates from Panel a as a function of the corresponding false alarm rates. The ROC area, shown in gray, was estimated using a nonparametric technique (see text for details). (c) ROC area as a function of PZ half-width and best-fitting psychometric functions. ROC areas are shown as symbols; the best-fitting psychometric functions as lines. The different symbols and line styles indicate different PZ conditions (6, 10, and 14 ST). (d) Detection thresholds estimated based on the psychometric functions shown in Panel c for the different PZ half-widths. The error bars indicate 95% CIs around the mean. The dashed line indicates that the 95% CIs did not overlap between the smallest and the two larger PZ conditions.

It is important to acknowledge that although the results have been discussed in terms of informational masking, a possible contribution of energetic masking cannot be completely eliminated. With the moderate stimulus level (50 dB SPL per tone) and wide PZ widths (12, 20, and 28 ST) used in this experiment, the contribution of energetic masking would probably have been minimal in humans because even the smallest (12 ST) PZ widths are roughly 10 times the ERB in normal-hearing listeners (Moore, 2003). As mentioned above, in the various animal species in which auditory-filter bandwidths have been measured behaviorally, these bandwidths have been found to be somewhat larger than in humans. However, for energetic masking to significantly limit the detection of the target tones in the current experiment, which involved PZ widths of 1 octave or more, frequency selectivity would have to be considerably less in ferrets than in humans.

The increases in ROC area over time following the onset of the stimulus sequence, which were seen in Figure 5, are qualitatively consistent with the results of psychophysical studies in humans,

which indicate that the target tones become more detectable over time within the course of the stimulus sequence (Gutschalk, Micheyl, & Oxenham, 2008; Kidd et al., 2003; Micheyl, Shamma, et al., 2007). This effect may be related to the phenomenon known as the “build-up of segregation,” whereby the probability of hearing a sequence of alternating tones as two separate streams instead of a single coherent stream increases gradually (over several seconds) following the stimulus onset (Anstis & Saida, 1985; Bregman, 1978; van Noorden, 1975). The increasing probability of detecting the target tones amid the randomly varying maskers may be related to increasing segregation of the target tones from the background tones over time. From this point of view, the present findings suggest that stream segregation takes some time in ferrets, as it does in humans. However, because the changes in ROC area shown in Figure 5 occurred over approximately 1 s, this effect can also be explained in terms of response time—more specifically, decision time—without necessarily implicating the build-up of stream segregation. The ferret may have needed more time to

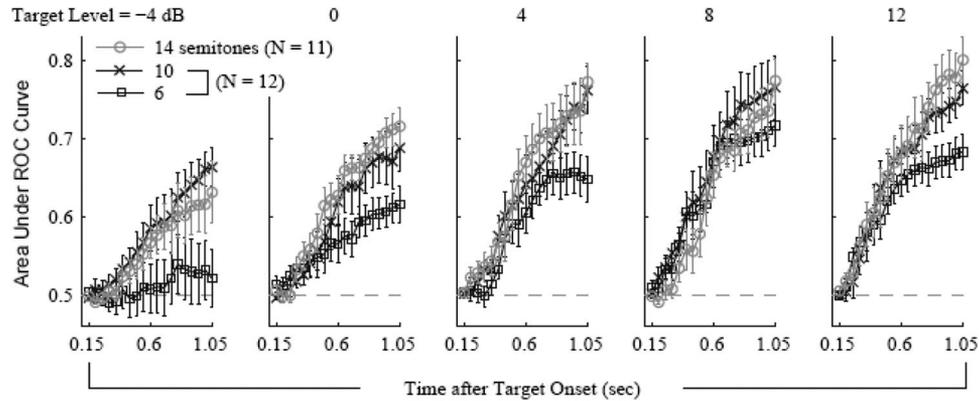


Figure 5. Area under the receiver operating characteristic (ROC) curve as a function of time after target onset in Experiment 2. The different panels correspond to different target levels (relative to the masker), from low (left) to high (right). The error bars are standard errors of the mean across daily sessions.

respond in conditions in which the target tones were harder to detect. Thus, the trends observed in Figure 5 could be reproduced, for instance, by a diffusion model of response time in which noisy sensory evidence accumulates toward a bound, and the rate of accumulation is determined by the strength of the sensory evidence (e.g., Ratcliff, Van Zandt, & McKoon, 1999). These trends could also be accounted for by probability-summation or multiple-looks models (Green & Swets, 1966; Viemeister & Wakefield, 1991), in which the probability of correct detection increases with the number of signals. Thus, these data do not allow us to conclusively dissociate components of response time that are unrelated to the build-up of segregation from components that are related to it.

The decrease in ROC area with decreasing target repetition rate in Figure 6 is qualitatively consistent with psychophysical data in humans (Kidd et al., 2003; Micheyl, Shamma, et al., 2007). These data show a decrease in d' as the target presentation rate decreased

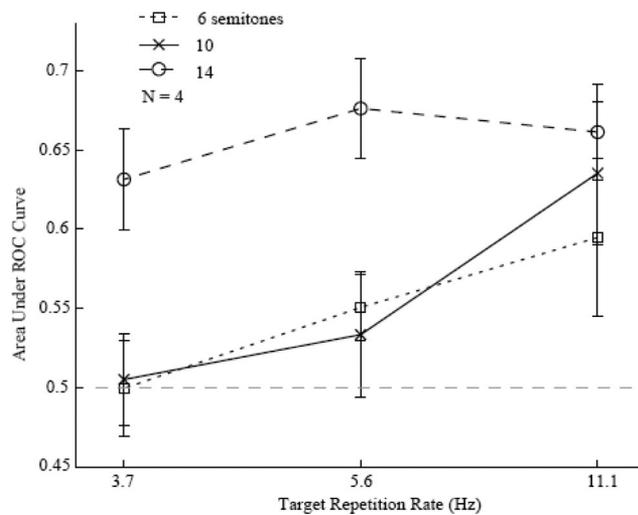


Figure 6. Area under the receiver operating characteristic (ROC) curve as a function of target repetition rate in Experiment 2. The different line styles indicate different protected zone (PZ) widths. The error bars denote standard errors of the mean across daily sessions.

from 16.7/s to 5.6/s, a range that partially overlaps with that tested in the ferret (11.1/s to 3.7/s). A similar effect was observed by Kidd et al. (1994), who measured masked detection thresholds rather than d' . These authors found that thresholds for the detection of four or eight tone bursts inside a randomly varying multitone background improved by about 15 dB as the interval between consecutive target bursts decreased from 400 ms (which in that study yielded a target rate of 2.2/s) to 50 ms (a rate of 9.1/s). However, the trend illustrated in Figure 6 may also be explained by an increase in detectability of the target tones as a function of their number—an effect discussed in the preceding paragraph and consistent with probability-summation or multiple-looks models. This is because, in the current experiment, target sequence length was kept constant, independent from tone repetition rate, so that the total number of target tones in the stimulus sequence increased with the repetition rate. However, this probability-summation or multiple-looks models explanation is contradicted by the experiments of Kidd et al. (2003), which used similar stimuli to those in Experiment 2, except for the masker components being synchronous with each target tone. They compared the target detection threshold under different repetition rates, but with the number of target tones fixed. They found that thresholds decreased with increasing rates, a finding inconsistent with a simple version of the multiple-looks model, but instead in favor of a perceptual segregation of the signal from the masker. Accordingly, it is also unlikely that the effect seen in our experiment is purely due to probability-summation or multiple-looks models.

General Discussion

The results indicate that the behavioral performance of ferrets in two auditory perception tasks, which have been used to measure auditory streaming in humans, varies as a function of stimulus parameters in a way that is qualitatively consistent with the human data. Specifically, in Experiment 1, higher performance and lower thresholds in the detection of frequency shifts between target tones at a given frequency were observed when temporally interleaved tones (interferers) were either absent or at a remote frequency compared with when the interfering tones were closer in frequency to the targets. This finding is qualitatively consistent with the

human psychophysical data of Micheyl, Carlyon, et al. (2005), which were explained as resulting from stream segregation allowing listeners to attend selectively to the target tones. Selective attention to the target sounds presumably allows the characteristics of these sounds (e.g., pitch or loudness) to be perceived more acutely, whereas other background sounds are analyzed less thoroughly by the auditory system. Accordingly, thresholds for the detection or discrimination of changes in the frequency (subjectively, pitch) of the B tones are expected to be smaller under stimulus conditions that promote stream segregation between the A and B tones. The increase in performance in the target-frequency discrimination task with increasing A–B frequency separation is consistent with stream segregation becoming easier as the A–B frequency separation increases. The finding of increasing performance, and decreasing thresholds, with increasing size of the PZ width in Experiment 2, where the task was to detect regularly repeating target tones among randomly varying masker tones, is also consistent with psychophysical data in humans (Micheyl, Shamma, et al., 2007). Here, the effect has been interpreted as resulting from wider frequency separations between the target and masker tones facilitating the perception of the target tones as a separate stream.

Although we cannot ascertain that the ferrets experienced the stimuli in these experiments in the same way as humans do, the finding of similar trends in behavioral performance as a function of stimulus parameters in the two species indicates that the perceptual organization of these stimuli varies in qualitatively the same way in the ferrets as it does in humans. At the same time, the present data indicate important quantitative differences in the way in which performance, or thresholds, in the two considered tasks vary as a function of stimulus parameters in ferrets and humans. In general, the ferrets needed larger frequency separations in Experiment 1 and larger PZ widths in Experiment 2, to be able to perform the tasks above chance. Moreover, even under the most favorable stimulus conditions (i.e., very large spectral separations), performance in the ferrets was still well below ceiling, and thresholds were still considerably larger than those measured in humans. This cannot be due solely to insufficient training, because the ferrets received fairly extensive training and performed these tasks or a simpler version of them repeatedly over the course of several months. This suggests that these tasks are intrinsically difficult for ferrets. A likely reason for this is that both tasks require selective attention, in addition to basic auditory detection and discrimination abilities, such as the ability to discriminate the frequencies of tones. Thus, although behavioral studies have found that the performance of various animal species in basic auditory detection or discrimination tasks can equal and sometimes exceed that of humans, these studies almost invariably used simpler stimuli than those used in the experiments described here. It is important to note that both of the tasks used in this study required the ferret to be able to sustain selective attention to a subset of stimuli within a sequence of sounds that contained other, irrelevant sounds. Unfortunately, such a selective-attention ability is required, to some extent, by any task that aims to measure performance in separating auditory streams.

Two limitations of the present study must be acknowledged. First, although the patterns of results that were observed as a function of stimulus parameters in both experiments are qualitatively consistent with those that have been observed and attributed

to auditory streaming in comparable experiments in humans, there remain numerous other parameters whose direct or indirect effects on streaming need to be investigated. For example, it is assumed that selective attention plays a key role in task performance in both of our tasks here, but it is of course difficult to assess precisely the role it plays in determining experimental thresholds. Instead, we are aware that it is extremely difficult, if not impossible, to measure auditory stream segregation without involving some form of selective attention. In addition, as in vision, the ability to attend selectively to certain dimensions or features of a stimulus is closely related to, as well as constrained by, perceptual grouping. Conversely, attention can influence auditory streaming—although the extent to which it does is still debated (see Carlyon, Cusack, Foxtan, & Robertson, 2001; Sussman, Horvath, Winkler, & Orr, 2007). Selective attention in frequency or some other sound dimension almost certainly played a role in previous behavioral measures of auditory streaming as well. Nevertheless, we believe that, to the extent that the psychophysical results that have been obtained using comparable stimuli and tasks in humans are related to auditory streaming (which, introspectively, they appear to be), the observation of similar trends in performance as a function of various stimulus parameters in an animal is a good indication that the animal is experiencing a similar perceptual phenomenon.

A second limitation of the current study, which future studies should aim to overcome, relates to the fact that in both tasks used in this study, stream segregation was beneficial to performance. It is known that auditory streaming depends on listener's intention or "attentional set." In particular, the A–B frequency separation required for a listener to experience two streams is smaller if the listener is actively trying to hear two separate streams than if the listener is trying to "hold on" the percept of a single stream for as long as possible (van Noorden, 1975, 1977). Therefore, an important goal for future studies is to measure animals' performance in tasks that promote stream integration, rather than segregation. Such tasks have already been devised and tested in human listeners. In particular, performance in tasks in which listeners have to judge accurately the relative timing of sounds within a sequence appears to be dramatically affected by factors that promote stream segregation and prevent stream integration. For instance, thresholds for the detection of a shift in the timing of the B tones relative to the temporally adjacent A tones in a repeating AB or ABA sequence have been found to be substantially higher when the A and B tones are widely separated in frequency and are perceived as separate streams than when frequency separation is small and all tones are heard as part of the same stream (Micheyl, Hunter, & Oxenham, in press; Roberts, Glasberg, & Moore, 2002, 2008; Vliegen, Moore, & Oxenham, 1999).

The development of behavioral tasks, which can be used to induce an animal to experience one of two bistable percepts (e.g., hear a sequence of tones as one stream or as two streams), and to indirectly verify that this percept was actually experienced through performance measures, has potentially important implications for studies of the neural correlates of perceptual experience (Logothetis & Schall, 1989). Over the past decade, a rapidly increasing number of studies have been devoted to unraveling the brain basis and neural mechanisms of auditory stream formation in both humans and animals (for reviews, see Carlyon, 2004; Micheyl, Carlyon, et al., 2007; Snyder & Alain, 2007). In particular, recordings of neural responses to alternating-tone sequences similar to

those used in studies of auditory streaming in humans have started to reveal potential neural correlates of auditory streaming at the single-unit level (Bee & Klump, 2004; Elhilali et al., 2009; Fishman et al., 2001; Itatani & Klump, 2009; Kanwal et al., 2003; Micheyl, Tian, et al., 2005; Pressnitzer et al., 2008). However, the conclusions of these studies are limited by the lack of behavioral data on auditory streaming under directly comparable stimulus conditions in the same species. Therefore, the two tasks described here could prove particularly useful in investigations into the neural basis of auditory streaming in animals. In particular, one advantageous feature of randomly varying multitone stimuli such as those used in Experiment 2 is that they can also be used to measure spectrotemporal receptive fields (e.g., Noreña, Gourévitch, Aizawa, & Eggermont, 2006; Noreña, Gourévitch, Pienkowski, Shaw, & Eggermont, 2008).

References

- Anstis, S., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 257–271.
- Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, *61*, 467–480.
- Baru, A. V. (1967). *Mechanisms of hearing* (in Russian). Leningrad: Nauka.
- Bee, M. A., & Klump, G. M. (2004). Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology*, *92*, 1088–1104.
- Bee, M. A., & Micheyl, C. (2008). The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, *122*, 235–251.
- Bey, C., & McAdams, S. (2002). SchemA-Based processing in auditory scene analysis. *Perception & Psychophysics*, *64*, 844–854.
- Bey, C., & McAdams, S. (2003). Postrecognition of interleaved melodies as an indirect measure of auditory stream formation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 267–279.
- Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 380–387.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, *89*, 244–249.
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, *8*, 465–471.
- Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 115–127.
- Carlyon, R. P., & Gockel, H. (2008). Effects of harmonicity and regularity on the perception of sound sources. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Auditory perception of sound sources* (pp. 191–213). New York: Springer.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.
- Dai, H. P., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *Journal of the Acoustical Society of America*, *89*, 2837–2842.
- David, S. V., Fritz, J. B., & Shamma, S. A. (2008, February). *Dynamics of rapid plasticity during positively and negatively reinforced behavior in auditory cortex*. Poster session at the meeting of 31st mid-winter meeting of the Association for Research in Otolaryngology, Phoenix, AZ.
- Dowling, W. J. (1968). Rhythmic fission and perceptual organization. *Journal of the Acoustical Society of America*, *44*, 369.
- Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*, 322–337.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- 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.
- Elliott, D. N., Stein, L., & Harrison, M. J. (1960). Determination of absolute-intensity thresholds and frequency-difference thresholds in cats. *Journal of the Acoustical Society of America*, *32*, 380–384.
- Fay, R. R. (1988). Comparative psychoacoustics. *Hearing Research*, *34*, 295–306.
- Fay, R. R. (1998). Auditory stream segregation in goldfish (*Carassius auratus*). *Hearing Research*, *120*, 69–76.
- Fay, R. R. (2000). Spectral contrasts underlying auditory stream segregation in goldfish. *Journal of the Association for Research in Otolaryngology*, *1*, 120–128.
- Fay, R. R. (2008). Sound source perception and stream segregation in nonhuman vertebrate animals. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Auditory perception of sound sources* (pp. 307–323). New York: Springer.
- 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.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*, 437–455.
- Fritz, J. B., Shamma, S. A., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, *6*, 1216–1223.
- Gourevitch, G. (1970). Delectability of tones in quiet and in noise by rats and monkeys. In W. C. Stebbins (Ed.), *Animal psychophysics: The design and conduct of sensory experiments* (pp. 67–97). New York: Plenum Press.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Krieger.
- Greenberg, G. S., & Larkin, W. D. (1968). Frequency-response characteristics of auditory observers detecting signals of a single frequency in noise: The probe-signal method. *Journal of the Acoustical Society of America*, *44*, 1513–1523.
- Gutschalk, A., Micheyl, C., Melcher, J. R., Rupp, A., Scherg, M., & Oxenham, A. J. (2005). Neuromagnetic correlates of streaming in human auditory cortex. *The Journal of Neuroscience*, *25*, 5382–5388.
- Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, *6*, e138.
- Gutschalk, A., Oxenham, A. J., Micheyl, C., Wilson, E. C., & Melcher, J. R. (2007). Human cortical activity during streaming without spectral cues suggests a general substrate for auditory stream segregation. *The Journal of Neuroscience*, *27*, 13074–13081.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.
- Heffner, R. S., Heffner, H. E., & Masterton, B. (1971). Behavioral measurements of absolute and frequency difference thresholds in guinea pig. *Journal of the Acoustical Society of America*, *49*, 1888–1895.
- Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: Stream segregation of birdsong by

- European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology*, *111*, 3–13.
- Itatani, N., & Klump, G. M. (2009). Auditory streaming of amplitude-modulated sounds in the songbird forebrain. *Journal of Neurophysiology*, *101*, 3212–3225.
- Izumi, A. (2002). Auditory stream segregation in Japanese monkeys. *Cognition*, *82*, 113–122.
- Kalluri, S., Depireux, D. A., & Shamma, S. A. (2008). Perception and cortical neural coding of harmonic fusion in ferrets. *Journal of the Acoustical Society of America*, *123*, 2701–2716.
- 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.
- Kelly, J. B. (1970). *Effects of lateral lemniscal and neocortical lesions on auditory absolute thresholds and frequency difference thresholds of the rat* (unpublished doctoral thesis). Vanderbilt University.
- Kidd, G. J., Mason, C. R., Deliwala, P. S., Woods, W. S., & Colburn, H. S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, *95*, 3475–3480.
- Kidd, G. J., Mason, C. R., & Richards, V. M. (2003). Multiple bursts, multiple looks, and stream coherence in the release from informational masking. *Journal of the Acoustical Society of America*, *114*, 2835–2845.
- Kidd, G. J., Mason, C. R., Richards, V. M., Gallun, F. J., & Durlach, N. I. (2008). Informational masking. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Auditory perception of sound sources* (pp. 143–189). New York: Springer.
- Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, *63*, 1421–1455.
- Logothetis, N. K., & Schall, J. D. (1989, August 18). Neuronal correlates of subjective visual perception. *Science*, *245*, 761–763.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q., & White, W. (1998). Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences. *Journal of the Acoustical Society of America*, *103*, 3581–3587.
- Micheyl, C., & Carlyon, R. P. (1998). Effects of temporal fringes on fundamental-frequency discrimination. *Journal of the Acoustical Society of America*, *104*, 3006–3018.
- Micheyl, C., Carlyon, R. P., Cusack, R., & Moore, B. C. J. (2005). Performance measures of auditory organization. In D. Pressnitzer, A. de Cheveigné, S. McAdams, & L. Collet (Eds.), *Auditory signal processing: Physiology, psychoacoustics, and models* (pp. 203–211). New York: Springer.
- Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A. J., Rauschecker, J. P., . . . Wilson, E. C. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Research*, *229*, 116–131.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36–47.
- Micheyl, C., Hunter, C., & Oxenham, A. J. (in press). Auditory stream segregation and the perception of across-frequency synchrony. *Journal of Experimental Psychology: Human Perception and Performance*.
- Micheyl, C., Moore, B. C., & Carlyon, R. P. (1998). The role of excitation-pattern cues and temporal cues in the frequency and modulation-rate discrimination of amplitude-modulated tones. *Journal of the Acoustical Society of America*, *104*, 1039–1050.
- Micheyl, C., Shamma, S., & Oxenham, A. J. (2007). Hearing out repeating elements in randomly varying multitone sequences: A case of streaming? In B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, & J. Verhey (Eds.), *Hearing: From basic research to applications* (pp. 267–274). Berlin: Springer.
- Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron*, *48*, 139–148.
- Miller, G. A., & Heise, G. A. (1950). The trill threshold. *Journal of the Acoustical Society of America*, *22*, 637–638.
- Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America*, *54*, 610–619.
- Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press.
- Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United With Acustica*, *88*, 320–333.
- Moore, B. C. J., Hafter, E. R., & Glasberg, B. R. (1996). The probe-signal method and auditory-filter shape: Results from normal- and hearing-impaired subjects. *Journal of the Acoustical Society of America*, *99*, 542–552.
- Neff, D. L., & Green, D. M. (1987). Masking produced by spectral uncertainty with multicomponent maskers. *Perception & Psychophysics*, *41*, 409–415.
- Nelson, D. A., & Kiester, T. E. (1978). Frequency discrimination in the chinchilla. *Journal of the Acoustical Society of America*, *64*, 114–126.
- Nienhuys, T. W., & Clark, G. M. (1979). Critical bands following the selective destruction of cochlear inner and outer hair cells. *Acta Otolaryngologica*, *88*, 350–358.
- Noreña, A. J., Gourévitch, B., Aizawa, N., & Eggermont, J. J. (2006). Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. *Nature Neuroscience*, *9*, 932–939.
- Noreña, A. J., Gourévitch, B., Pienkowski, M., Shaw, G., & Eggermont, J. J. (2008). Increasing spectrotemporal sound density reveals an octave-based organization in cat primary auditory cortex. *The Journal of Neuroscience*, *28*, 8885–8896.
- Pickles, J. O. (1979). Psychophysical frequency resolution in the cat as determined by simultaneous masking and its relation to auditory-nerve resolution. *Journal of the Acoustical Society of America*, *66*, 1725–1732.
- Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology*, *18*, 1124–1128.
- Prosen, C. A., Moody, D. B., Sommers, M. S., & Stebbins, W. C. (1990). Frequency discrimination in the monkey. *Journal of the Acoustical Society of America*, *88*, 2152–2158.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261–300.
- Richards, V. M., & Tang, Z. (2006). Estimates of effective frequency selectivity based on the detection of a tone added to complex maskers. *Journal of the Acoustical Society of America*, *119*, 1574–1584.
- Roberts, B., Glasberg, B. R., & Moore, B. C. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *Journal of the Acoustical Society of America*, *112*, 2074–2085.
- Roberts, B., Glasberg, B. R., & Moore, B. C. (2008). Effects of the build-up and resetting of auditory stream segregation on temporal discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 992–1006.
- Rose, M. M., & Moore, B. C. J. (2000). Effects of frequency and level on auditory stream segregation. *Journal of the Acoustical Society of America*, *108*, 1209–1214.
- Rose, M. M., & Moore, B. C. J. (2005). The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects. *Hearing Research*, *204*, 16–28.
- Seaton, W. H., & Trahotis, C. (1975). Comparison of critical ratios and critical bands in the monaural chinchilla. *Journal of the Acoustical Society of America*, *57*, 193–199.
- Sek, A., & Moore, B. C. J. (1995). Frequency discrimination as a function of frequency, measured in several ways. *Journal of the Acoustical Society of America*, *97*, 2479–2486.
- Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of

- human cochlear tuning from otoacoustic and behavioral measurements. *Proceedings of the National Academy of Sciences, USA*, 99, 3318–3323.
- Sinnott, J., & Brown, C. (1993). Effects of varying signal and noise levels on pure tone frequency discrimination in humans and monkeys. *Journal of the Acoustical Society of America*, 93, 1535–1540.
- Sinnott, J. M., Brown, C. H., & Brown, F. E. (1992). Frequency and intensity discrimination in Mongolian gerbils, African monkeys and humans. *Hearing Research*, 59, 205–212.
- Sinnott, J. M., Petersen, M. R., & Hopp, S. L. (1985). Frequency and intensity discrimination in humans and monkeys. *Journal of the Acoustical Society of America*, 78, 1977–1985.
- Sloan, A. M., Dodd, O. T., & Rennaker, R. L. (2009). Frequency discrimination in rats measured with tone-step stimuli and discrete pure tones. *Hearing Research*, 251, 60–69.
- Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin*, 133, 780–799.
- 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.
- Sussman, E. S., Horvath, J., Winkler, I., & Orr, M. (2007). The role of attention in the formation of auditory streams. *Perception & Psychophysics*, 69, 136–152.
- Sussman, E., Ritter, W., & Vaughan, H. G. J. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, 36, 22–34.
- Syka, S., Rybalko, N., Brožek, G., & Jilek, M. (1996). Auditory frequency and intensity discrimination in pigmented rats. *Hearing Research*, 100, 107–113.
- Talwar, S. K., & Gerstein, G. L. (1998). Auditory frequency discrimination in the white rat. *Hearing Research*, 126, 135–150.
- Talwar, S. K., & Gerstein, G. L. (1999). A signal detection analysis of auditory-frequency discrimination in the rat. *Journal of the Acoustical Society of America*, 105, 1784–1800.
- van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences* (unpublished doctoral thesis). University of Technology, Eindhoven, the Netherlands.
- van Noorden, L. P. A. S. (1977). Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. *Journal of the Acoustical Society of America*, 61, 1041–1045.
- Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, 90, 858–865.
- Vliegen, J., Moore, B. C., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *Journal of the Acoustical Society of America*, 106, 938–945.
- Walker, K. M., Schnupp, J. W., Hart-Schnupp, S. M., King, A. J., & Bizley, J. K. (2009). Pitch discrimination by ferrets for simple and complex sounds. *Journal of the Acoustical Society of America*, 126, 1321–1335.
- Watson, C. S., Kelly, W. J., & Wroton, H. W. (1976). Factors in the discrimination of tonal patterns: II. Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America*, 60, 1176–1186.
- Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, 61, 178–184.
- Wilson, E. C., Melcher, J. R., Micheyl, C., Gutschalk, A., & Oxenham, A. J. (2007). Cortical fMRI activation to sequences of tones alternating in frequency: Relationship to perceived rate and streaming. *Journal of Neurophysiology*, 97, 2230–2238.
- Wisniewski, A. B., & Hulse, S. H. (1997). 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*, 111, 337–350.
- Yin, P., Fritz, J. B., & Shamma, S. A. (2009). *Do ferrets perceive relative pitch?* Manuscript submitted for publication.
- Yin, P., Mishkin, M., Sutter, M., & Fritz, J. B. (2008). Early stages of melody processing: Stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. *Journal of Neurophysiology*, 100, 3009–3029.

Received August 7, 2009

Revision received October 26, 2009

Accepted November 11, 2009 ■