Adaptation Reveals Multiple Levels of Representation in Auditory Stream Segregation

Joel S. Snyder
University of Nevada Las Vegas and Veterans Affairs Boston
Healthcare System/Harvard Medical School

Olivia L. Carter
Harvard University

Erin E. Hannon
University of Nevada Las Vegas and Harvard University

Claude Alain
Baycrest Centre for Geriatric Care

When presented with alternating low and high tones, listeners are more likely to perceive 2 separate streams of tones ("streaming") than a single coherent stream when the frequency separation (Δf) between tones is greater and the number of tone presentations is greater ("buildup"). However, the same large-Δf sequence reduces streaming for subsequent patterns presented after a gap of up to several seconds. Buildup occurs at a level of neural representation with sharp frequency tuning. The authors used adaptation to demonstrate that the contextual effect of prior Δf arose from a representation with broad frequency tuning, unlike buildup. Separate adaptation did not occur in a representation of Δf independent of frequency range, suggesting that any frequency-shift detectors undergoing adaptation are also frequency specific. A separate effect of prior perception was observed, dissociating stimulus-related (i.e., Δf) and perception-related (i.e., 1 stream vs. 2 streams) adaptation. Viewing a visual analogue to auditory streaming had no effect on subsequent perception of streaming, suggesting adaptation in auditory-specific brain circuits. These results, along with previous findings on buildup, suggest that processing in at least 3 levels of auditory neural representation underlies segregation and formation of auditory streams.

Keywords: auditory scene analysis, adaptation, buildup, frequency shift detector, cross-modal

Understanding how preceding context affects auditory perception is a central problem because real-world stimuli such as speech and music usually occur in a temporal sequence. For example, spoken words are typically preceded by other words in a phrase and musical notes are typically preceded by other notes in a melody. In the laboratory, manipulating context is a powerful methodology for determining the nature of the psychological and neural representations underlying perception. For example, classic studies in visual spatial perception that adapted participants to

context patterns provided evidence for neurons in the human visual system that are specifically tuned to size and spatial orientation (e.g., Blakemore & Campbell, 1969). Studies of the adapting effects of context have also been applied fruitfully in the auditory domain (Repp, 1985). Measuring the adapting effects of context can provide evidence for the existence of multiple levels of processing, thus providing insight into the hierarchy of processing levels in a specific perceptual phenomenon.

We recently reported a novel context effect on auditory stream segregation that likely occurs at sensory levels of processing (Snyder, Carter, Lee, Hannon, & Alain, 2008). The auditory stream segregation paradigm is used as a probe for understanding general principles of auditory segregation and perceptual formation of auditory objects (i.e., auditory scene analysis; Bregman, 1990). In this paradigm, low tones (A), high tones (B), and silences (–) are presented in a repeating ABA– pattern (Bregman & Campbell, 1971; Van Noorden, 1975). When the frequency difference between A and B (Δf) is small and the repetition rate is slow, listeners typically report a single ABA– "galloping" pattern. When Δf is large enough and the repetition rate is sufficiently fast, listeners report hearing two separate streams of tones, each in a metronome-like rhythm (i.e., A–A–A–A– . . . and B–B– . . .). At intermediate Δf, the stimulus is ambiguous and perception can alternate between interpretations of one and two streams (Pressnitzer & Hupé, 2006).

Perceiving two streams ("streaming") does not occur instantly. Instead, listeners initially hear one stream that perceptually splits into two streams after several seconds of repeating ABA– patterns, a facilitative context effect called *buildup* (Anstis & Saida, 1985;

Carlyon, Cusack, Foxton, & Robertson, 2001). Despite the fact that attention can modulate buildup (Carlyon et al., 2001; Cusack, Deeks, Aikman, & Carlyon, 2004; Snyder, Alain, & Picton, 2006), psychophysical and neurophysiological evidence suggests that buildup begins in early stages of the ascending auditory system. For example, Anstis and Saida (1985) showed that buildup can be reset by shifting the A and B tones up or down by just a few semitones, suggesting adaptation in brain regions with narrow frequency tuning typical of early brain areas along the ascending auditory pathway. Consistent with this psychophysical result, Pressnitzer, Micheyl, Sayles, and Winter (2007) measured neural correlates of buildup in the cochlear nucleus, the first stage of the auditory central nervous system that receives input from the ear, providing direct evidence for early auditory processing underlying buildup.

In contrast to the known facilitation of streaming caused by a large Δf and more ABA– repetitions (i.e., buildup), Snyder et al. (2008) showed that streaming during a given trial was reduced if the listener had been exposed to a large Δf in previous trials. This contrastive context effect (cf. Holt, 2005, 2006) was strongest when the current trial had an intermediate Δf, suggesting that only ambiguous patterns are highly susceptible to this influence of context. The effect lasted for tens of seconds but began decaying immediately after the pattern ended, suggesting the involvement of long-lasting sensory adaptation (Micheyl, Tian, Carlyon, & Rauschecker, 2005; Ulanovsky, Las, Farkas, & Nelken, 2004; Ulanovsky, Las, & Nelken, 2003) or long auditory sensory memory (Cowan, 1984). Currently, it is not known what representations undergo adaptation resulting in the contrastive context effect we observed, but understanding this may provide general insights into what levels of processing are important in segregation of sequential patterns and formation of auditory streams.

The current study consists of four experiments that presented a context sequence, followed by a short silent period and a test sequence. We manipulated the relationship between the context and test sequences to test hypotheses regarding the representations that undergo adaptation and lead to the context effect. Specifically, we investigated whether the mechanisms undergoing adaptation (a) were narrowly tuned in frequency like buildup (Anstis & Saida, 1985) or more widely tuned, which could provide information about the stages of auditory processing involved; (b) whether they were tuned to an explicit representation of Δf (e.g., neurons tuned to different size Δf; cf. Anstis & Saida, 1985; Demany & Ramos, 2005; Van Noorden, 1975) that might undergo adaptation during the context and influence subsequent perception during the test; and (c) whether they were tuned to the auditory modality or whether they reside in multimodal brain areas (cf. Cusack, 2005).

We predict that if the neurons undergoing adaptation during the context effect are narrowly tuned as in buildup (Anstis & Saida, 1985), shifting the frequency range of tones by a small amount between the context and test periods should disrupt the context effect. If the neurons undergoing adaptation are more broadly tuned, it would be necessary to shift the frequency range by a larger amount to detect any disruption of the context effect. Demonstrating different tuning widths for neurons involved in buildup versus the contrastive context effect would provide evidence for multiple levels of representation in streaming as proposed by recent theoretical frameworks (Denham & Winkler, 2006; Snyder & Alain, 2007b), which contradict the theory that streaming primarily occurs in the auditory periphery (Beauvois & Meddis, 1996; Hartmann & Johnson, 1991; Van Noorden, 1975).

In all four experiments described below, the test sequence was an ABA– pattern with a Δf of six semitones between the A and B tones, an ambiguous pattern that could be heard as one or two streams by most listeners in our previous report (Snyder et al., 2008). The current study also included trials in which the context and test sequences were identical to each other, which allowed us to test for perceptual adaptation (rather than stimulus-specific adaptation) as the basis for the context effect by controlling the influence of previous Δf. Specifically, it is possible that there are populations of neurons in the auditory system that represent perception of one versus two streams and that perceiving one of these percepts causes adaptation of that particular percept, independent of the specific stimuli presented. Therefore, if perceiving one stream versus two streams during the context sequence influences perception during the test in the absence of any difference in Δf between the two sequences, this would demonstrate an effect of perceptual context on perception of streaming. We also included trials in which the context period consisted of silence to determine whether presenting a small and a large Δf each had a measurable context effect relative to silence.

Experiment 1A

Previous research has shown that the adaptation underlying streaming buildup (i.e., leading up to the initial switch from one stream to two streams) is highly frequency specific (Anstis & Saida, 1985), suggesting that buildup relies on adaptation of neurons in stages of the auditory system that are narrowly tuned to the individual tone frequencies. The purpose of the current experiment was to determine the frequency specificity of the contrastive context effect we observed previously (Snyder et al., 2008). On each trial, we presented a context sequence consisting of an ABA– pattern with a small or large Δf, which was followed after a short break by the perceptually ambiguous (intermediate Δf) test sequence. The test sequence was either shifted to a different frequency range or had the same A tone as the preceding context sequence. Thus, if the contrastive context effect reflects adaptation to the repeated presentation of the specific A tone frequency, changing the A tone between the two sequences should disrupt the contrastive context effect. Experiment 1A tests the strength of the context effect when the A tone frequency changes from the context to the test by a relatively small amount; Experiment 1B tests the context effect when the A tone changes by a larger amount. Thus, Experiments 1A and 1B together test how widely tuned to frequency are the neurons undergoing adaptation during the context sequence.

Method

Participants. Eleven normal-hearing adults (6 men and 5 women, age range = 24–37 years, mean age = 30.3 years) from the Harvard University community participated after giving written informed consent according to the guidelines of the Faculty of Arts and Sciences at Harvard University. Joel S. Snyder and Olivia L. Carter participated.

Materials and procedure. Stimuli were generated and behavioral responses were collected by a custom Matlab script with

functions from the Psychtoolbox (Brainard, 1997), running on a Pentium 4 computer with a SoundMAX Integrated Digital Audio card. Sounds were presented binaurally through Koss UR-30 headphones at 65 dB SPL. Pure tones were 50 ms in duration, including 10-ms rise–fall times. The time between adjacent A and B tone onsets within each ABA– cycle was 120 ms, as was the silent duration between ABA triplets.

Each trial consisted of a 6.72-s context period, a 1.44-s silent period, and a 6.72-s test period. The intertrial interval was 6 s. The test period contained 14 ABA– repetitions with Δf equal to 6 semitones with one of the following two sets of A and B tone frequencies: (a) A = 300 Hz and B = 424 Hz or (b) A = 500 Hz and B = 707 Hz. This Δf of 6 semitones was chosen because it usually leads to a bistable percept in which it is possible to hear one or two streams. Table 1 shows the tone frequencies of the context sequences, each lasting 14 ABA–repetitions (6.72 s). There were six different context sequence types for each of the two test sequences: (a) 6.72-s silence; (b) same as the test sequence; (c) smaller (3 semitone) Δf and the same A tone frequency as the test sequence; (d) larger (12 semitone) Δf and the same A tone frequency as the test sequence; (e) smaller (3 semitone) Δf and a different A tone frequency from the test sequence; and (f) larger (12 semitone) Δf and a different A tone frequency from the test sequence. With two test sequences and six types of context sequence, there were 12 different trial types, each presented once each per block in a random order without replacement. Six blocks were presented. Thus, each context–test combination was presented six times to each participant.

Participants were seated in a dimly lit room and were asked to maintain fixation on a white cross on a gray background in the center of a computer screen throughout the experiment. Participants were asked to press and hold down the down-arrow key when they perceived one stream and to press and hold down the left-arrow key if they perceived two streams. Participants were instructed to hold down both keys during any periods of confusion or perception of both one and two streams and to release all keys during intervals between the stimulus presentations. Participants were not explicitly told that there were context and test periods; instead, they were told to respond whenever the sound stimuli were played to them. Participants were encouraged to let their perception take a natural time course and not to bias their perception in favor of one stream or two streams. The button presses were sampled synchronously with the A tones (every 240 ms), resulting in 62 data points per trial.

Data Analysis

We quantified the proportion of time during which participants perceived two streams at each time point, making it possible to graph the time course of streaming. To quantify streaming for statistical analysis, we calculated the proportion of total time that participants reported two streams for the context and test periods, separately. These values were obtained for each participant across each of the 12 different trial types. The values for trials with 3 and 12 semitone context sequences were entered into a three-way repeated measures analysis of variance (ANOVA) to test for differences in perception of streaming during the test period depending on the test sequence (300–424–400– or 500–707–500–), Δf of the context sequence (3 or 12 semitones), and whether the A tone frequency was the same or different in the context and test periods. The degrees of freedom were adjusted with the Greenhouse-Geisser epsilon, and all reported probability estimates were based on the reduced degrees of freedom.

To test the effect of previous percept, we analyzed trials that had identical test and context sequences. These trials were sorted according to whether participants perceived one stream or two streams at the end of the context period. The mean proportion of streaming during the test sequence was then compared across the two percepts (previously perceiving one stream vs. two streams) using a one-way repeated measures ANOVA.

Results and Discussion

Our main interest in this experiment was to test the frequency specificity of the contrastive context effect. First, however, we present data replicating the typical effect of Δf on perception of two streams (Bregman & Campbell, 1971; Van Noorden, 1975) and the context effect we reported previously (Snyder et al., 2008). Figure 1 shows the proportion of time participants heard streaming over the course of the trial, including the context, silent, and test periods. When the context sequence had a small 3-semitone Δf, participants heard less streaming compared with when the context sequence had a large 12-semitone Δf. Participants heard intermediate amounts of streaming when the context sequence had a 6-semitone Δf (same condition). This effect of Δf reached significance, $F(1, 10) = 197.85$, $p < .001$, $\eta^2 = 0.95$, and similar results were obtained regardless of the A tone frequency during the context sequence. No streaming was reported when the context was silence.

Table 1
Frequencies (Hertz) of Tones Used in Context and Test Sequences for Experiment 1A

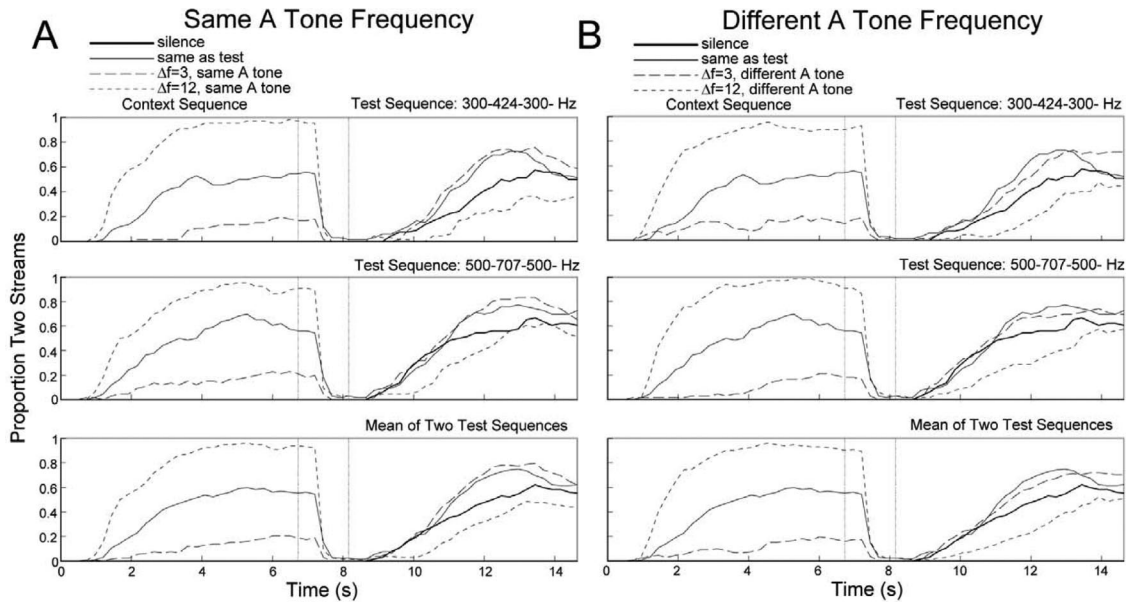|	Test sequence			
	A = 300 Hz, B = 424 Hz		A = 500 Hz, B = 707 Hz	
Context sequence	A tone	B tone	A tone	B tone
Silence	—	—	—	—
Same as test	300	424	500	707
Δf = 3 semitones/same A	300	357	500	595
Δf = 12 semitones/same A	300	600	500	1000
Δf = 3 semitones/different A	500	595	300	357
Δf = 12 semitones/different A	500	1000	300	600

Figure 1. Effect of stimulus context in Experiment 1A. Proportion of time the auditory tone sequences were heard as two streams (streaming) for the context period (left portion of panels) and test period (right portion of panels). The test sequence always had a Δf of 6 semitones. The top panels include trials in which the test sequence was 300–424–300–, the middle panels include trials in which the test sequence was 500–707–500–, and the bottom panels show the mean across these two test sequences. The context period contained silence (thick solid line), the same sequence as the test sequence (thin solid line), a sequence with a 3-semitone Δf (long dashed line), or a sequence with a 12-semitone Δf (short dashed line). A: Trials in which the context sequences with 3 and 12 semitone Δf had the same A tone frequency as the test sequence. B: Trials in which the context sequences with 3 and 12 semitone Δf had a different A tone frequency than the test sequence. Note that the traces for the silence and same conditions are exactly the same for the left and right panels.

During the test period, which always consisted of a 6-semitone Δf pattern, participants showed the same contrastive context effect that we reported for the first time recently (Snyder et al., 2008). Specifically, when the context sequence had a 12-semitone Δf (i.e., biased toward perception of two streams), participants reported less perception of two streams during the test sequence. When the context had a 3-semitone Δf (i.e., biased toward perception of one stream), participants reported more perception of two streams during the test sequence. The difference among these conditions was reflected by a significant main effect of Δf, $F(1, 10) = 38.07$, $p < .001$, $\eta^2 = 0.79$. Similar amounts of streaming during the test were reported whether the context had a Δf of 3 or 6 semitones (same condition), and streaming was enhanced in these conditions compared with when no context sequence was presented (i.e., silence condition). This may reflect the fact that a buildup effect occurred when presenting identical context and test sequences, which has a similar magnitude effect as presenting a small Δf (3 semitones) during the context. Presenting a 12-semitone context resulted in less streaming than when no context was presented, suggesting that presenting both small and large Δf sequences produced context effects relative to silence.

The current study was designed to identify factors responsible for the contrastive context effect. The main finding we report is that the same magnitude of contrastive effect was seen regardless of whether the A and B tones in the test sequence fell within the same frequency range as the context sequence ($\Delta f \times$ A Tone

interaction), $F(1, 10) = 0.27$, $p = .61$, $\eta^2 = 0.03$. Thus, the contrastive context effect is not highly frequency specific and is unlikely to occur in neurons narrowly tuned to the A and B tones, unlike buildup (Anstis & Saida, 1985). Instead, the context effect appears to generalize to an ambiguous ABA– sequence in a different frequency range.

One possible explanation for the context effect is that it is driven by perception during the context period rather than by the stimulus properties during the context. In our previous study, we addressed this by sorting behavioral responses that occurred at the end of trials on the basis of whether listeners perceived one or two streams on the previous trial. We found no evidence that perception of two streams on the previous trial caused less streaming on the current trial (Snyder et al., 2008). Rather, we found the opposite tendency, with participants reporting hearing two streams more often when they had also perceived two streams on the previous trial, similar to the well-known buildup effect (Anstis & Saida, 1985; Carlyon et al., 2001). However, this analysis had insufficient data to perform statistical analysis.

The current experiment allowed us to more rigorously test the influence of perception during the context period on subsequent perception during the test period. To control the influence of Δf during the context period, we only analyzed the condition in which the ambiguous test sequence was preceded by the identical ambiguous 6-semitone context sequence. We sorted the data from the test sequence according to whether participants perceived one

stream or two streams at the end of the context sequence. One participant was excluded from this analysis because of always reporting two streams at the end of the context sequence. When participants reported hearing one stream at the end of the context period, they were also more likely to report hearing the same one stream percept during the test sequence (Figure 2, solid line). Similarly, when two streams were heard at the end of the context sequence, this same two-stream percept was more likely to be heard during the test period (Figure 2, broken line).

In summary, this experiment showed that larger Δf during the context sequence caused less streaming during the test sequence, consistent with our previous findings (Snyder et al., 2008). This effect occurred regardless of whether the context and test sequences shared the same A tone frequency. Simply perceiving more streaming at the end of the context sequence resulted in slightly (but nonsignificantly) more streaming during the test sequence. This facilitative effect of percept provides evidence that the relevant factors responsible for the contrastive context effect are likely to be related to stimulus-specific adaptation rather than perceptual adaptation.

Experiment 1B

Although Experiment 1A showed no reliable diminution of the contrastive context effect as a result of changing the frequency of the A tone between the context and test sequences, this does not completely rule out the possibility that the adaptation leading to the context effect occurs in auditory brain areas that have frequency-specific neurons. For instance, it is possible that the neurons undergoing adaptation had relatively broad frequency tuning that did not result in a reliable diminution of the context effect because the difference in frequency range between the context and test sequences was too small. To test this hypothesis, we used similar stimuli and procedure as in Experiment 1A but with context and test sequences that were much farther separated in frequency.

Method

Participants. Fifteen normal-hearing adults (8 men and 7 women, age range = 18–37 years, mean age = 27.7 years) from the Harvard University community participated after giving written informed consent according to the guidelines of the Faculty of Arts and Sciences at Harvard University. Joel S. Snyder and Olivia L. Carter participated.

Materials and procedure. As shown in Table 2, this experiment was constructed as in Experiment 1A except the test sequences had frequencies as follows: (a) A = 300 Hz and B = 424 Hz or (b) A = 1300 Hz and B = 1838 Hz. The context sequences were adjusted accordingly.

Data Analysis

We analyzed the data in the same manner as in Experiment 1A.

Results and Discussion

Figure 3 shows the proportion of time during which participants heard streaming over the course of the trial. As in Experiment 1A, participants heard more streaming during the context period when the context sequence had a larger Δf, $F(1, 14) = 137.02$, $p < .001$, $\eta^2 = 0.91$. During the test period, which always consisted of a 6-semitone Δf pattern, participants again showed a contrastive context effect, as confirmed by a significant main effect of context Δf, $F(1, 14) = 17.53$, $p < .001$, $\eta^2 = 0.56$. However, unlike in Experiment 1A, the difference between the 3- and 12-semitone Δf conditions was reduced when the context and test sequences did not share the same A tone frequency ($\Delta f \times$ A Tone interaction), $F(1, 14) = 9.60$, $p < .01$, $\eta^2 = 0.41$.

As in Experiment 1A, we assessed the effect of prior perception. Figure 4 shows the proportion of time participants heard two streams for trials in which the context and test sequences were identical. When one stream was perceived at the end of the context period, participants were more likely to report hearing the same one-stream percept during the test sequence (Figure 4, solid line). Similarly, when two streams were heard at the end of the context
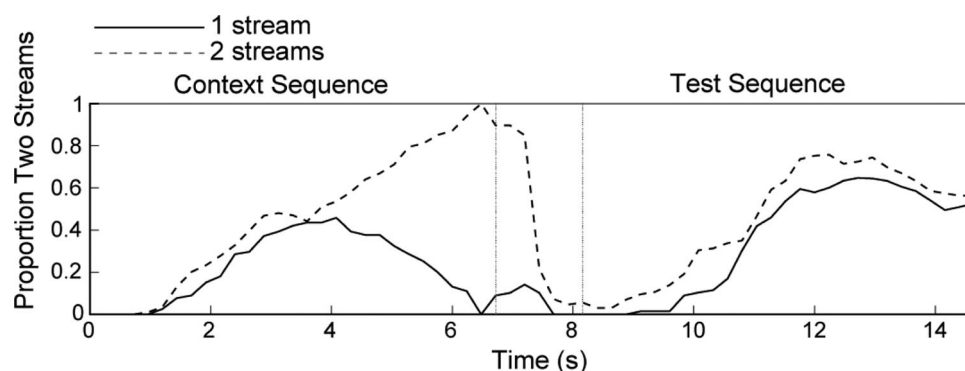

Figure 2. Effect of perceptual context in Experiment 1A. Proportion of time trials heard as streaming for the context period (left portion of panel) and test period (right portion of panel) showing the effect of different perceptual contexts ("previous percept") on perception of streaming during the test period. Data, shown only for trials in which the context and test sequences were the same, are collapsed across the two test sequences. The time course of perceptual dominance is shown both for trials in which the dominant percept at the end of the context period was one stream (solid line) and two streams (dashed line).

Table 2
Frequencies (Hertz) of Tones Used in Context and Test Sequences for Experiment 1B

	Test sequence			
	A = 300 Hz, B = 424 Hz		A = 1300 Hz, B = 1838 Hz	
Context sequence	A tone	B tone	A tone	B tone
Silence	—	—	—	—
Same as test	300	424	1300	1838
Δf = 3 semitones/same A	300	357	1300	1546
Δf = 12 semitones/same A	300	600	1300	2600
Δf = 3 semitones/different A	1300	1546	300	357
Δf = 12 semitones/different A	1300	2600	300	600

sequence, this same two-stream percept was more likely to be heard during the test period (Figure 4, broken line). Unlike in Experiment 1A, this effect of perception reached significance in the current experiment, $F(1, 14) = 14.87$, $p < .005$, $\eta^2 = 0.52$.

In summary, this experiment showed that larger Δf during the context sequence caused less streaming during the test sequence and that this contrastive context effect was reduced when the context and test sequences were presented in very distant frequency ranges. However, perceiving streaming at the end of the context sequence resulted in significantly more streaming during the test sequence. Thus, the Δf-dependent context effect is contrastive in nature, meaning that presenting a large Δf during the context results in more streaming during the context but less streaming during the test sequence; conversely, the perception-dependent context effect is facilitative in nature because perceiving two streams at the end of the context sequence also makes perceiving two streams more likely during the test sequence.

These data suggest on one hand the existence of both stimulus-specific and percept-specific adaptation, which may occur at different levels of representation with opposite effects on perception of streaming. On the other hand, in vision, a recent series of behavioral and computational studies have shown that the apparently opposite phenomena of perceptual stability (facilitation) and perceptual reversals (contrastive context effect) may be driven by a single process involving multiple timescales of adaptation (Brascamp et al., 2008; Brascamp, Knapen, Kanai, van Ee, & van den Berg, 2007; Noest, van Ee, Nijs, & van Wezel, 2007). Further studies are therefore needed to determine whether the same relationship can be seen in the timescales of the perceptual facilitation and contrastive effects in the auditory domain.

Experiments 1A and 1B together demonstrate that the adaptation underlying the contrastive context effect generalized to different frequencies and was still present, although reduced, when the context and test frequencies were very distant in frequency.
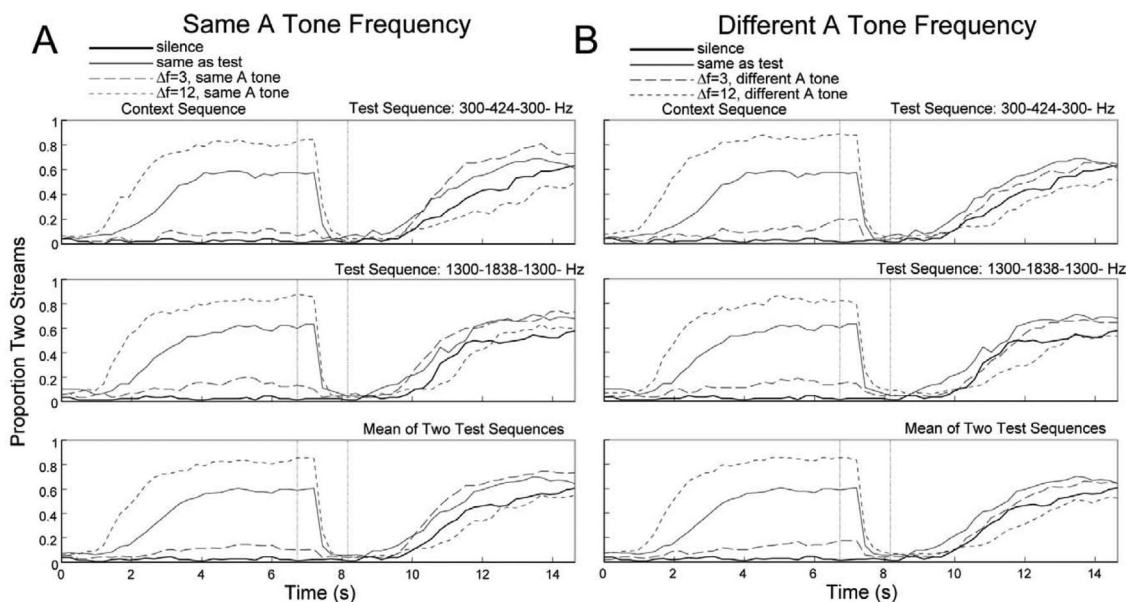

Figure 3. Effect of stimulus context in Experiment 1B. Similar to Figure 1 except the top panels include trials in which the test sequence was 300–424–300–, the middle panels include trials in which the test sequence was 1300–1838–1300– and the bottom panels show the mean across these two test sequences.
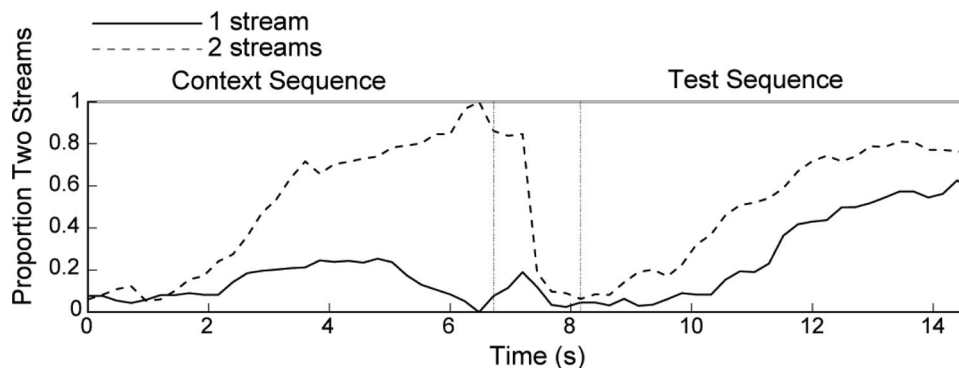
Figure 4. Effect of perceptual context in Experiment 1B. Proportion of time trials heard as streaming for the context period (left portion of panel) and test period (right portion of panel) showing the effect of different perceptual contexts ("previous percept") on perception of streaming during the test period. Data, shown only for trials in which the context and test sequences were the same, are collapsed across the two test sequences. The time course of perceptual dominance is shown both for trials in which the dominant percept at the end of the context period was one stream (solid line) and two streams (dashed line).

This finding suggests that the contrastive context effect occurs at a late stage of processing in the auditory system because early stages of processing show relatively sharp frequency specificity (Kaas & Hackett, 2000). Thus, together with evidence from studies of buildup showing adaptation of neurons narrowly tuned to frequency (Anstis & Saida, 1985), Experiments 1A and 1B provide evidence for at least three levels of adaptation in stream segregation by showing adaptation of neurons broadly tuned in frequency and adaptation of perception.

Experiment 2

One possible explanation for the contrastive context effect is that during the context period, neurons selective for Δf (in addition to being widely tuned to frequency) undergo adaptation. Specifically, presenting a context sequence with a small Δf would adapt a specific subset of neurons tuned to small Δf. Subsequently presenting a sequence with an intermediate Δf would result in fewer than normal responses from neurons tuned to small Δf, which would result in a greater likelihood of the listener hearing two streams. This would be analogous to the adaptation underlying motion aftereffects in vision, illustrated by the waterfall illusion in which stationary objects appear to move up after viewing of downward motion for a number of seconds (Anstis, Verstraten, & Mather, 1998). Evidence for Δf-specific adaptation during streaming would provide support for the existence of neurons that are tuned to frequency shifts (cf. Demany & Ramos, 2005) that do not depend on repeated stimulation of a specific part of a tonotopic map.

Method

Participants. Eight normal-hearing adults (5 men and 3 women, age range = 22–37 years, mean age = 29.8 years) from the Harvard University community participated after giving written informed consent according to the guidelines of the Faculty of Arts and Sciences at Harvard University. Joel S. Snyder and Olivia L. Carter participated.

Materials and procedure. On each trial, the A and B frequencies during the test period were held constant (A = 300 Hz and B = 424 Hz). The context sequences consisted of either 6.72 s of silence or a 6.72-s sequence in which each ABA– repetition had an A tone frequency that was randomly selected from the range 600–2600 Hz and a B tone frequency that was either 3 or 12 semitones above the A tone frequency. With one test sequence and three types of context, there were three different trial types, each presented four times per block in a random order without replacement. Six blocks were presented. Thus, each trial type was presented 24 times to each participant.

The procedure was similar to Experiments 1A and 1B with the following exception. Participants were told that some of the sequences they would hear would have tones that randomly move around in frequency and that they might not perceive two streams but that if they did they should press the button corresponding to this percept.

Data Analysis

We analyzed the data using a one-way repeated-measures ANOVA with context (silence, 3 semitones, or 12 semitones) as the factor.

Results and Discussion

Figure 5 shows the proportion of time participants heard streaming during the context, silent, and test periods. During the random-frequency ABA– context, the likelihood of reporting two streams was similarly low for both the 3- and 12-semitone Δf conditions, $F(1, 7) = 0.40$, $p = .55$, $\eta^2 = 0.06$. During the test period with the constant-frequency ABA– sequences, participants heard similar amounts of streaming, regardless of the prior context, $F(1, 7) = 1.88$, $p = .20$, $\eta^2 = 0.21$. These data suggest that to the extent that the contrastive context effect observed in Experiments 1A and 1B and in our previous investigation (Snyder et al., 2008) was because of adaptation of frequency-shift detectors (Demany & Ramos, 2005), these frequency-shift detectors require repeated stimulation
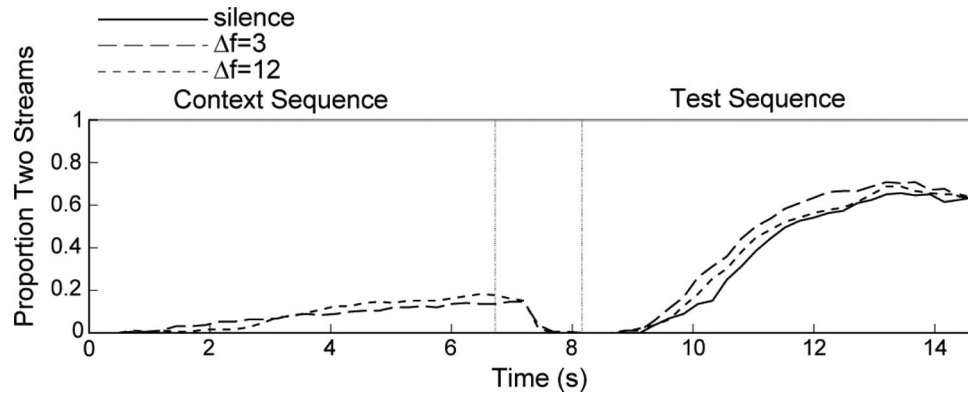
Figure 5. Effect of stimulus context in Experiment 2. The test sequence was 300–424–300– and the context period consisted of silence (solid line) and a sequence of ABA– repetitions in which each A tone was chosen randomly from the range 600–2600 Hz and the B tone was 3 (long dashed line) or 12 (short dashed line) semitones higher.

of a particular area of a tonotopic map. Another possible reason for the lack of a Δf-based context effect in the current experiment is that the context stimuli used might not have engaged the same mechanisms that are typically recruited to process ABA– patterns, resulting in adaptation in different neural pathways than those activated by the test stimuli. Thus, future experiments are needed to further evaluate the possibility of frequency-shift detectors undergoing adaptation during streaming.

Experiment 3

The previous experiments provided evidence for the existence of two different context effects on auditory stream segregation, a contrastive effect of stimulus context (i.e., small Δf vs. large Δf) and a facilitative effect of perceptual context (i.e., one stream vs. two streams). The purpose of Experiment 3 was to understand whether the two context effects arise from brain circuits that are auditory specific or whether they reflect adaptation at a multimodal level of processing that is selective for the relative number of stimulus components (Burr & Ross, 2008) or a general stimulus quality such as coherence versus segregation (cf. Cusack, 2005). Although the existence of such category-general adaptation would be surprising, it would be consistent with observations that exposure to looming (expanding) visual motion that appears to be coming toward the observer results in the perception of a constant tone sounding like it was moving away from the observer (Kitagawa & Ichihara, 2002). To test for the existence of similar multimodal influences in auditory streaming, we presented one of the test sequences that we used in the previous experiments but instead of presenting auditory context sequences, we presented moving visual plaid stimuli that could be perceived as either a coherent pattern moving vertically or two superimposed gratings sliding over each other in opposite directions.

The plaid stimulus shows similar perceptual dynamics as in auditory stream segregation (Pressnitzer & Hupé, 2006). For example, the coherent percept generally dominates initially, but later the two grating components will segregate and appear to slide independently. Analogous to the effect of Δf on auditory stream segregation, the plaid stimulus can be systematically biased toward

the segregated sliding percept by increasing the difference between the two motion directions (α). At intermediate values of alpha, the stimulus is ambiguous and the coherent and segregated percepts tend to alternate indefinitely (Hupe & Rubin, 2003). An alpha-based contrastive context effect, similar to the Δf-based effect observed in the auditory modality (Snyder et al., 2008) occurs when both the context and the test stimuli consist of moving plaids (Carter, Snyder, Rubin, & Nakayama, 2007).

Method

Participants. Eight normal-hearing adults (6 men and 2 women, age range = 22–34 years, mean age = 29.8 years) from the Harvard University community participated after giving written informed consent according to the guidelines of the Faculty of Arts and Sciences at Harvard University. Joel S. Snyder and Olivia L. Carter participated.

Materials and procedure. As shown in Figure 6, the visual plaid stimuli were composed of superimposed rectangular-wave gratings with a duty cycle of 0.33 (one third dark gray 14 cd/m^2 and two thirds light gray 22 cd/m^2) and a spatial frequency of 0.3 cycles/degree of visual angle. The dark gratings moved at a speed of 2°/s and appeared transparent (the intersection regions were visibly darker, 7 cd/m^2). The plaid pattern was presented within a circular aperture of 13° in diameter, and the background outside this aperture was also gray (14 cd/m^2). A red 0.2° fixation dot was presented centrally, within a dark gray exclusion zone (9 cd/m^2) with a diameter of 1.5°. The direction of component motion was manipulated by rotating the gratings symmetrically clockwise or counterclockwise, relative to the vertical axis. In all trials, the coherent plaid appeared to move upward.

On each trial, the auditory test sequence had frequencies as follows: A = 300 Hz and B = 424 Hz. The visual context lasted for 14.88 s and consisted of (a) a uniform light gray surface or (b) the visual plaid stimulus described above with alpha values of 50°, 105°, or 150°. The fixation point and the dark gray background were presented throughout the entire block of auditory and visual testing.
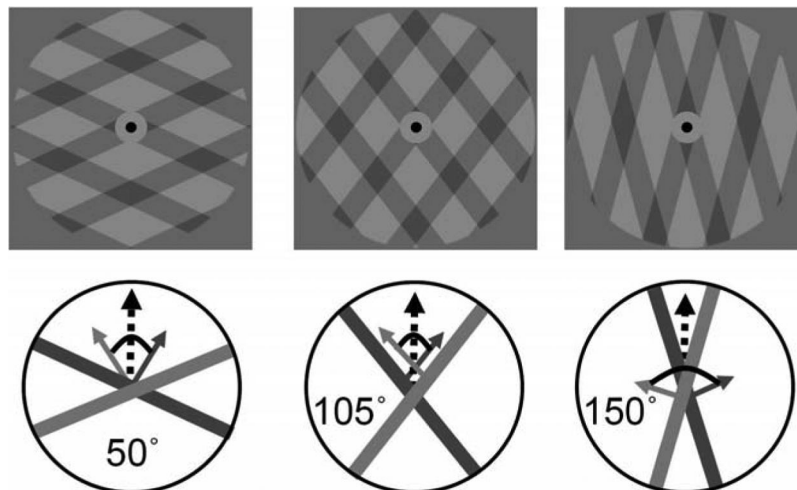
Figure 6. Plaid stimuli presented during the context period in Experiment 3. The angle (α) between the direction of motion when perceiving one coherent plaid stimulus (vertical) and the direction of motion of the two grating components could have values of 50°, 105°, or 150°. Larger angle leads to more perception of two visual gratings, analogously to the effect of increasing Δf on perception of two auditory streams.

With one test sequence and four types of context, there were four different trial types, each presented three times each per block in a random order without replacement. Six blocks were presented. Thus, each trial type was presented 18 times to each participant.

The procedure was similar to the previous experiments with the following exception. Observers viewed the visual fixation and plaid stimuli from 57 cm, with head position maintained using an adjustable chin rest.

Data Analysis

We analyzed the data using a one-way repeated-measures ANOVA with context (blank screen, $\alpha = 50°$, $\alpha = 105°$, or $\alpha = 150°$) as the factor.

Results and Discussion

Figure 7A shows the proportion of time participants perceived two segregated stimulus components over the course of the trial, including the visual context, silent–blank, and auditory test periods. Despite large stimulus-related effects on visual perception during the context period,[1] there was very little effect of visual stimulus context on auditory stream segregation, $F(1, 7) = 1.78$, $p = .20$, $\eta^2 = 0.20$. This demonstrates that the stimulus-based contrastive context effect does not generalize from the visual to the auditory modality, suggesting that context effects on auditory streaming are mediated by auditory-specific brain circuits, below any level of multimodal perception of segregation. This is despite the fact that a contrastive context effect occurs when both the context and the test stimuli consist of moving plaids (Carter et al., 2007). It is therefore likely that the contrastive context effect on streaming is because of adaptation of auditory-specific neurons.

To test whether the facilitative effect of perceptual context on streaming is also the result of modality-specific processing, we reanalyzed the trials in which the context pattern had an intermediate alpha value of 105°. Specifically, we sorted the data from the test sequence according to whether participants perceived one moving plaid or two gratings sliding past each other at the end of the context sequence to test for any effect on auditory perception. Figure 7B shows the proportion of time participants heard two streams depending on the visual percept at the end of the context period. There was no effect of visual perception on subsequent auditory perception during the test period, $F(1, 7) = 0.33$, $p = .58$, $\eta^2 = 0.05$.

This finding suggests that the facilitative effect of perceptual context on streaming occurs in auditory-specific brain circuits, as does the contrastive effect of stimulus context. This fails to support the proposal that the adaptation effects seen in auditory streaming reflect the involvement of an abstract non–modality-specific brain area responsible for perception of object number (Burr & Ross, 2008) or relative stimulus coherence (Cusack, 2005). Although the negative results observed here do not rule out the possible involvement of multimodal perceptual mechanisms, our results are consistent with findings in vision that perceptual stabilization depends on the continuity of low-level factors (Chen & He, 2004; Maier, Wilke, Logothetis, & Leopold, 2003).

General Discussion

To reveal the nature of the representation underpinning perceptual organization of sequential auditory patterns, the current series of experiments presented trials in which a test ABA– pattern that could be heard as one stream or two streams was preceded by

---

[1] During the context period, the perception of two objects increased with increasing values of alpha, consistent with previous reports (Hupe & Rubin, 2003) and analogous to the effect of Δf on perception of streaming. This was confirmed by a significant effect of context type on perception during the context period, $F(1, 7) = 79.23$, $p < .001$, $\eta^2 = 0.92$. Moreover, the effect of alpha on visual perception was similar to that of Δf on auditory perception in Experiments 1A, 1B, and 2.
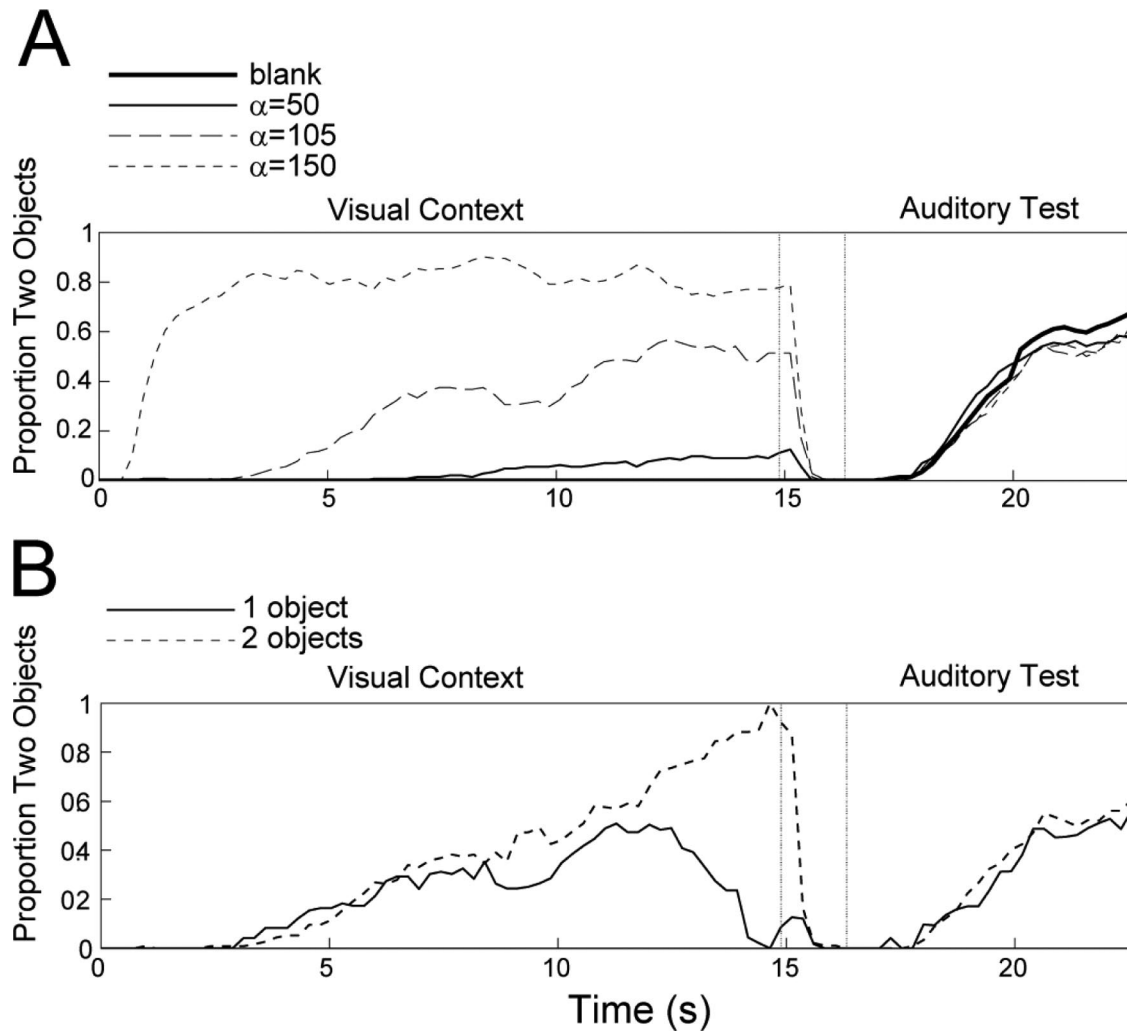
Figure 7. A: Effect of visual stimulus context on auditory perception in Experiment 3. The test sequence was 300–424–300–, and the context period consisted of a blank screen (thick solid line), a moving plaid stimulus with angle (α) = 50° (thin solid line), α = 105° (long dashed line), or α = 150° (short dashed line). B: Effect of visual perceptual context on auditory perception in Experiment 3. Data are shown only for trials in which the context was ambiguous (α = 105°). The time course of perceptual dominance is shown both for trials in which the dominant percept at the end of the context period was one object (solid line) and for trials in which it was two objects (dashed line).

various types of adapting context sequences. In Experiments 1A and 1B, we replicated our previous finding that preceding an ambiguous test sequence with a context sequence having a large Δf caused less perception of streaming compared with when the context sequence had a small Δf (a contrastive effect of context). Experiment 1B showed a significant enhancement of perceiving two streams during the test period when participants perceived two streams at the end of the context (a facilitative effect of perceptual context). The same trend for a facilitative effect was seen in Experiment 1A, but the effect size was not as large and did not reach significance. Beyond demonstrating a facilitative effect of prior percept, this rules out the possibility that the contrastive effect resulted from a response bias on the part of participants to switch responses between the end of the context period and the start of the test period. Experiments 1A and 1B further showed that

when the context and test sequences were separated in frequency range by a small amount, the contrastive effect of context was unchanged compared with when they were in the same frequency range; however, when the frequency range of the context and test differed by a larger amount, the context effect was significantly diminished although still present.

Relation of Context Effects to Buildup

These findings suggest that the contrastive effect of context is caused by adaptation of neurons that have relatively wide tuning to sound frequency. This is in contrast to the case of Δf-based segregation and buildup (i.e., the process leading up to the initial switch from perceiving one to perceiving two streams), both of which rely on adaptation of neurons with narrow frequency tuning

(Anstis & Saida, 1985; Gutschalk et al., 2005; Micheyl et al., 2005; Pressnitzer et al., 2007; Snyder & Alain, 2007a; Snyder et al., 2006). It is therefore likely that the contrastive context effect is a result of adaptation in different levels of the auditory system compared with Δf-based segregation and buildup. Evidence from neurophysiological studies in non-human animals showed neural adaptation that can predict human perceptual buildup occurring as early as the ventral cochlear nucleus (Pressnitzer et al., 2007) and also in primary auditory cortex (Micheyl et al., 2005), suggesting that Δf-based segregation and buildup are computed at the first stage of central auditory processing and that this information is preserved up to the first stage of auditory cortical processing. Thus, it is likely that the neural mechanisms of the contrastive context effect operate in brain areas that integrate information across larger frequency ranges. Such frequency integration could occur as early as the inferior colliculus (Sinex & Li, 2007). Another possibility is that feedback from higher brain areas could modulate activity in earlier anatomical points in the ascending auditory pathway, causing the observed context effect (Suga, 2008). Another difference between the contrastive context effect and buildup is that buildup can be enhanced by prior exposure to a single tone stream and does not depend on prior exposure to an alternating-tone pattern (Beauvois & Meddis, 1997; Rogers & Bregman, 1993, 1998; also see Sussman & Steinschneider, 2006). Finally, buildup and the Δf-based effect can be distinguished from each other by the fact that they act in opposite directions on perception of streaming, the former being facilitative and the latter being contrastive.

Like buildup, the effect of perceptual context reported here is a facilitative context effect. Thus, an important question for future studies is the extent to which the facilitative effect of perceptual context shares mechanisms with buildup. This could be addressed by testing how the facilitative context effect is influenced by changes in factors that affect buildup such as changes in sound location (Rogers & Bregman, 1993, 1998) or sound level (Rogers & Bregman, 1998) and by ignoring the ABA– sequence (Carlyon et al., 2001; Cusack et al., 2004; Snyder et al., 2006).

Frequency-Shift Detectors

Given the likelihood that the contrastive context effect relies on mechanisms that integrate information across frequency, this raised the possibility that the neurons undergoing adaptation are selective for higher order auditory features. Experiment 2 therefore tested the hypothesis that neurons undergoing adaptation are selectively tuned to Δf without regard to frequency range. To test this hypothesis, we presented a context sequence in which each ABA– repetition had a randomly selected A tone that was far removed from the frequency range of the test sequence and a B tone frequency that was either 3 or 12 semitones above the A tone frequency. Presenting a small- or a large-Δf context sequence of this type did not affect subsequent perception of streaming. These data therefore do not lend support to the idea that there are auditory neurons that are tuned to frequency shifts of various sizes as proposed by several investigators (Anstis & Saida, 1985; Demany & Ramos, 2005; Van Noorden, 1975; cf. Rogers & Bregman, 1993). However, it is possible that such frequency-shift detectors are present and undergo adaptation during streaming but require repeated stimulation of a single frequency region. To generate the

context effects observed in the present study, such frequency-shift detectors would have to be present in brain areas that integrate across sound frequencies and across time (e.g., Brosch & Schreiner, 1997; Malone & Semple, 2001).

Modality Specificity of Context Effects

Given the possibility that the contrastive context effect is generated at relatively high levels of the auditory system and the fact that many brain areas that were traditionally considered to be auditory specific also show responses to stimuli from other modalities (Schroeder & Foxe, 2005), Experiment 3 tested whether the representations undergoing adaptation are multimodal. We presented visual context patterns consisting of a moving plaid stimulus that can be perceived as one or two objects and shows similar perceptual dynamics as auditory stream segregation (Pressnitzer & Hupé, 2006), followed by an ambiguous auditory ABA– sequence. Presenting a visual stimulus that was likely to be perceived as one object or two objects did not produce a contrastive context effect on streaming. Thus, although the neurons undergoing adaptation during the auditory contrastive context effect might be in relatively high-level brain areas, we found no evidence that these neurons are the same as those undergoing adaptation during the visual contrastive context effect (Carter et al., 2007). Furthermore, whether participants perceived one or two objects at the end of the visual context did not produce a facilitative context effect on streaming. The lack of contrastive or facilitative effects in Experiment 3 also rules out response biases as an explanation for context effects observed when both the context and the test sequences were auditory in Experiments 1A and 1B. In the case of the contrastive effect, it is not the case that simply pressing one button during the context causes participants to press the opposite button during the test; and in the facilitative effect, it is not the case that pressing one button during the context causes participants to press the same button during the test.

Despite the fact that no positive evidence was found for multimodal brain areas undergoing adaptation, it is still possible that similar computations in different brain areas underlie auditory and visual perception (Pressnitzer & Hupé, 2006). Thus, it may be useful to consider predictions from models of visual perception for auditory streaming. For example, a recent computational model of visual bistable perception (Noest et al., 2007) predicts that unambiguous stimuli lead to subsequent perceptual alternations. This is consistent with the basic finding we have reported that presenting large-Δf sequences resulted in less subsequent perception of two streams and more subsequent perception of two streams for small-Δf sequences. Work from the same group also demonstrated that a single process may underlie both perceptual stabilization and alternations, if it is assumed that adaptation exists over multiple timescales (Brascamp et al., 2007, 2008). Further experiments are need to test the extent to which this model can explain the effects observed here. The fact that our previous study showed that the contrastive context effect is quite robust, even for very long silences between sequences (Snyder et al., 2008), suggests that some characteristics may be different between modalities and even between specific paradigms within a modality.

Multiple Levels of Auditory-Specific Representation in Streaming

A prominent theory of auditory stream segregation called the *peripheral channeling hypothesis* proposes that adaptation in tonotopically organized peripheral structures (i.e., cochlea and auditory nerve) is sufficient to explain streaming (Beauvois & Meddis, 1996; Hartmann & Johnson, 1991; Van Noorden, 1975). More recently, however, several lines of evidence using behavioral, computational, and neurophysiological techniques have cast doubt on this theory. On the basis of a review of these data, we recently proposed that perception of one or two streams is the result of a complex set of processes occurring at multiple levels in both the peripheral and the central auditory systems (Snyder & Alain, 2007b; also see Denham & Winkler, 2006; Moore & Gockel, 2002). The current study supports this hypothesis by identifying the existence of two different types of context effect on streaming, a contrastive stimulus-based effect and a facilitative perception-based effect, the former of which is likely to be computed in auditory brain areas with coarse frequency resolution.

Summary

Four experiments were performed to identify the nature of the representations underlying contextual influences on auditory stream segregation. The data indicated that the contrastive effect of previous Δf on streaming was the result of adaptation in auditory neurons with wide frequency tuning, unlike streaming buildup. We found no evidence that this contrastive context effect resulted from adaptation of neurons tuned to specific Δf sizes, although such adaptation may require continual stimulation of a particular tonotopic region. The contrastive context effect of Δf could be distinguished from a facilitative effect of previous perception on streaming. Both the contrastive and the facilitative context effects were based on auditory-specific representations. Thus, along with previous findings on buildup, the current study suggests the existence of at least three levels of representation within the auditory system underlying perceptual organization of sequential auditory patterns.

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.

Anstis, S., Verstraten, F. A. J., & Mather, G. (1998). The motion aftereffect: A review. *Trends in Cognitive Science, 2,* 111–117.

Beauvois, M. W., & Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *Journal of the Acoustical Society of America, 99,* 2270–2280.

Beauvois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception & Psychophysics, 59,* 81–86.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology (London), 203,* 237–260.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brascamp, J. W., Knapen, T. H., Kanai, R., Noest, A. J., van Ee, R., & van den Berg, A. V. (2008). Multi-timescale perceptual history resolves visual ambiguity. *PLoS ONE, 3,* e1497.

Brascamp, J. W., Knapen, T. H., Kanai, R., van Ee, R., & van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision, 7,* 1–12.

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.

Brosch, M., & Schreiner, C. E. (1997). Time course of forward masking tuning curves in cat primary auditory cortex. *Journal of Neurophysiology, 77,* 923–943.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18,* 425–428.

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.

Carter, O., Snyder, J., Rubin, N., & Nakayama, K. (2007). Contrastive effects of prior context on visual and auditory grouping. *International Brain Research Organization Abstracts,* POS-SUN-166.

Chen, X., & He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Current Biology, 14,* 1013–1017.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370.

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience, 17,* 641–651.

Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 643–656.

Demany, L., & Ramos, C. (2005). On the binding of successive sounds: Perceiving shifts in nonperceived pitches. *Journal of the Acoustical Society of America, 117,* 833–841.

Denham, S. L., & Winkler, I. (2006). The role of predictive models in the formation of auditory streams. *Journal of Physiology (Paris), 100,* 154–170.

Gutschalk, A., Micheyl, C., Melcher, J. R., Rupp, A., Scherg, M., & Oxenham, A. J. (2005). Neuromagnetic correlates of streaming in human auditory cortex. *Journal of Neuroscience, 25,* 5382–5388.

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception, 9,* 155–184.

Holt, L. L. (2005). Temporally nonadjacent nonlinguistic sounds affect speech categorization. *Psychological Science, 16,* 305–312.

Holt, L. L. (2006). The mean matters: Effects of statistically defined nonspeech spectral distributions on speech categorization. *Journal of the Acoustical Society of America, 120,* 2801–2817.

Hupe, J. M., & Rubin, N. (2003). The dynamics of bi-stable alternation in ambiguous motion displays: A fresh look at plaids. *Vision Research, 43,* 531–548.

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America, 97,* 11793–11799.

Kitagawa, N., & Ichihara, S. (2002, March 14). Hearing visual motion in depth. *Nature, 416,* 172–174.

Maier, A., Wilke, M., Logothetis, N. K., & Leopold, D. A. (2003). Perception of temporally interleaved ambiguous patterns. *Current Biology, 13,* 1076–1085.

Malone, B. J., & Semple, M. N. (2001). Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *Journal of Neurophysiology, 86,* 1113–1130.

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.

Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United With Acustica, 88,* 320–333.

Noest, A. J., van Ee, R., Nijs, M. M., & van Wezel, R. J. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: A low-level neural model. *Journal of Vision, 7,* 1–14.

Pressnitzer, D., & Hupé, J. M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology, 16,* 1351–1357.

Pressnitzer, D., Micheyl, C., Sayles, M., & Winter, I. M. (2007). Responses to long-duration tone sequences in the cochlear nucleus [Abstract 1048]. *Association for Research in Otolaryngology Abstracts,* 131.

Repp, B. H. (1985). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92,* 81–110.

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics, 53,* 179–189.

Rogers, W. L., & Bregman, A. S. (1998). Cumulation of the tendency to segregate auditory streams: Resetting by changes in location and loudness. *Perception & Psychophysics, 60,* 1216–1227.

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, "unisensory" processing. *Current Opinion in Neurobiology, 15,* 454–458.

Sinex, D. G., & Li, H. (2007). Responses of inferior colliculus neurons to double harmonic tones. *Journal of Neurophysiology, 98,* 3171–3184.

Snyder, J. S., & Alain, C. (2007a). Sequential auditory scene analysis is preserved in normal aging adults. *Cerebral Cortex, 17,* 501–512.

Snyder, J. S., & Alain, C. (2007b). 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.

Snyder, J. S., Carter, O. L., Lee, S.-K., Hannon, E. E., & Alain, C. (2008). Effects of context on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1007–1016.

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 194,* 169–183.

Sussman, E., & Steinschneider, M. (2006). Neurophysiological evidence for context-dependent encoding of sensory input in human auditory cortex. *Brain Research, 1075,* 165–174.

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience, 24,* 10440–10453.

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

Van Noorden, L. P. A. S. (1975). Temporal coherence in the perception of tone sequences. *Unpublished doctoral dissertation, Eindhoven University of Technology, Eindhoven.*