Build-up of auditory stream segregation induced by tone sequences of constant or alternating frequency and the resetting effects of single deviants

Nicholas R. Haywood and Brian Roberts

Psychology, School of Life and Health Sciences, Aston University, Birmingham B4 7ET, UK

Running Title: Build-up and resetting of auditory streaming

Version Date: February 15, 2013

Contact details:

Email: n.haywood@ucl.ac.uk

Telephone: +44 20 7679 8981

Email: b.roberts@aston.ac.uk

Telephone: +44 121 204 3887

Current address / address for correspondence:

Nicholas Haywood, Ph.D.

UCL Ear Institute

332 Grays Inn Road

London

WC1X 8EE, UK

ABSTRACT

A sequence of constant-frequency tones can promote streaming in a subsequent sequence of alternating-frequency tones, but why this effect occurs is not fully understood and its time course has not been investigated. Experiment 1 used a 2.0-s-long constant-frequency inducer (10 repetitions of a low-frequency pure tone) to promote segregation in a subsequent, 1.2-s-long test sequence of alternating low- and high-frequency tones. Replacing the final inducer tone with silence substantially reduced reported test-sequence segregation. This reduction did not occur when either the 4th or 7th inducer was replaced with silence. This suggests that a change at the induction/test-sequence boundary actively resets build-up, rather than less segregation occurring simply because fewer inducer tones were presented. Furthermore, Experiment 2 found that a constant-frequency inducer produced its maximum segregation-promoting effect after only 3 tones – this contrasts with the more gradual build-up typically observed for alternating-frequency sequences. Experiment 3 required listeners to judge continuously the grouping of 20-s-long test sequences. Constant-frequency inducers were considerably more effective at promoting segregation than alternating ones; this difference persisted for ~10 s. In addition, resetting arising from a single deviant (longer tone) was associated only with constant-frequency inducers. Overall, the results suggest that constant-frequency inducers promote segregation by capturing one subset of test-sequence tones into an on-going, pre-established stream, and that a deviant tone may reduce segregation by disrupting this capture. These findings offer new insight into the dynamics of stream segregation, and have implications for the neural basis of streaming and the role of attention in stream formation.

KEYWORDS: Auditory grouping, stream segregation, tone sequences, build-up, resetting, deviant tone

INTRODUCTION

An important aspect of the perceptual representation of acoustic stimuli is the ability to integrate sounds separated in time but arising from a common source into coherent perceptual streams (see, e.g., Bregman, 1990). Accurate perceptual representation relies on temporally overlapping sounds from different sources being excluded from the stream of interest. This parsing process is known as auditory stream segregation, and is typically studied using sequences of pure tones alternating rapidly between low (L) and high (H) frequencies (e.g., Miller & Heise, 1950; Bregman & Campbell, 1971). For such stimuli, a larger frequency separation or a faster rate of presentation promotes stream segregation (e.g., Bregman & Campbell, 1971; van Noorden, 1975). More generally, any salient difference between sequentially presented sounds may lead to streaming (for reviews, see Moore & Gockel, 2002; 2012).

For an unchanging, repeating sequence of L and H tones, the likelihood of segregation increases over time (van Noorden, 1975). This "build-up" of stream segregation is most rapid over the first few seconds of a tone sequence but continues over intervals of at least a minute (Bregman, 1978; Anstis & Saida, 1985). However, even for very long tone sequences with large HL frequency separations, segregation is never heard exclusively; perception continues to switch between integration and segregation (e.g., Pressnitzer & Hupé, 2006; Denham & Winkler, 2006). On the basis of these findings, it has been suggested that the perception of stream segregation is bi-stable after build-up has occurred, and that build-up actually reflects a bias towards hearing a prolonged integrated percept at sequence onset. This alternative account of build-up, in terms of the bi-stability of stream segregation, is considered further in the General Discussion. Whatever the underlying mechanism, note that the experiments reported here used tone sequences ranging from a few to a few tens of seconds, for which build-up should continue throughout.

Rogers and Bregman (1993) studied another form of build-up, one that occurs in the absence of frequency alteration. Their stimuli comprised a relatively long induction sequence (4.8 s) followed immediately by a test sequence of three HLH– triplets (1.2 s). The properties of the test sequence were kept constant, and so differences in reported segregation were attributed directly to the effect of the inducer. An inducer comprising a repeating H-tone arrangement was highly effective at promoting segregation in the subsequent test sequence, despite the absence of frequency alternation. Such a constant-frequency (CF) inducer was most effective at promoting stream segregation when the H-tone density and number of onsets matched those for the H tones of the test sequence. Although this finding has been replicated in several studies (Beauvois & Meddis, 1997; Rogers & Bregman, 1998; Roberts, Glasberg, & Moore, 2008; Haywood & Roberts, 2010; 2011b), we are unaware of any direct comparison between the segregation-promoting effect of a CF induction sequence and the build-up of segregation that occurs during an on-going sequence of alternating-frequency (AF) tones.

Build-up can be "reset" following an abrupt change in sequence properties, such that integration is perceived once more. This resetting effect was originally observed in the context of AF inducers, but it has also been reported for CF inducers. Anstis and Saida (1985) presented a long, repeating LH tone sequence and altered its properties for 1 s (test sequence) after every 4 s of the standard sequence. Listeners adjusted the sequence rate to the point at which integration was heard; an adjustment to a slower rate was taken as evidence that another factor was promoting segregation. Either changing the ear of presentation, or applying a frequency offset of more than about 2 semitones to both the L and H tones, led to reduced segregation. Rogers and Bregman (1993) also demonstrated that a change in ear of presentation between AF induction and AF test sequences (both ears vs. right ear only) led to listeners reporting less segregation than when they were unchanged (right ear only). This was attributed to the change in intensity (left ear) and/or perceived lateralization at the induction/test-sequence boundary causing a resetting of the prior build-up of streaming. Recently, Kondo et al. (2012) have demonstrated that changes in lateralization cues can evoke resetting even if they arise from self-induced head motions, suggesting that stream segregation is directly influenced by a listener's "active sensing" of their environment, such as orienting the head towards relevant acoustic stimuli.

Rogers and Bregman (1998) presented a repeating HLH– induction sequence and subsequent test sequence to both ears, and found similar resetting effects for sudden changes in loudness (level), perceived lateralization (interaural time difference, ITD), or perceived location (via a loudspeaker array). Resetting effects largely consistent with these findings were observed in the context of CF inducers by Roberts et al. (2008), who used performance at detecting a delay in the onset of the H tones as an index of stream segregation (cf., Vliegen, Moore, & Oxenham, 1999; Cusack & Roberts, 2000; Roberts, Glasberg, & Moore, 2002). Rogers and Bregman (1998) offered two possible explanations for the resetting observed following a change in sequence properties. First, build-up may fail to transfer between sequences with different tonal characteristics. Second, the change may itself trigger "active" resetting. More specifically, they proposed that a sudden change may be interpreted as evidence that a new event has occurred, which in turn triggers a re-analysis of the entire auditory scene. There is some evidence in favor of this account, as Rogers and Bregman (1998) found that the magnitude of resetting following a level change depended on the direction of the change – a soft-to-loud transition had a much greater resetting effect than a loud-to-soft transition. They reasoned that an increase in level could signify a new event, whereas a decrease in level could not.

Haywood and Roberts (2010) provided further evidence in support of an active resetting mechanism. A "standard" induction sequence of ten L tones was used to promote subsequent segregation (test sequence = 3 LHL– cycles). Resetting was measured when *only* the final induction tone was altered on some dimension (frequency, level, duration, or replacement with silence). Each type of "deviant" tone tended to reset the build-up of stream segregation compared with the standard induction case. This effect was often substantial, typically a loss of build-up of between one and two thirds relative to the standard (i.e., 0.0) and no-inducer (i.e., 1.0) cases, although the trend towards resetting following the level change failed to reach significance. Given that the nine tones preceding the deviant were unchanged relative to the standard inducer case, the substantial reduction in segregation observed could not be attributed simply to a failure of the deviant tone to contribute to the build-up occurring during the induction sequence. These findings were interpreted as evidence that a single change to an on-going sequence can actively reset the build-up of stream segregation (see also Haywood, 2009; Haywood & Roberts 2011a). In addition, Haywood and Roberts (2010) found that relatively small changes in frequency could trigger resetting, and so concluded that resetting does not require a change in magnitude likely to be perceived as a new source. The idea that a single deviation to a sequence can trigger resetting is also consistent with Cusack et al.'s (2004) finding that reported stream segregation in an on-going LHL– sequence was greatly reduced following a silent gap. The extent of this resetting did not vary significantly with gap duration (range tested = 1–10 s), and so they concluded that resetting must occur much more rapidly than the build-up of the tendency to hear segregation.

As noted above, to our knowledge the dynamics of the build-up in the tendency to hear stream segregation in a subsequent test sequence have not been examined before in the context of CF inducers. Despite the lack of systematic study, it has been suggested that the segregation-promoting effect of a CF inducer may reflect a process different from the build-up that occurs during an AF sequence. Rogers and Bregman (1993) proposed that a CF inducer may promote segregation by capturing a subset of test-sequence tones into the perceptual stream already formed from the induction sequence. An alternative suggestion by Thompson et al. (2011) is that exposure to a CF inducer may bias listeners to attend the novel subset of tones in the test sequence, and that this bias may promote segregation. They also suggested that segregation following a CF inducer may be due to selective adaptation of neurons tuned to the frequency of the inducer tones. In the light of these suggestions, the findings of Haywood and Roberts (2010) concerning the extent to which a single deviant tone brings about resetting may not necessarily generalize from CF to AF inducers. The three experiments reported here attempt to elucidate the dynamics of build-up and resetting in these two different stimulus contexts.

EXPERIMENT 1

This experiment investigated further the resetting effect of including a single deviant in the induction sequence by manipulating its serial position relative to the boundary between the inducer and the test sequence. If a deviant tone actively resets a cumulative build-up process, resetting should be most evident when the deviant tone occurs later in the induction sequence, and hence closer in time to the test sequence.

Method

Listeners

Eight listeners (3 males, mean age = 24.6 years, SD = 3.9) took part in Experiment 1; all reported normal hearing. Four of these listeners had previous experience of stream-segregation experiments. This research was approved by the Aston University Ethics Committee.

Stimuli and conditions

Each trial sequence comprised a combination of an induction sequence and a subsequent test sequence. The test sequence comprised three LHL– triplets. The L tones were set to 1 kHz, and the H tones were 4, 6, 8, 10, 12, or 14 ST higher in frequency (1260, 1414, 1587, 1782, 2000, or 2245 Hz, respectively). All tones were presented diotically and at 70 dB SPL. Each tone was 100-ms long (including 10-ms raised cosine ramps at onset and offset); the silence at the end of each triplet was also 100-ms long. Hence, the duration per triplet cycle was 400 ms and the total duration of the test sequence was 1.2 s. These test sequences are identical to those used by Haywood and Roberts (2010, 2011b).

There were five induction conditions. The standard induction sequence comprised 10 L-tone repetitions, for which the duration and timing of the L tones was identical to those of the test sequence (i.e., the silent intervals between successive L tones were 100 ms long). Note that two L-tone repetitions corresponds to one triplet cycle for which the H tone has been replaced by silence. The standard-induction condition was included as a control to measure the maximum segregation-promoting effect of an unaltered, CF induction sequence. A second control, the no-induction condition, was included to measure test-sequence streaming in the absence of any prior induction sequence (and hence of any build-up). The three experimental conditions used modifications of the standard induction sequence; for each, a single tone was replaced with a silent interval of equivalent duration. This "silent deviant" could replace the 4^{th}, 7^{th}, or the 10^{th} (i.e., the last) tone of the induction sequence. For each of these conditions, the properties of all the other induction tones were identical to those of the standard induction case, and so the inclusion of a silent deviant resulted in a 300-ms silence between the two adjacent L tones. A schematic illustrating the set of induction conditions is shown in Figure 1. Note that the standard-induction, no-induction, and 10^{th}-silent induction conditions were all exact replications of conditions tested by Haywood and Roberts (2010).

Procedure

On each trial, a combination of an induction sequence (or no inducer) followed by a test sequence was presented once. Listeners were instructed to report their perception of the final LHL– triplet of the test sequence. They were asked to avoid trying to hear either integration or segregation, but instead to report which percept was more dominant. Listeners responded via a computer keyboard to indicate a perception of either “one stream” (i.e., integration) or “two streams” (i.e., segregation). To reduce the possibility of errors, listeners were required to confirm their response by pressing “enter”, after which there was a 3 s pause before the next trial began automatically. This pause was to ensure that any build-up of segregation would decay before the onset of the next trial (cf. Bregman, 1978). For the main experiment, trials were presented in 20 blocks, each comprising a combination of all five induction conditions with all six HL frequency separations for the test sequence (i.e., 30 trials per block). For each listener, the order of trial presentation within a block was randomized anew for each block. After an initial explanation of the task, listeners were first presented with examples of clearly integrated and segregated test sequences. Listeners then completed a brief training session, comprising two trial blocks only.

All stimuli were synthesized with 16-bit resolution using MITSYN (Henke, 1997), and played back via a Turtle Beach Santa Cruz sound card at 20 kHz sampling rate. The stimuli were presented over Sennheiser HD480-13II earphones; the overall output level of the sound card was set using the on-board analogue attenuator for coarse adjustment and digital multiplication for fine adjustment. The setup was calibrated using a sound-level meter (Brüel & Kjaer, type 2209) coupled to the earphones by an artificial ear (type 4153). Listeners completed the experiment either in a double-walled sound-attenuating chamber (Industrial Acoustics 1201A) or in a single-walled chamber (Industrial Acoustics 401A) that was housed within a quiet room.

Results

For each listener, the percentage of trials on which stream segregation was reported was computed for each condition. Responses across all listeners were averaged to give an overall indication of streaming for each condition; these mean values are shown in Figure 2. The mean percentages reported as segregated for the six frequency separations were: 4 ST = 1.3%, 6 ST = 10.4%, 8 ST = 30.3%, 10 ST = 57.3%, 12 ST = 63.9%, and 14 ST = 88.5%. The means for the five induction conditions were: standard = 51.0%, no-inducer = 23.0%, 4^{th}-silent = 51.1%, 7^{th}-silent = 50.2%, and 10^{th}-silent = 34.2%.

Data were analyzed using a two-way, repeated-measures analysis of variance (ANOVA). The ANOVA confirmed significant main effects of both frequency separation [$F(5,35)=49.07$, $p<0.001$, $\eta^2_p=0.87$] and induction condition [$F(4,28)=37.40$, $p<0.001$, $\eta^2_p=0.84$]. There was also a significant interaction between these two variables [$F(20,140)=9.75$, $p<0.001$, $\eta^2_p=0.58$]. Two-tailed pairwise comparisons were conducted using the restricted least-significant-difference test (Snedecor & Cochran, 1967; Keppel, 1991). Floor and ceiling effects were observed at frequency separations of 4 and 14 ST, and so these data were excluded from the pairwise comparisons. There was a substantial and significant difference in reported segregation between the standard and no-induction conditions [difference in percentage points[1] = 37.8%; $t(7)=9.17$, $p<0.001$]. There was also a substantial and significant difference between the standard induction and 10^{th}-silent conditions [23.1%; $t(7)=4.99$, $p<0.001$]. In contrast, the difference between the standard and 4^{th}-silent conditions [0.6%; $t(7)=0.19$, $p>0.05$], or between the standard and 7^{th}-silent conditions [1.1%; $t(7)=0.34$, $p>0.05$], was negligible. Furthermore, reported segregation for both the 4^{th}- and 7^{th}-silent conditions was significantly different from that for the 10^{th}-silent condition ([4^{th} vs. 10^{th} = 22.5%; $t(7)=4.11$, $p<0.005$]; [7^{th} vs. 10^{th} = 22.0%; $t(7)= 4.62$, $p<0.005$]). Finally, the loss of build-up associated with replacing the final inducer tone with silence was not complete, as the remaining difference from the no-inducer case was substantial and significant [14.7%; $t(7)=8.77$, $p<0.001$].

In summary, these results show that the standard induction sequence was effective at promoting stream segregation, and that replacing the 4^{th} or 7^{th} induction tone with silence had no appreciable impact on this effectiveness. Reported segregation was reduced substantially only when the 10^{th} (the last) induction tone was replaced with silence.

Discussion

Listeners' responses to the test sequences were consistent with the known effect of frequency separation on stream segregation (e.g., van Noorden, 1975). As expected (Haywood & Roberts, 2010, 2011b), the standard induction sequence had a strong segregation-promoting effect, and the least streaming was observed for the no-induction condition. Also consistent with the results of Haywood and Roberts (2010), reported segregation was significantly reduced compared with the standard induction case when the 10^{th} (final) induction tone was replaced with silence. As well as considering the direct change in the percentage of trials reported as segregated, Haywood and Roberts (2010) also calculated an "extent of resetting" measure for the "silent" (and other) deviant conditions. This measure reflected a proportional shift in the pattern of responses away from that for the standard induction condition and towards that for the no-induction condition (i.e., standard induction = 0.0, no-induction = 1.0). Haywood and Roberts (2010) reported a mean extent of resetting following the silent deviant of 0.65 (when averaged across test-sequence frequency separations of 6, 8, 10, and 12 ST). In the current study, the extent of resetting for the same frequency separations following the 10th-silent deviant was 0.68.

If reduced stream segregation in the nine-tone induction conditions was merely a result of reduced overall build-up compared with the standard ten-tone case, one might predict a similar reduction irrespective of the serial position of the tone that was replaced with silence. Instead, reduced segregation was only observed when the final (10th) induction tone was replaced with silence; the same change applied to either the 4th or 7th induction tone did not lead to any decrease in reported segregation. This outcome, and the magnitude of the effect, confirms that the observed reduction in segregation is not due merely to the presence of fewer induction tones. Given that the decrease in segregation following the silent deviant is too large to be explained as a gradual decay of the tendency to hear stream segregation (Bregman, 1978; Beauvois & Meddis, 1997), we conclude that the silent tone at the inducer/test boundary must act to reset build-up. Indeed, Haywood and Roberts (2010) also demonstrated resetting for cases where the final inducer was altered on one acoustic dimension (as opposed to being replaced by silence). The current finding that reduced segregation cannot be explained simply by the presence of fewer standard inducer tones can also be generalized to conditions in which the deviant remains present but is altered in its acoustic properties. Hence, the current results support the hypothesis that a noticeable deviation to a sequence must actively reset the build-up of stream segregation.

It is perhaps surprising that there was *no* evidence of resetting when either the 4th or the 7th induction tone was replaced with silence. One possible explanation is that some resetting did take place in those cases, but that this was obscured by build-up re-occurring during the remainder of the induction sequence. If so, this recovery must have occurred rapidly, as only three tones were present after the deviant in the 7th-silent case. This may be an indication that the rate of build-up during the CF induction sequence is much more rapid than that which occurs during an AF sequence.

EXPERIMENT 2

This experiment investigated the rate at which a CF induction sequence promotes subsequent stream segregation by manipulating the number of tones in the induction sequence prior to the onset of the test sequence.

Method

Eight listeners (6 males, mean age = 24.9 years, SD = 5.0) took part, all of whom reported normal hearing. None of them had taken part in Experiment 1 or in any of our other auditory streaming experiments (Haywood & Roberts, 2010; 2011a; 2011b). This experiment included the standard- and no-induction conditions that were used in Experiment 1. The induction sequences for the three experimental conditions differed from the standard case in that they contained 6, 3, or 1 L-tone repetition(s). To preserve the overall length of each induction sequence at 2 s, the missing tone repetitions were replaced by filling the initial portion of the stimulus with a continuous band-pass filtered noise of matching duration. The noise continued until 100 ms before the onset of the first remaining L tone. Specifically, six L-tone repetitions lasted for 1.2 s, and so the noise for the 6-inducers condition was set to 0.7 s. Similarly, three repetitions lasted for 0.6 s (3-inducers condition: noise = 1.3 s), and one repetition lasted for 0.2 s (1-inducer condition: noise = 1.7 s). Fig. 3 illustrates these induction conditions.

The noise used was centered on 1 kHz and filtered with a bandwidth of 4 ST (pass-band = 891 Hz - 1122 Hz). Band-pass noise was used rather than broadband noise in order to focus the stimulation in the frequency region around the L tones; nonetheless, this stimulus sounded more noise-like than tonal, owing to its third-octave bandwidth. The noise was created digitally by combining sinusoids with random starting phases distributed at 2-Hz intervals; these sinusoids were equal in amplitude across the pass-band but were progressively attenuated outside (spectral roll-off = 80 dB/oct). As for the pure tones, the noise had 10-ms raised cosine ramps at onset and offset, and the steady-state portion of the noise was presented at 70 dB SPL. Continuous white noise is known not to induce subsequent stream segregation (Bregman, 1978; Rogers & Bregman, 1993, 1998), and a pilot study indicated that the same was true for the narrower noise band used here. More generally, any long, continuous stimulus (even a pure tone) with only a single onset and offset prior to the test sequence is ineffective at promoting subsequent stream segregation (e.g., Rogers & Bregman, 1993; Roberts et al., 2008; Haywood & Roberts, 2011b). Trials were organized in the same way as for Experiment 1 (5 induction conditions × 3 frequency separations = 15 trials per block), and the same exemplars of integration and segregation were used. The training and main experiment comprised two and 20 blocks of trials, respectively.

Results

Responses from all listeners were averaged, and the mean percentages of trials heard as segregated are shown in Fig. 4. The percentage of trials reported as segregated for the six frequency separations were: 4 ST = 2.0%,6 ST = 27.6%, 8 ST = 61.5%, 10 ST = 80.5%, 12 ST = 89.2%, and 14 ST = 96.6%. The means for the five induction conditions were: standard = 64.2%, no-induction = 45.2%, 6-inducers = 67.2%, 3-inducers = 65.3%, and 1-inducer = 56.0%. A two-way, repeated-measures ANOVA confirmed significant main effects of frequency separation [$F(5,35)=56.71$, $p<0.001$, $\eta^2_p=0.89$] and induction condition [$F(4,28)=7.69$, $p<0.001$, $\eta^2_p=0.52$]. There was also a significant interaction [$F(20, 140)=3.59$, $p<0.001$, $\eta^2_p=0.34$].

The standard induction, 6-inducers, and 3-inducers conditions all promoted a strong and broadly similar tendency for listeners to report stream segregation. For the 1-inducer condition, reported segregation was roughly midway between that for the standard and no-induction cases, indicating that the segregation-promoting effect of a single inducer tone was about half the size of that evoked by a standard ten-tone induction sequence. Pairwise comparisons showed that there was a significant difference between the standard and no-induction conditions [difference in reported segregation (percentage points) = 26.1%; $t(7)=2.73$, $p<0.05$], but that the difference between the standard and 3-inducers conditions was not [1.9%; $t(7)=0.51$, $p>0.05$]. Although the difference between the standard and 6-inducers conditions was significant, it was small and in the opposite direction to the predicted change [4.1%; $t(7)=2.57$, $p<0.05$]. Despite the larger difference in means, corresponding to a reduction in segregation of about a half, the difference between the standard and 1-inducer conditions was not significant [12.5%; $t(7)=1.54$, $p>0.05$]. This outcome is a consequence of the greater variability across listeners observed for the 1-inducer condition. Nonetheless, all four L-tone induction conditions promoted significantly more stream segregation than the no-induction case ($p<0.05$ in all cases).

Discussion

The induction sequences comprising either six or three L-tone repetitions were broadly as effective at promoting segregation as the standard, ten-tone case. Indeed, the mean percentages of trials heard as segregated following the 3- and 6-inducers conditions were actually a little higher than for the standard induction case. Furthermore, even the induction sequence containing only a single L-tone repetition increased the perception of stream segregation with respect to the no-induction condition – albeit to a lesser extent (~half) than all of the other induction conditions. Given that, statistically, the six-tone induction sequence promoted more stream segregation than did the standard, ten-tone case, one might speculate that the preceding band-pass filtered noise itself had some small segregation-promoting effect in the experimental induction conditions. However, we contend that this explanation is unlikely, given that adding a single extended sound has previously been shown to be largely – if not entirely – ineffective at promoting test-sequence segregation (Bregman, 1978; Rogers & Bregman, 1993, 1998; Haywood and Roberts, 2011b), irrespective of whether the sound is narrowband (pure tone) or wideband (≥1-octave-wide noise).

In Experiment 1, no evidence of resetting was observed when either the 4th or 7th tone of a ten-tone induction sequence was replaced with an equivalent-duration silent interval. Note, however, that this does not necessarily mean that changes at these serial positions did not cause resetting, because the current results indicate that build-up during the subsequent induction tones can occur rapidly enough to obscure such an effect. Hence, it is difficult to determine whether or not resetting occurs when a deviant tone is included in an early serial position in the induction sequence. Induction sequences containing ten, six, or three L-tone repetitions all promoted a similar degree of reported segregation; this contrasts strongly with what is known about the rate of build-up during an on-going, AF tone sequence (Bregman, 1978; Anstis & Saida, 1985). For such stimuli, build-up is typically progressive over the first 5-10 s after the onset of the sequence. Subsequently, the rate of build-up becomes more gradual but still remains apparent over the full course of a 60-s sequence (Anstis & Saida, 1985). In contrast, it appears that the cumulative segregation-promoting effect of a CF induction sequence has very different dynamic properties. The current results could indicate that the segregation-promoting effect of a CF induction sequence reaches a maximum level after only three rapid L-tone repetitions, and that subsequent

L tones do not promote any further increase in segregation. However, it is unknown whether stream segregation would increase further if more than ten L-tone inducers were presented.

Rogers and Bregman (1993) noted that a single integrated stream must always be perceived during a CF induction sequence, as there is no viable alternative perceptual representation. This pre-established stream may then be maintained during the test sequence, provided that there is a good continuation of tonal properties between the induction tones and a subset of tones in the test sequence. More generally, these authors suggested that once a distinct percept emerges from an auditory scene, properties derived from that percept are fed back to control the on-going analysis of that auditory scene. In the current experiments, an L-tone stream must be heard during the induction sequence, and so the L tones of the test sequence would likely be 'captured' into this on-going stream (cf. Bregman & Rudnicky, 1975). In contrast, the H tones would be less likely to be integrated into this stream owing to their different frequency. If so, the likelihood of H-tone segregation should increase when an L-tone stream has been established during the prior induction sequence. This hypothesis is capable of explaining the current results, as it may require only relatively few repetitions of an inducer tone before the corresponding stream is established.

The proposal that a CF inducer captures the corresponding tones in the test sequence into a pre-established stream is supported by evidence from electrophysiological mismatch negativity (MMN) studies. The MMN can be elicited by any noticeable change in a repetitive auditory stimulus, and is thought to reflect a deviance-detection process which is based on memory of the regularities occurring in an acoustic stimulus (see Näätänen et al., 2001, for a review). Importantly, the MMN component correlates strongly with perceptual organization (Sussman et al., 1999; 2005; 2007; Ritter et al., 2000; Winkler et al. 2003). A deviant sound will only elicit an

MMN component if it has been preceded by a minimum of three repetitions of an identical sound (Cowan et al., 1993; Winkler et al., 1996). This is taken as evidence that at least three repetitions of a regular sound are required before the pattern is firmly established in memory.

In the context of the CF inducers used here, comprising three or more L-tone repetitions, the first H tone of the test sequence is likely to be interpreted as a novel event, deviating from the established pattern of the preceding inducers. There is behavioral evidence suggesting that deviants which elicit an MMN response may capture attention involuntarily (Schröger, 1996; Schröger & Wolff, 1998; Parmentier, 2008). Hence, the first H tone of the test sequence may have a greater attention-capturing effect when it is introduced after a CF inducer (Thompson et al., 2011). This may cause the listener to maintain attention on the subsequent H tones, and so promote stream segregation. According to this hypothesis, the listener switches attention from the L-tone-only stream (heard during the induction sequence) to the novel H tones of the test sequence. Note, however, that it is unclear how best to reconcile this suggestion with the body of evidence that a shift in attention can trigger resetting and so reduce stream segregation (Carlyon et al., 2001, 2003; Cusack et al. 2004; Thompson et al., 2011; see also Haywood & Roberts, 2010). Further research is needed to evaluate whether deviation-based attentional capture can promote subsequent stream segregation.

EXPERIMENT 3

This experiment was designed to directly compare the segregation-promoting effect of a CF induction sequence with the build-up that occurs during an on-going AF sequence. This was achieved in two ways – first by measuring test-sequence segregation following both CF and AF induction sequences; second, the test sequence was extended to 50 LHL– triplets (20 s) and listeners were asked to continuously report their perception. Hence, the extent of stream segregation heard immediately after the inducer could be compared with that heard later in the test sequence, where any effect of the inducer should have diminished. Another aim of Experiment 3 was to measure deviant-tone resetting in an on-going AF sequence, as to date this type of resetting has only been tested and demonstrated when the deviant occurs at the end of a CF induction sequence (Haywood & Roberts 2010; 2011a).

Method

Listeners

Twelve listeners (6 males, mean age = 24.5 years, SD = 5.6) successfully completed Experiment 3; all reported normal hearing. One listener was excluded from the final data set and replaced, because his responses did not conform to the well-established effects of pure-tone frequency separation on streaming judgments, or show any signs of build-up (van Noorden, 1975; Anstis & Saida, 1985). Five of the twelve accepted listeners had either taken part in Experiment 2, or had previous experience of our other auditory streaming studies (Haywood & Roberts, 2010; 2011a; 2011b).

Stimuli and conditions

As before, each trial comprised a combination of an induction sequence and a test sequence, but in this experiment the test sequence was extended to 50 LHL– cycles. The properties of the individual tones and silences were identical to those used previously, except that a frequency separation of 3 ST, 6 ST, or 9 ST was used (L tone = 1 kHz; H tone = 1189, 1414, or 1682 Hz, respectively). Given that the overall duration of each LHL– triplet was 0.4 s and there were 50 triplets in total, the test sequence was 20 s long.

Five different induction sequences were tested, all of which were 2.0 s long (see Fig. 5). The "silent" induction condition was intended as a measure of test-sequence streaming in the absence of any prior build-up. Only a single L tone was presented at the onset of the stimulus; the remaining portion of the induction sequence was filled by 1.9 s of silence. The single L tone acted as a "warning tone", to help listeners anticipate the onset of the test sequence. Note that any segregation-promoting effect of this short tone (cf. the one-inducer condition in Experiment 2) should decay completely during the subsequent silent interval (Bregman, 1978; Cusack et al., 2004). In effect, the silent-inducer condition differs from the no-inducer condition used in Experiments 1 and 2 only in that it preserves the 2 s interval between inducer onset and test-sequence onset. The CF induction condition was identical to the standard-induction condition used in Experiments 1 and 2 (i.e., 10 L-tone repetitions were used). The CF-deviant induction condition differed from the CF condition only in that the last L tone was extended in duration from 100 ms to 150 ms, and the subsequent inter-tone silence was reduced from 100 ms to 50 ms, in order to preserve the regular onset-to-onset time between successive L tones. Haywood and Roberts (2010) tested an identical deviant-tone arrangement, and observed a substantial resetting effect for a subsequent, short test sequence (1.2 s).

For the AF induction condition, a sequence of five LHL– triplets was used (i.e., the equivalent of 10 L-tone repetitions). These triplets were identical to those presented in the test sequence (L tones = 1 kHz; H tones = 3, 6, or 9 ST higher, chosen to match the properties of the test sequence). Note also that all properties of the L tones – including tone density and timing – were identical to those of their counterparts in the CF induction condition (see Rogers & Bregman, 1993). Hence, the induction and test sequences were seamless and so the induction/test-sequence boundary was defined arbitrarily in the AF induction condition. It was assumed that build-up would occur at the same rate during the (arbitrary) induction sequence as it would during the test sequence. The AF-deviant induction condition was identical to the AF condition, except that the final L tone of the last (fifth) inducer triplet was extended in duration from 100 ms to 150 ms (i.e., as for the deviant tone in the CF-deviant condition). Note that using an extended tone, rather than replacement with silence, in the deviant conditions preserves the pattern of pitch changes within the final LHL– triplet of the AF-deviant inducer. The AF-deviant condition was included to measure the resetting effect of a single deviant tone on the build-up of stream segregation in an on-going, AF tone sequence. In this case, the end of the triplet containing the deviant tone defined the inducer/test boundary.

Procedure

On each trial, a single combination of an induction sequence and a test sequence was presented. Each trial was initiated 1 s after the listener pressed “enter” on the computer keyboard. A visual cue indicated the start of the test sequence and prompted listeners to begin responding. Listeners indicated as soon as they could whether they were hearing integration (one stream) or segregation (two streams) by pressing either “1” or “2”, respectively. During the rest of the test sequence, listeners were asked to press the appropriate key every time their perception of the test sequence changed. They were asked to avoid listening actively for either integration or segregation, but simply to report which of the two percepts was more dominant at that moment. At the end of each sequence, there was a 5-s pause before listeners could begin the next trial, to allow for any build-up to decay before the onset of the next trial (cf. Bregman, 1978).

Trials were organized into blocks in a manner consistent with the previous experiments (i.e., 5 induction conditions × 3 frequency separations = 15 trials per block). The training session comprised a single trial block and the main experiment comprised 10 trial blocks. The two parts together typically took 2 – 2½ hours to finish, and so the experiment was completed over two separate sessions. The apparatus and set-up used were identical to those for Experiments 1 and 2, except that the current experiment was run using the Media Control Functions (MCF) stimulus presentation software (Ahad, 2000). This software supports the precision measurement of the timing of key presses necessary for accurate estimates of the extent of stream segregation.

Results

Response data from each trial were divided into twenty 1-s time bins (i.e., 0-1 s, 1-2 s . . . 19-20 s). For each time bin, the percentage of time during which the test sequence was heard as segregated was calculated from the timings of individual key presses. This percentage was recorded only if the listener's initial response had occurred before the current time bin or within the first 0.5 s of that time bin. For the 0-1 s time bin, only 15% of all trials met this criterion (compared with 75% for the 1-2 s time bin). Owing to the limited data available, the 0-1 s time bin was excluded from all subsequent analysis and graphical representation; all other time bins were included. For each listener, the data for a given time bin were averaged across trial blocks separately for each combination of induction condition and frequency separation. Each mean value was calculated only from the trials for which the time bin met the acceptance criteria described above. Finally, the data were averaged across the twelve listeners to yield an overall percentage heard as segregated as a function of time for each of the induction conditions. These data are displayed in Fig. 6; a separate panel is used for each frequency separation.

The greatest differences between induction conditions were evident during the first ~11 s of the test sequence. A three-way, repeated-measures ANOVA conducted on the first 10 s of response data available for analysis (frequency separation × induction condition × time interval: *time bins 1-2 s to 10-11 s, inclusive*), confirmed significant main effects of frequency separation

[$F(2,20)=39.16$, $p<0.001$, $\eta^2_p=0.80$], induction condition [$F(4,40)=13.75$, $p<0.001$, $\eta^2_p=0.58$], and time interval [$F(9,90)=28.33$, $p<0.001$, $\eta^2_p=0.74$]. Clearly, all three factors influenced stream segregation during the first half of the test sequence – segregation was greater for larger frequency separations, tended to change over time (usually increased), and tended to be greater for CF-type than for AF-type inducers. Each two-way interaction term was also significant ([frequency separation × induction condition: $F(8,80)=2.83$, $p<0.01$, $\eta^2_p=0.22$], [frequency separation × time interval: $F(18,180)=7.11$, $p<0.001$, $\eta^2_p=0.42$], [induction condition × time interval: $F(36,360)=8.86$, $p<0.001$, $\eta^2_p=0.47$]), and so was the three-way interaction term [$F(72,720)=4.35$, $p<0.001$, $\eta^2_p=0.30$]. The origin of these interactions is evident in Fig. 6 – changes over time in stream segregation depend not only on the induction condition but also on the extent of frequency separation in the test sequence. Note also that all three main effects, and all interactions except frequency separation × time interval, remained significant when the silent-inducer condition was excluded from the analysis. This shows that the significant effects found in the main analysis were not simply due to the inclusion of the silent-inducer condition.

A similar analysis (including the silent-inducer case) was conducted on the response data for the final 9 s of the test sequence (frequency separation × induction condition × time interval*: time bins from 11-12 s to 19-20 s, inclusive*). This confirmed significant main effects of frequency separation [$F(2,22)=6.83$, $p<0.005$, $\eta^2_p=0.38$] and time interval, [$F(8,88)=2.73$, $p<0.01$, $\eta^2_p=0.20$]. However, in contrast with the analysis for the first half of the test sequence, there was no significant main effect of induction condition [$F(4,44)=0.49$, $p>0.05$, $\eta^2_p=0.04$]. These outcomes indicate that: 1) though more slowly, reported stream segregation continued on average to rise in the latter portion of the test sequence; 2) the extent of reported segregation continued to be influenced by frequency separation, even after the period of most substantial change in the tendency to hear two streams was over; 3) the properties of the induction sequence did not influence responses in the latter portion of the test sequence, presumably because the effects of the inducer on responses to the test sequence had largely dissipated. The interaction term frequency separation × time interval was also significant [$F(16,176)=9.09$, $p<0.001$, $\eta^2_p=0.45$]; this primarily reflects the greater tendency for build-up to continue increasing for the 3-ST frequency separation. No other interaction term was significant ($p>0.3$ in all cases).

Before the outcomes for the different induction conditions are compared in detail, the relationship between the silent- and AF-inducer cases merits consideration. The AF induction sequence was identical to the subsequent test sequence, and lasted for 2 s. If, as expected, the act of making an initial response part way through an on-going sequence does not affect streaming judgments, then mean stream segregation in the AF condition should be very similar to that occurring 2 s later in the silent-inducer condition. To test this, the AF data were offset by +2 s and compared with the silent-inducer case. These data are displayed in Figure 7. A three-way ANOVA confirmed significant main effects of frequency separation [$F(2,22)=22.518$, $p<0.001$, $\eta2_p=0.672$], and time interval [$F(16,176)=59.685$, $p<0.001$, $\eta2_p=0.844$]. Crucially, there was not a significant main effect of induction condition [$F(1,11)=1.422$, $p>0.05$, $\eta2_p=0.114$], and none of the interaction terms involving the induction-condition variable were significant. As anticipated, this outcome also supports the view that the warning tone at the onset of trials in the silent-inducer condition had little or no effect on subsequent judgments of streaming.

Having established a significant main effect of induction condition for the first half of the test sequence, including when the silent-inducer condition was removed from the analysis, we now consider the differences between induction conditions. These differences were explored using the results from the first time-bin available for analysis (1-2 s), for which the pattern of responses should be most affected by the properties of the induction sequence. The mean percentages of stream segregation for the three frequency separations were: 3 ST = 10.3%, 6 ST = 41.5%, and 9 ST = 69.4% (averaged across induction conditions). The means for the five induction conditions were: silent = 20.5%, CF = 62.3%, CF-deviant = 49.7%, AF = 33.6%, and AF-deviant = 35.9% (averaged across frequency separations). A two-way, repeated-measures ANOVA (frequency separation × induction condition) confirmed significant main effects of frequency separation [$F(2,22)=75.73$, $p<0.001$, $\eta 2p=0.87$], and of induction condition [$F(4,44)= 20.89$, $p<0.001$, $\eta_{2p}=0.655$]. The interaction between these two variables was also significant [$F(8,88)=8.16$, $p<0.001$, $\eta_{2p}=0.426$]. Most probably, this interaction was driven by a partial floor effect (i.e., primarily integrated responses) for the 3-ST cases. Note that the significance levels reported for the two main effects and the interaction term remained high ($p<0.001$) when the silent-inducer condition was removed from the analysis.

One key purpose of the experiment was to examine the relative ability of comparable CF and AF inducers to promote subsequent stream segregation. Pairwise comparison showed that the CF inducer promoted substantially more segregation at test-sequence onset than did the AF inducer [28.6%; $t(11) = 6.52$, $p<0.001$]. Another key purpose of the experiment was to examine the resetting effect of a single deviant tone in the context of both the CF- and AF-type induction sequences. When a CF-type inducer was used, stream segregation for the first time bin decreased by 10% or more when the final inducer was extended in duration; this effect was significant [12.6%; $t(11)=2.33$, $p<0.01$]. For the AF-deviant condition, the same duration increase was applied to the deviant tone, but this had only a small and inconsistent effect on initial streaming judgments across the frequency separations tested. The analysis reflected this; overall there was a small nominal increase in mean segregation from the AF to the AF-deviant condition, but this difference was not significant [2.3%; t(11)=0.066, p>0.05]. Finally, there was a significant difference between the CF-deviant and AF-deviant conditions [13.8%; t(11) = 4.61, p<0.001].

We have already established that initial stream segregation was significantly greater in the CF condition than in the AF condition. In addition, reported segregation in the CF condition appears actually to decrease over time for frequency separations of 6 or 9 ST. Paired-sample t-tests were run to determine whether this trend was significant. For each frequency separation, the t-tests compared initial segregation (1-2 s bin) with that 10 s later (10-11 s bin). For the CF inducer, stream segregation *increased* significantly over time for the 3-ST case [+31.4%; t(11)=3.275, p<0.01] but *decreased* significantly over time for the 9-ST case [-12.0%; t(11)=2.572, p<0.05]; the change for the 6-ST case was not significant [-4.2%; t(11)=0.707, p>0.05]. For the AF inducer, reported segregation always increased from the 1-2 s time bin to the 10-11 s time bin. This effect was significant for frequency separations of 3 ST [+40.0%; t(11)=6.406, p<0.001] and 6 ST [+36.6%; t(11)=7.424, p<0.001], though not for the largest frequency separation tested [9 ST: +9.7%; t(11)=1.226, p>0.05].

Discussion

The silent-inducer condition was not expected to have an appreciable influence on the perception of the test sequence, and so a typical pattern of build-up was predicted. Similarly, the properties of the AF inducer were exactly the same as those of the subsequent test sequence, and so build-up was expected to occur at the same rate as during the test sequence. This prediction was confirmed when the data were offset by 2 s, to account for the build-up occurring during the (arbitrary) induction sequence. For both conditions, the dynamics of build-up during the test sequence were broadly consistent with the findings of previous studies (Bregman, 1978; Anstis & Saida, 1985; Carlyon et al., 2001). More specifically, the rate of build-up over the first ~10 s increased with frequency separation, which is in good accord with previous research (e.g., Anstis & Saida, 1985). Also, the overall tendency to report segregation (or the percentage of trials heard as segregated near asymptote) rose as frequency separation was increased – from ~60 % two-stream percepts for the 3-ST case to ~75% – 80% for the 9-ST case. While this trend is not quite as pronounced as that observed in some other studies (e.g., Carlyon et al., 2001), the current results are largely consistent with those of Cusack (2005), who observed only a modest increase in stream segregation near asymptote across the range of frequency separations 3–7 ST.

Perhaps the most notable finding of the current experiment is that the CF induction sequence promoted substantially more stream segregation over the first several seconds of the test sequence than did the AF inducer. Indeed, following the CF inducer, the tendency to report segregation actually *decreased* over the course of the test sequence for the largest frequency separation tested (9 ST). What might account for this difference in outcomes between the CF and AF induction conditions? The results of Experiment 2 suggested that a CF inducer may promote segregation by capturing one subset of test-sequence tones into an on-going stream established during the induction sequence (Rogers & Bregman, 1993; see also Bregman & Rudnicky, 1975). Given that as few as three tones may be required to establish an L-tone stream in the CF condition, the tendency to hear a separate H-tone stream when the test sequence begins is likely to be strong, particularly when the HL frequency separation is large. Presumably, this effect decays over several seconds, during which the classical build-up associated with the rapid alternation of L and H tones also develops. By this account, it is the changing balance between these factors over time that shapes the response profiles for the CF conditions across different frequency separations.

Sussman and Steinschneider (2006) presented results consistent with our finding that a CF sequence has a greater segregation-promoting effect than an AF sequence. They measured the segregation-promoting effect of a CF induction sequence by using the MMN as an indication of stream segregation in a task in which listeners were not required to attend the tone sequence. The authors presented a test sequence comprising 4 LHHH repetitions. The level of each H tone was randomized (from 67 to 87 dB SPL), whereas the L tones were fixed at 71 dB SPL. 10% of the L tones were "probes", which were increased to 83 dB SPL. Generally, for this sequence arrangement, no intensity regularity would be heard when the L and H tones formed a single stream. Therefore, when the sequence was heard as integrated, the probe tone would not be expected to elicit an MMN component (as an established regular pattern is a pre-requisite for MMN elicitation). In contrast, for a segregated percept, the probe should be heard as a deviation from the otherwise regular intensity of the L-tone-only stream, and so should elicit the MMN component. The authors found evidence of the MMN when the test sequence was preceded by an L-tone-only induction sequence, but not when it was preceded by an alternating-frequency LHHH sequence comprising the same number of L tones (in which the H tones were either matched to those of the test sequence or presented at a frequency intermediate between those of the H and L tones). This outcome suggests that the CF induction sequence had a greater segregation-promoting effect than an AF sequence of equivalent duration.

Experiment 3 also measured the resetting effect of an alteration applied to the final L tone of the induction sequence. The deviant tone was created by extending its duration from 100 ms to 150 ms. Using a related subjective measure, Haywood and Roberts (2010) demonstrated that this change substantially reset the segregation-promoting effect of a CF induction sequence. These CF-type induction conditions were replicated in Experiment 3, and a significant deviant-tone resetting was observed – despite the differences in the task and the overall duration of the test sequence. Of particular interest was whether the same deviant tone would also have a resetting effect when inserted into an on-going, AF sequence (where the deviant tone replaced the final L tone of the LHL– induction sequence). For this condition, there was no evidence of resetting – reported segregation was similar to that following the AF induction sequence for which no deviant tone was present. The only hint of a trend towards resetting was observed at the 9-ST frequency separation, and this effect was small (see Fig. 6). There are two possible explanations for the lack of resetting in the AF-deviant condition. The first is that a deviant tone may only be capable of triggering resetting in the context of a CF induction sequence. If a CF inducer promotes segregation by capturing one subset of test-sequence tones into an on-going stream, then a single deviant may reduce segregation by disrupting this process. This capturing effect is not present in an AF sequence, and so deviant-tone resetting may not occur in that context.

A second possible explanation (and not necessarily exclusive) for the lack of resetting in the AF-deviant condition relates to how the AF induction sequence was perceived. Listeners were most likely to hear integration over the first several seconds of the test sequence, and it is highly likely that the perception of integration would have been stronger still during the initial AF induction sequence. If we consider an integrated percept, each LHL– tone triplet is heard as a single object – a gallop-like percept. If only one of these tones is changed within a single triplet, then the salience of the deviant tone would be reduced, as the perception of the "gallop" would remain relatively unaffected (at least for the temporal change tested here). One might predict, in the LHL– context, that a *deviant triplet* would have a much more substantial resetting effect than would a single deviant tone. For example, the properties of all tones in a single LHL– triplet could be altered (e.g., parallel reduction in frequency). Even when the induction sequence is heard as integrated, this abrupt change should be highly salient. Alternatively, a single deviant tone could be changed in frequency – so that the standard LHL– triplet is heard as an LHH$^+$– arrangement (where H$^+$ = an even higher-frequency tone, so that an ascending percept is heard). Unlike the temporal change used here, this deviant tone would substantially alter the perception of the triplet, and so this more salient change may have a greater resetting effect. Note that, for both of these proposals, only a brief change would need to be applied to the on-going sequence, and so any reduction in segregation would be evidence in favor of active resetting in the context of AF sequences. This kind of approach could be explored in future research.

GENERAL DISCUSSION

CF induction reflects perceptual capture, not a contrast effect

The experiments reported here suggest that a CF induction sequence may promote subsequent segregation by capturing one subset of test-sequence tones into a pre-established stream (Rogers & Bregman, 1993). This capturing effect appears to be distinct from the progressive build-up that occurs gradually during an AF sequence; the capturing effect of a CF sequence may be linked to selective attention. During the CF induction sequence, attention must be drawn to the L tones (as these are the only stimuli present), and this attentional focus may continue into the test sequence. After being primed to selectively attend the L tones, the listener may be more likely to exclude the novel H tones from the pre-established stream. Alternatively, the H tones of the test sequence may be processed as a deviation from a pattern established during the induction sequence, and so attention may switch *towards* these novel tones. Any such selective attending to the H tones may also promote segregation. These suggestions relate to the hierarchical decomposition model proposed by Cusack et al. (2004). Specifically, they proposed that although there may be some automatic segregation in a multi-source listening environment, attentional focus strongly influences which source is subject to a more complete elaboration of its perceptual representation (see also Brochard, Drake, Botte, & McAdams, 1999).

Before drawing a firm conclusion that the segregation-promoting effect of a CF inducer results from the perceptual capture of a subset of the test-sequence tones into a separate stream, another kind of context effect requires discussion. Snyder et al. (2008) presented a repeating LH sequence for 10.8 s on each trial. They found that less streaming was reported in the current trial with increasing frequency separation for the previous trial (see also Snyder et al., 2009a; 2009b). This contrast effect occurred regardless of listening "set" – i.e., whether listeners were instructed to attempt to hear integration or segregation – and was not simply due to response bias, as the *perception* of segregation during the previous trial (as opposed to the use of a larger physical frequency separation) did not cause less streaming to be reported during the current trial. Snyder et al. (2008) noted that the long interval over which this contrast effect occurred was similar to the duration of auditory sensory memory (Cowan, 1984). These findings merit consideration in relation to the current CF induction condition, as Snyder et al. (2008) demonstrated that this contrast effect even occurs following a sequence of CF tones (i.e., the case where the HL frequency separation was 0 ST). In principle, this result might be able to explain why the current L-tone-only induction sequence was so effective at promoting test-sequence segregation. However, there are several aspects of Snyder et al.'s findings suggesting that the current results cannot be explained primarily in terms of a contrast effect. These are considered below.

Snyder et al. (2008) found that a prior CF sequence increased reported segregation in the current trial by a similar extent for the entire duration of the test sequence. A very different pattern was observed in the current study. The segregation-promoting effect of the CF induction sequence was most apparent at the onset of the test sequence; this effect diminished over time and was essentially lost after ~10 s. Indeed, the CF induction sequence could alter drastically the dynamics of streaming, as is most evident from the observed decay of stream segregation for the 9-ST case. This difference between the two studies supports the view that our CF induction condition influenced perception through grouping with a subset of the test sequence tones, as previously discussed. One would not expect such an effect in Snyder et al.'s (2008) study, as each sequence was separated by a relatively large silent interval (minimum = 1.44 s). In summary, the current results appear to reflect perceptual capture, rather than some form of comparison between the induction and test sequence (although it cannot be ruled out entirely that the latter may have had some influence on our results).

Implications for neural-adaptation accounts of the build-up and resetting of stream segregation

It has long been proposed that the build-up of stream segregation in an AF sequence may be due to the adaptation of hypothetical "frequency-jump" detectors (van Noorden, 1975; Anstis & Saida, 1985). Anstis and Saida (1985) reasoned that perceptual integration may occur when such detectors register the frequency change between successive tones. Hence, if these detectors were to adapt over time and no longer register this change, the perception of integration should break down. However, Rogers and Bregman (1993) noted that the concept of frequency-jump detectors could not account for their finding that a CF induction sequence promoted segregation in a subsequent test sequence, because frequency-jump detectors would not respond – and so would not adapt – during that type of induction sequence. The results of the current study provide further evidence against a role for frequency-jump detectors in the build-up of streaming. Experiment 3 showed that a CF inducer promoted stronger segregation than an AF inducer of equivalent duration. Moreover, for the largest frequency separation tested (9 ST), the presence of a CF inducer led to a decay in stream segregation over the course of the test sequence. Contrary to this finding, any model of build-up based on the adaptation of frequency-jump detectors would predict that the tendency to hear two streams should only increase once an AF sequence begins.

The stream segregation of pure-tone sequences has been shown to correlate with changes in neural responses in primary auditory cortex (A1). Fishman et al. (2001) recorded single-unit responses in macaque A1 to a sequence of AF tones (a repeating AB arrangement). The A-tone frequency was set to the best-response frequency of the unit; both the B-tone frequency and the sequence presentation rate were varied. Stimulus manipulations that promoted perceptual segregation (greater frequency separation or faster rate) led to increased suppression of the unit's responses to the B tones. This suggests that successive sounds will be grouped into one stream when they excite overlapping populations of neurons, but will be heard as segregated when they excite two distinct neural populations (see also Fishman et al., 2004; Kanwal et al., 2003; Bee & Klump, 2004, 2005; Micheyl et al., 2007; Fishman & Steinschneider 2010). Indeed, recent studies have indicated that two sounds which excite distinct neural populations will be heard as segregated when presented sequentially (i.e., in an LHL arrangement), but not when the L and H frequency tones are presented simultaneously. This implies that different neural populations must be excited *at different times* in order for stream segregation to occur (Elhilali et al., 2009; Shamma & Micheyl, 2010; see also Fishman et al., 2012).

Micheyl et al. (2005) demonstrated that the suppression of B-tone responses for units in A1 increased over the course of a 10-s ABA- tone sequence. In other words, the units became increasingly less responsive to tones that were not presented at best frequency. Hence, these authors proposed that the build-up of segregation reflects a gradual adaptation of the frequency response of these neural units (see also Pressnitzer et al., 2008; Bee et al., 2010). This type of adaptation need not necessarily require stimulation away from the unit's best frequency and so may also occur during a CF induction sequence, leading to increased stream segregation in a subsequent AF test sequence (cf. Thompson et al., 2011). However, reconciling this account with the results of the current study presents significant challenges. Experiment 2 demonstrated that only three induction tones are required to promote maximum segregation in the test sequence, and Experiment 3 showed that a CF inducer is considerably more effective than an AF inducer at promoting segregation. Indeed, in Experiment 3, a CF inducer resulted in a decay of stream segregation during the 9-ST AF test sequence, rather than further build-up. Given these findings, any adaptation during a CF sequence would have to be large and considerably more rapid than that which occurs during an on-going AF sequence. The results of Experiments 1 and 3 also demonstrated that a single deviation at the end of a CF induction sequence can trigger a substantial resetting of build-up (see also Haywood and Roberts, 2010). If build up occurs as a result of neural adaptation, this implies that a deviation from an on-going CF sequence must trigger a rapid recovery from adaptation. In the absence of any physiological evidence for rapid adaptation (and recovery from adaptation) in the frequency response characteristics of central auditory units stimulated by CF tone sequences, the idea that neural adaptation can provide a complete account of the build-up and resetting of streaming should be regarded with caution.

The role of selective attention and switching in deviant-tone resetting

Another aim of the current experiments was to investigate further the resetting effect of a single deviant tone. Experiment 1 demonstrated that replacing the final inducer tone with silence caused resetting, but that no reduction in test-sequence streaming was observed when the same change was applied to an earlier tone in the sequence. This suggests that a deviant tone at the induction/test-sequence boundary has an active resetting effect (Haywood & Roberts, 2010). Such resetting may be linked to attentional factors. Carlyon et al. (2001, 2003) found that when attention was switched from a separate auditory task towards an on-going tone sequence, there was no evidence that build-up occurred during the unattended portion of the sequence (see also Thompson et al., 2011). However, there is evidence from EEG studies to suggest that unattended sound sequences are organized into perceptual streams (Sussman et al., 1999; Ritter et al., 2000; Winkler et al. 2003), and that stream segregation builds up in an unattended sequence (e.g., Sussman et al., 2007). Hence, it is currently unresolved whether attention is truly necessary for build-up to occur. Carlyon et al. (2001, 2003) and Thompson et al. (2011) noted that their findings could be reconciled with these EEG studies if the act of switching attention towards a previously unattended tone sequence triggers resetting (see also Cusack et al. 2004; Moore & Gockel, 2012). This hypothesis is supported by the findings of Cusack et al. (2004), who demonstrated that streaming was greatly reduced after only a brief switch in attention away from an on-going tone sequence. In the case of the resetting effect of a single deviant tone in a CF induction sequence, one might speculate that the alteration in sequence properties triggers an attentional shift at the induction/test-sequence boundary (Haywood & Roberts, 2010). This shift may disrupt (i.e., reset) the on-going L-tone stream, and so reduce the likelihood of this stream continuing into the test sequence.

Stream segregation as a bi-stable percept

Recent research has suggested that stream segregation may be a bi-stable percept. For example, Pressnitzer & Hupé (2006) presented a 240-s-long LHL– sequence and listeners were asked to continuously attend the sequence and to respond every time their perception switched from integration to segregation, and vice versa (a tracking procedure, cf. Anstis & Saida, 1985, and Experiment 3). They found that there was no significant difference between the duration of successive perceptual states, except for a prolonged initial percept of integration (see also Denham & Winkler, 2006; Hupé & Pressnitzer, 2012; Winkler et al., 2012). Hence, the apparent build-up in the tendency for stream segregation was attributed to an initial bias towards hearing an integrated percept. Subsequently, Denham et al. (in press) measured bi-stability in a variety of LHL– sequences and demonstrated that the duration of the initial integrated percept was longer at smaller frequency separations and at slower presentation rates. This finding is consistent with the slower rate of build-up usually observed for these sequence arrangements. An abrupt change to a sequence will increase the likelihood of hearing subsequent integration (i.e., resetting), but it is not known whether the duration of this integrated percept is prolonged in the same manner as the initial integrated percept, heard at sequence onset. Further research is needed to explore resetting in the context of perceptual bi-stability.

Finally, an initial bias towards hearing segregation established during a CF inducer might be responsible for the decrease over time in test-sequence segregation observed in Experiment 3. If so, any initial bias towards hearing segregation should decay over time, in much the same way that Pressnitzer and Hupé (2006) observed that an initial bias towards hearing integration was not maintained indefinitely. Unfortunately, this hypothesis cannot easily be evaluated using the current results, as the test sequence only lasted for 20 s in order to limit the overall length of the experiment for the listeners. Indeed, listeners typically did not report many switches in perception over the course of a trial (average = 2.9 percepts reported per trial), and the duration of the final percept was inevitably truncated by the end of a trial. Hence, an informative comparison of the duration of the initial percept with the durations of subsequent ones was not possible for the current dataset.

Concluding remarks

In conclusion, constant-frequency tone sequences are considerably more effective at promoting subsequent stream segregation than alternating ones; once established this difference persists for several seconds. Our results suggest that CF inducers comprising as few as three L-tone repetitions promote stream segregation by capturing one subset of test-sequence tones into an on-going, pre-established stream. This perceptual capture is much more rapid than the build-up of stream segregation associated with AF tone sequences, and may involve selective attention. Overall, these results suggest that contemporary accounts of stream segregation based on neural adaptation in auditory cortex do not provide a complete explanation of the dynamics of streaming. In addition, resetting arising from a single deviant tone appears to be associated only with CF-type inducers, at least within the range of parameters tested here. In that context, a single deviant tone close to the induction/test-sequence boundary may reduce stream segregation by disrupting the pre-established stream, and hence the perceptual capture of the corresponding tones in the test sequence. In more natural listening conditions, in which sound sequences usually comprise complex time-varying elements, it is likely that the auditory system requires larger deviations before an established stream is reset (Haywood & Roberts, 2010).

AUTHOR NOTES

Nicholas R. Haywood and Brian Roberts, Psychology, School of Life and Health Sciences, Aston University, Birmingham, UK. Nick Haywood is now at the UCL Ear Institute, London, UK. This research was supported by Aston University, which provided a Ph.D. studentship for Nick Haywood under the supervision of Brian Roberts. We are grateful to Peter Bailey, Caroline Witton, Karen Arnell, Rhodri Cusack, and two anonymous reviewers for their helpful comments on this research. Our thanks also go to Steve Holmes for his assistance with programming.

Correspondence concerning this article should be addressed to Nicholas R. Haywood, UCL Ear Institute, 332 Grays Inn Road, London, WC1X 8EE, UK, Email: n.haywood@ucl.ac.uk

FOOTNOTE

(1) Note that these values do not correspond to the differences between the condition means quoted above, because the data for the 4- and 14-ST cases have been excluded from these values and the associated pairwise comparisons. Our previous studies in which the same measure of streaming was used (Haywood & Roberts, 2010, 2011b) also excluded these cases from the pairwise comparisons, but the mean differences reported in those studies did not. The only effect of this discrepancy would have been a tendency to underestimate, in those studies, the magnitude of the differences between conditions across the critical 6- to 12-ST cases.

REFERENCES

Ahad, P.A. (2000). *Media Control Functions* (MCF, version 2.94) [computer software]. Montreal, Canada: Digivox (Author).

Anstis, S., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception & Performance, 11,* 257-271.

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

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., & Klump, G.M. (2005). Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. *Brain, Behavior and Evolution, 66,* 197-214.

Bee, M.A., Micheyl, C., Oxenham, A.J., & Klump, G.M. (2010). Neural adaptation to tone sequences in the songbird forebrain: Patterns, determinants, and relation to the build-up of auditory streaming. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 196,* 543-557.

Bregman, A.S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception & Performance, 4,* 380-387.

Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, Massachusetts: 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.

Bregman, A.S., & Rudnicky, A.I. (1975). Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception & Performance, 1,* 263-267.

Brochard, R., Drake, C., Botte, M.-C., & McAdams, S. (1999). Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation. *Journal of Experimental Psychology: Human Perception & Performance*, *25*, 1742-1759.

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 & Performance, 27,* 115-127.

Carlyon, R.P., Plack, C.J., Fantini, D.A., & Cusack, R. (2003). Cross-modal and non-sensory influences on auditory streaming. *Perception, 32,* 1393-1402.

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

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Short- and long-term prerequisites of the mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 909-921.

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 & Performance, 30,* 643-656.

Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics, 62,* 1112-1120.

Denham, S.L., Gyimesi, K., Stefanics, G., & Winkler, I. (in press). Perceptual bi-stability in auditory streaming: The role of stimulus features in perceptual organisation. *Learning and Perception*.

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

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.

Fishman, Y.I., Arezzo, J.C., & Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America*, *116*, 1656-1670.

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.

Fishman, Y.I., & Steinschneider, M. (2012). Searching for the mismatch negativity in primary auditory cortex of the awake monkey: Deviance detection or stimulus specific adaptation? *Journal of Neuroscience, 32,* 15747-15758.

Haywood, N.R. (2009). *Build-up and resetting of auditory stream segregation in quiet and in complex-tone backgrounds.* Doctoral thesis, Aston University, Birmingham, UK.

Haywood, N.R., & Roberts, B. (2010). Build-up of the tendency to segregate auditory streams: Resetting effects of single deviant tones. *Journal of the Acoustical Society of America, 128,* 3019-3031.

Haywood, N.R., & Roberts, B. (2011a). Sequential grouping of pure-tone percepts evoked by the segregation of components from a complex tone. *Journal of Experimental Psychology: Human Perception & Performance*, *37,* 1263-1274

Haywood, N.R., & Roberts, B. (2011b). Effects of inducer continuity on auditory stream segregation: Comparison of physical and perceived continuity in different contexts. *Journal of the Acoustical Society of America, 130*, 2917-2927

Henke, W.L. (1997). MITSYN: *A coherent family of high-level languages for time signal processing, software package* (version 11.01) [computer software]. Belmont, Massachusetts: Author. E-mail: mitsyn@earthlink.net; http://home.earthlink.net/~mitsyn

Hupé J.M. & Pressnitzer, D. (2012). The initial phase of auditory and visual scene analysis, *Philosophical Transactions of the Royal Society of London B, 367,* 942-953.

Kanwal, J.S., Medvedev, A.V., & Micheyl, C. (2003). Neurodynamics for auditory stream segregation: Tracking sounds in the mustached bat's natural environment. *Network: Computation in Neural Systems, 14,* 413-435.

Keppel, G. (1991). *Design and analysis: A researcher's handbook.* Upper Saddle River, New Jersey: Prentice-Hall.

Kondo, H.M., Pressnitzer, D., Toshima, I., & Kashino, M. (2012). Effects of self-motion on auditory scene analysis. *Proceedings of the National Academy of Sciences, 109,* 6775-6780.

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.

Micheyl, C., Carlyon, R.P., Gutschalk, A., Melcher, J.R., Oxenham, A.J., Rauschecker, J.P., Tian, B., and Wilson, C. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Reseach, 229*, 116-131.

Miller, G.A., & Heise, G.A. (1950). The trill threshold. *Journal of the Acoustical Society of America, 22,* 637-638.

Moore, B.C.J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica united with Acustica, 88,* 320-333.

Moore, B.C.J., & Gockel, H. (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal Society of London B, 367,* 919-931.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences, 24,* 283-288.

Parmentier, F.B. (2008). Towards a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition, 109,* 345-362.

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., Sayles, M., Micheyl, C., & Winter, I.M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology, 18,* 1124-1128.

Ritter, W., Sanctis, P., Molholm, S., Javitt, D.C., & Foxe, J.J. (2000). Preattentively grouped tones do not elicit MMN with respect to each other. *Psychophysiology, 43,* 423-430.

Roberts, B., Glasberg, B.R., & Moore, B.C.J. (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.J. (2008). Effects of the build-up and resetting of auditory stream segregation on temporal discrimination. *Journal of Experimental Psychology: Human Perception & Performance, 34*, 992-1006.

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.

Shamma, S.A., & Micheyl, C. (2010). Behind the scenes of auditory perception. *Current Opinion in Neurobiology, 20,* 361-366.

Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials Section, 100,* 517-526.

Schröger, E., & Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *NeuroReport, 9,* 3355-3358.

Snedecor, G.W., & Cochran, W.G. (1967). *Statistical methods (6th ed.).* Ames, Iowa: Iowa University Press.

Snyder, J.S., Carter, O.L., Hannon, E.E., & Alain, C. (2009b). Adaptation reveals multiple levels of representation in auditory stream segregation. *Journal of Experimental Psychology: Human Perception &* Performance, 35, 1232-1244.

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 & Performance, 34,* 1007-1016.

Snyder, J.S., Holder, W.T., Weintraub, D.M., Carter, O.L., & Alain, C. (2009a). Effects of prior stimulus and prior perception on neural correlates of auditory stream segregation. *Psychophysiology, 46,* 1208-1215.

Sussman, E., Bregman, A.S., Wang, W.J., & Khan, F.J. (2005). Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds in multistream auditory environments. Cognitive*, Affective, and Behavioral Neuroscience, 5,* 93-110.

Sussman, E., Horváth, 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. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology, 36,* 22-34.

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

Thompson, S.K., Carlyon, R.P. & Cusack, R. (2011). An objective measurement of the build-up of auditory streaming and of its modulation by attention. *Journal of Experimental Psychology: Human Perception & Performance, 37*, 1253-1262.

van Noorden, L.P.A.S. (1975). *Temporal coherence in the perception of tone sequences*. Doctoral thesis, Eindhoven University of Technology, The Netherlands.

Vliegen, J., Moore, B.C.J., & 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.

Winkler, I., Denham, S., Mill, R., Bőhm, T.M., & Bendixen, A. (2012). Multistability in auditory stream segregation: A predictive coding view. *Philosophical Transactions of the Royal Society of London B, 367,* 1001-1012.

Winkler, I., Horváth, J., Teder-Sälejärvi, W.A., Näätänen, R., & Sussman, E. (2003). Human auditory cortex tracks task-irrelevant sound sources. *NeuroReport, 14,* 2053-2056.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research, 742,* 239-252.

FIGURE CAPTIONS

(1) Stimuli for Experiment 1 – illustration of the induction conditions used. The schematic shows the standard induction sequence paired with a subsequent test sequence. The three arrows indicate the subset of inducer tones from which one was selected and replaced with a silent interval of equivalent duration to create the deviant-tone conditions.

(2) Results from Experiment 1 – the effects of induction condition (see insert) and frequency separation in the test sequence on reported stream segregation. Each point represents the mean percentage of trials (n=8) reported as segregated; error bars represent ±1 inter-subject standard error. For clarity, the data for different conditions are slightly offset along the abscissa.

(3) Stimuli for Experiment 2 – illustration of the induction conditions used. Each panel displays a different induction sequence paired with a subsequent test sequence. The solid lines represent pure tones and the shaded boxes represent band-pass filtered noise. The noise was included so that all induction sequences were the same duration, irrespective of how many inducer tones preceded the test sequence.

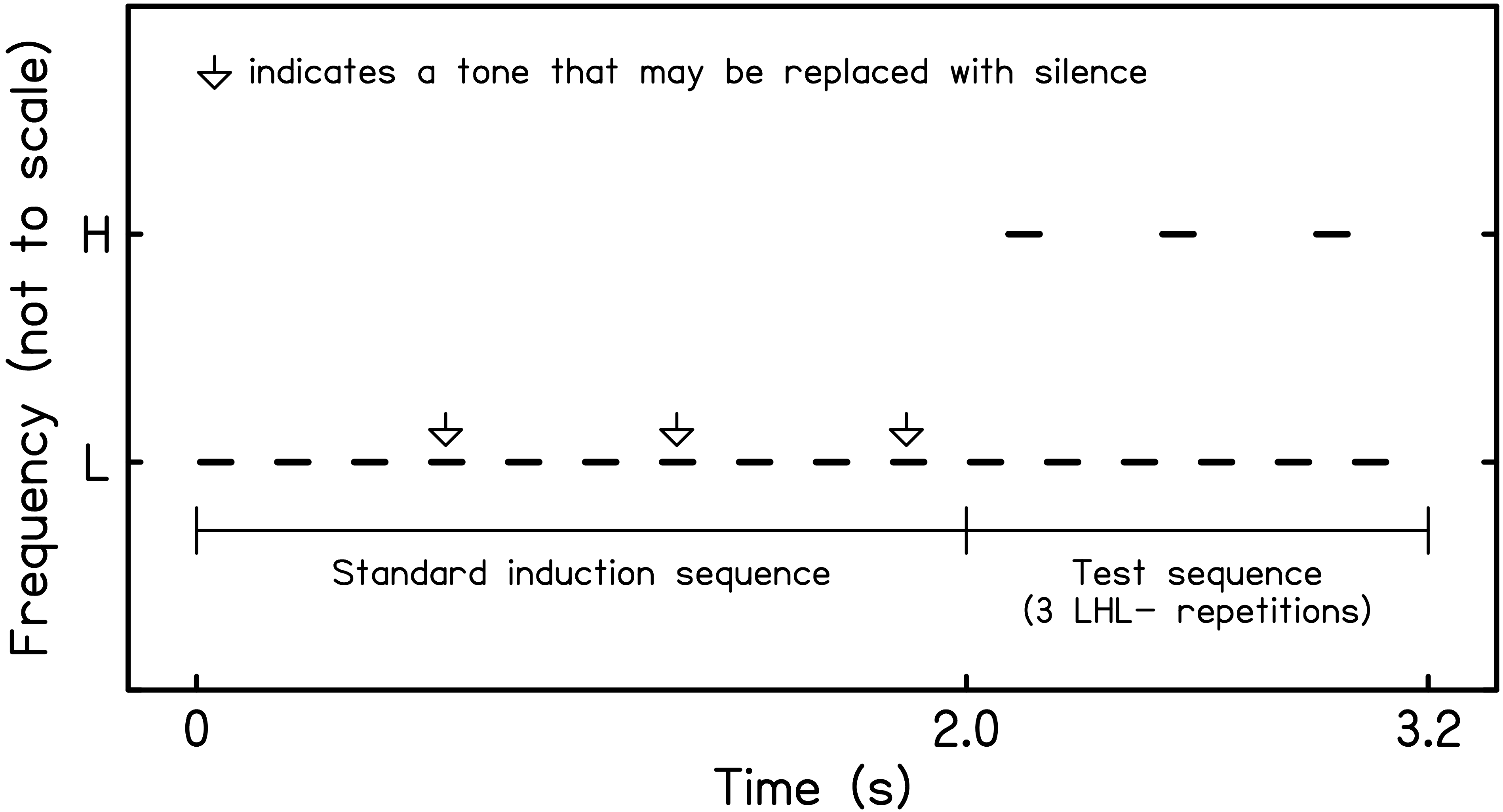
(4) Results from Experiment 2 – the effects of induction condition (see insert) and frequency separation in the test sequence on reported stream segregation. Each point represents the mean percentage of trials (n=8) reported as segregated; error bars represent ±1 inter-subject standard error. For clarity, the data for different conditions are slightly offset along the abscissa.

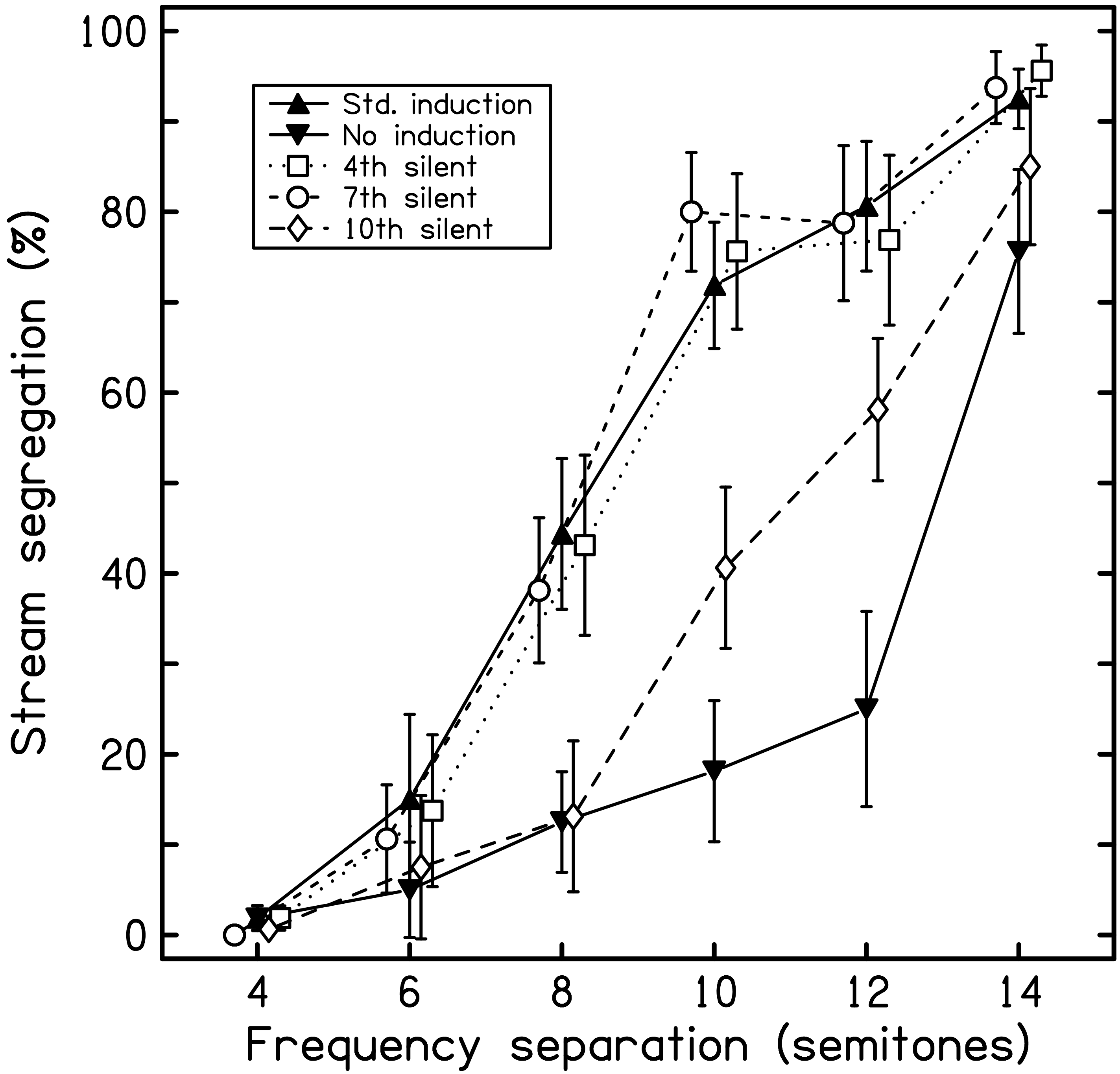
(5) Stimuli for Experiment 3 – illustration of the induction conditions used. Each panel displays a different induction sequence paired with a subsequent test sequence; note that the test sequence continued for 20 s. An arrow indicates an inducer tone made deviant by extending its duration

from 100 ms to 150 ms; the subsequent silent interval was reduced by 50 ms in order to preserve the rhythm of the sequence.

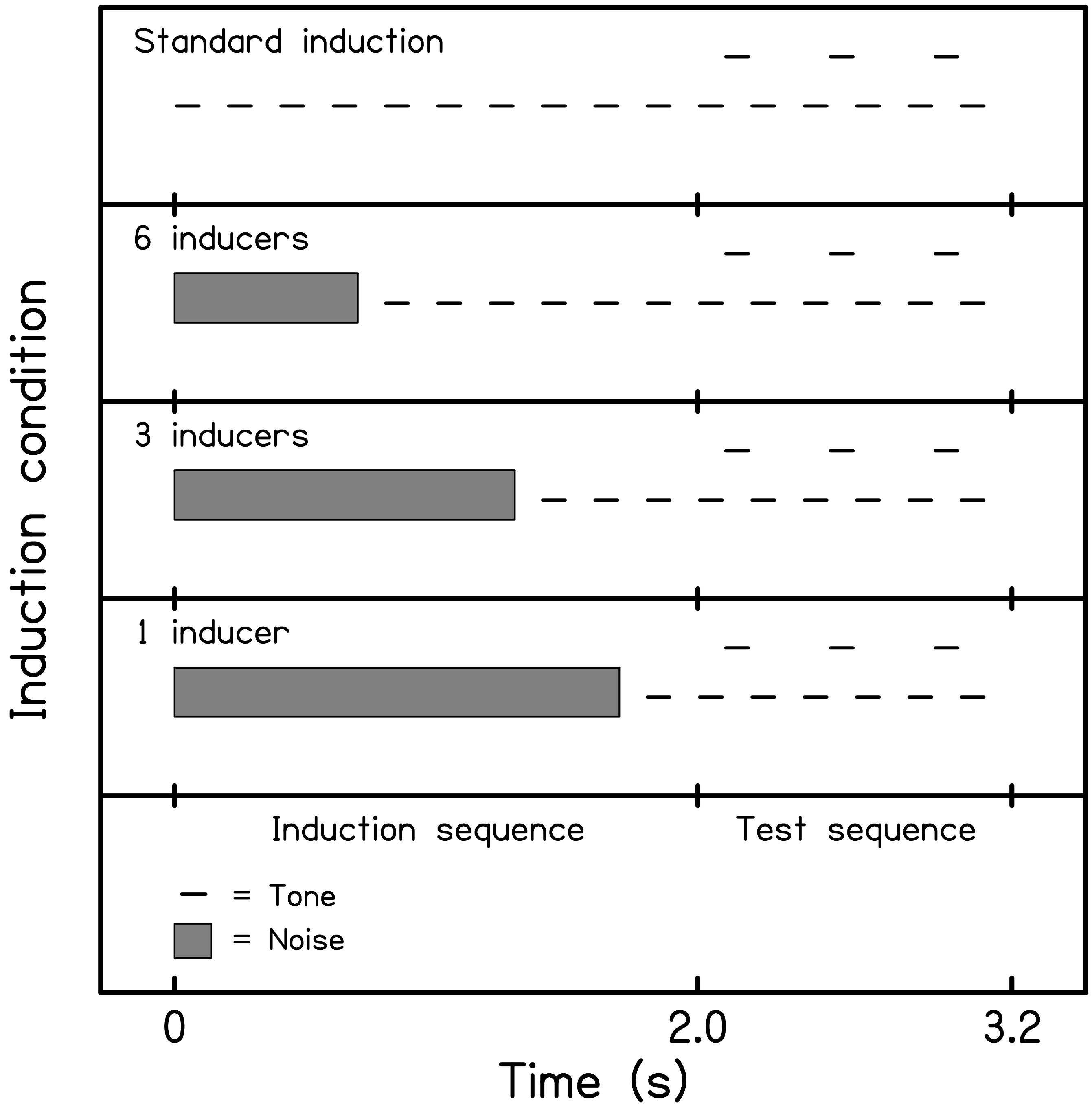
(6) Results from Experiment 3 – time bin analysis (n=12). Responses for each separate trial are divided into 1-s time bins, and then averaged across all trials. Note that the time indicated on the abscissa corresponds to the center of the corresponding time bin. Results for each frequency separation are displayed in separate panels; the insert in the right-hand panel identifies the different induction conditions. Data for the first time bin (0-1 s) are excluded owing to the limited number of responses made during this interval (see main text for a full explanation). For clarity, the mean values displayed are not accompanied by error bars. Instead, summary inter-subject errors are displayed for each frequency separation in an insert within each panel (left = minimum, center = mean, right = maximum).

(7) Results from Experiment 3 – comparison of time-aligned data for the silent and AF induction conditions (n=12). Responses for each separate trial are divided into 1-s time bins, and then averaged across all trials. Note that the time indicated on the abscissa corresponds to the center of the corresponding time bin. Results for all frequency separations are displayed in a single panel; the insert at the top left identifies the two induction conditions and the insert at the bottom right indicates the frequency separation in semitones. The data presented are the same as for Fig. 6, except that responses for the AF condition are offset by +2 s. This offset was applied to illustrate that build-up during the AF induction sequence was largely the same as that occurring during the (physically identical) test sequence.

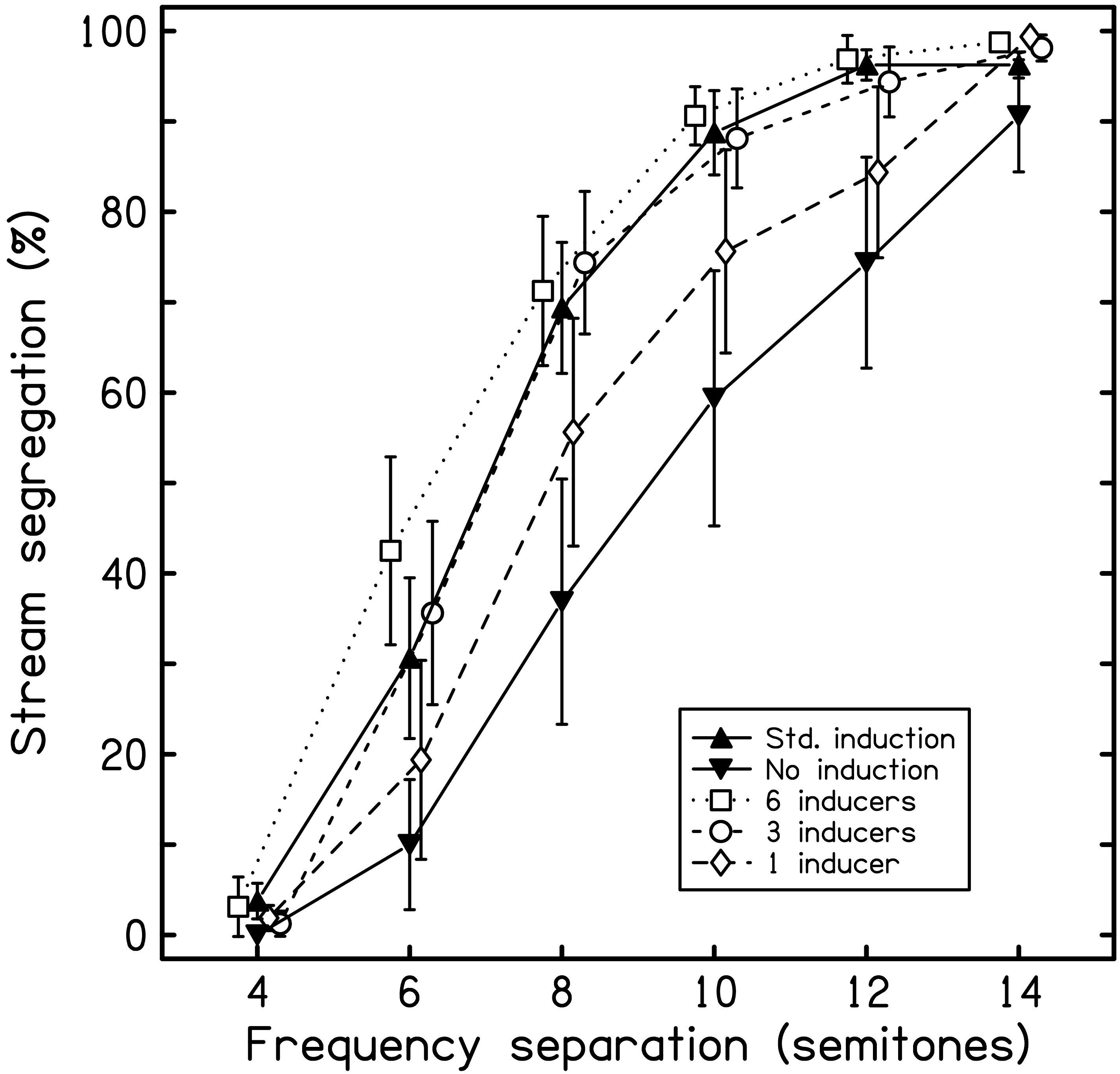

indicates a tone that may be replaced with silence
Frequency (not to scale)
H
L
Standard induction sequence
Test sequence
(3 LHL– repetitions)
0
2.0
3.2
Time (s)

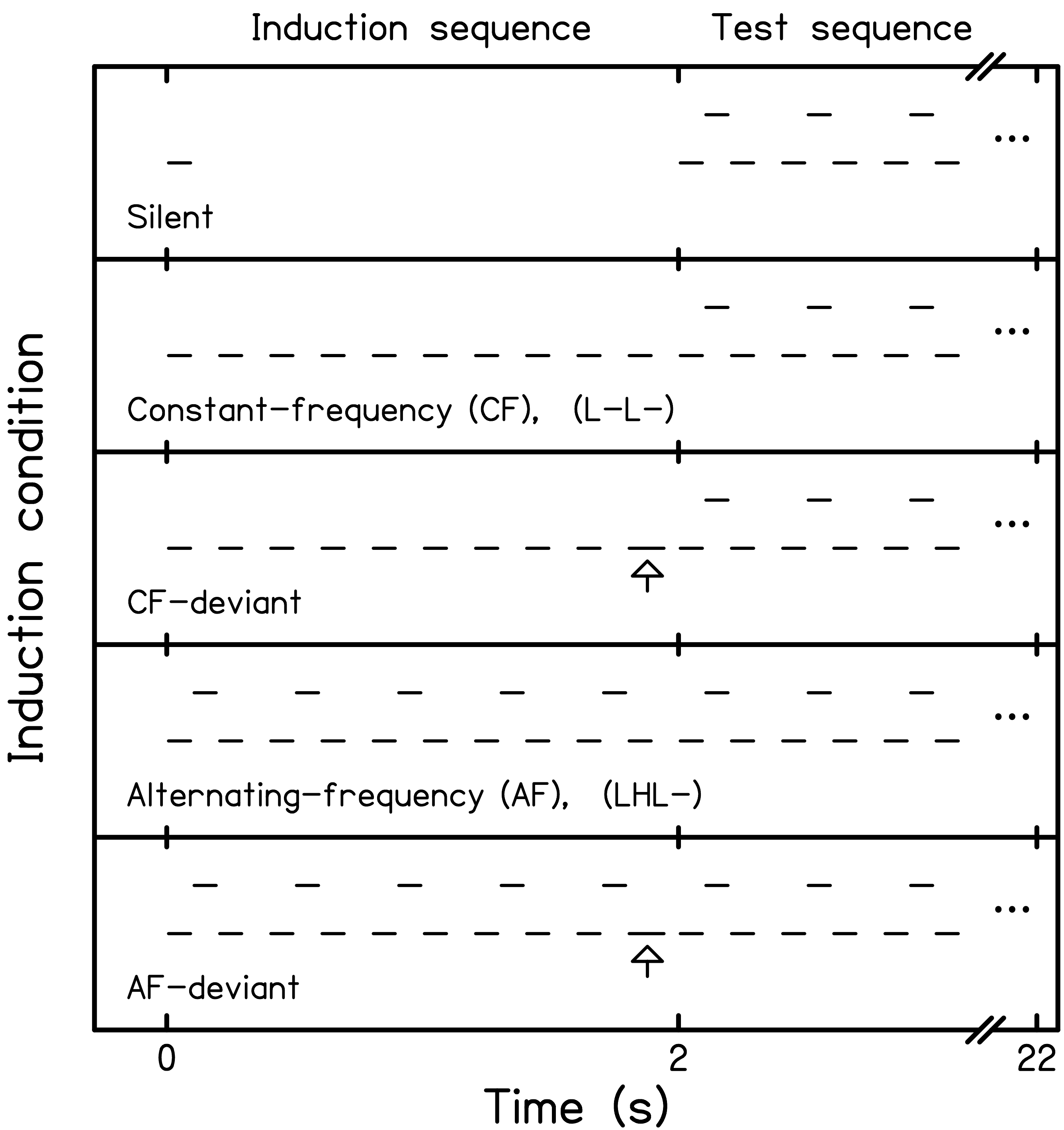


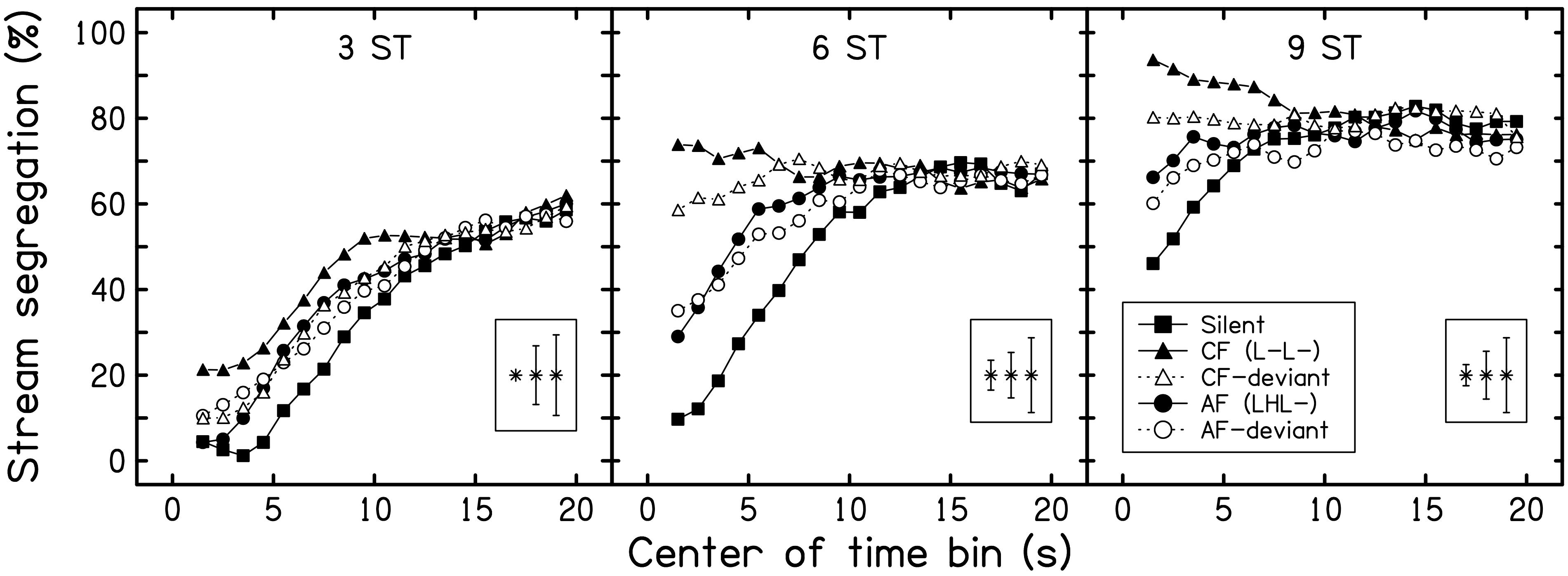
Std. induction
No induction
4th silent
7th silent
10th silent
100
80
60
40
20
0
4
6
8
10
12
14
Stream segregation (%)
Frequency separation (semitones)

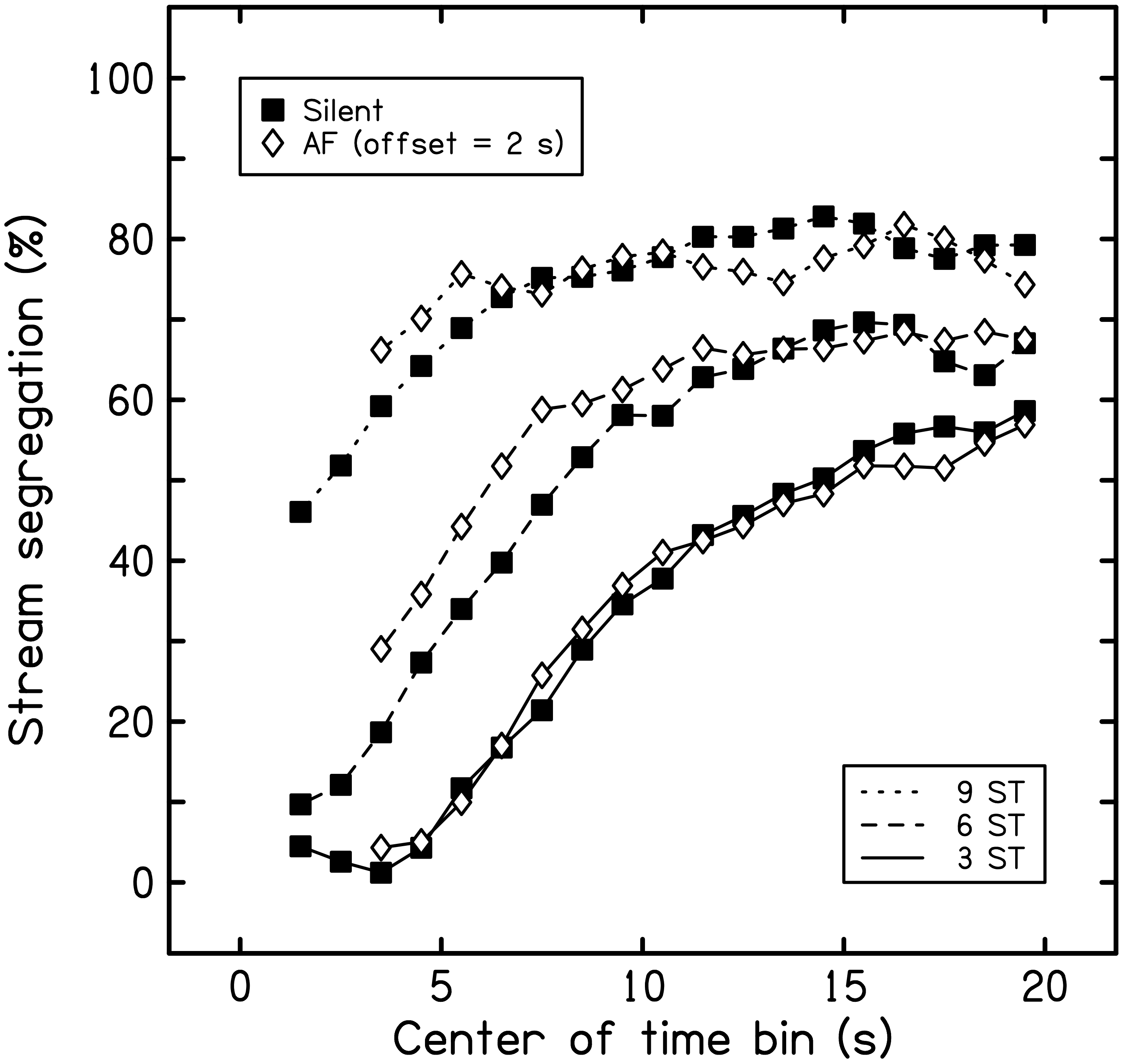


Induction condition
Standard induction
6 inducers
3 inducers
1 inducer
Induction sequence
Test sequence
= Tone
= Noise
0
2.0
3.2
Time (s)


Stream segregation (%)
0
20
40
60
80
100
4
6
8
10
12
14
Frequency separation (semitones)
Std. induction
No induction
6 inducers
3 inducers
1 inducer


Induction sequence
Test sequence
Silent
Constant-frequency (CF), (L-L-)
CF-deviant
Alternating-frequency (AF), (LHL-)
AF-deviant
Induction condition
0
2
22
Time (s)


3 ST
6 ST
9 ST
Stream segregation (%)
Center of time bin (s)
Silent
CF (L–L–)
CF–deviant
AF (LHL–)
AF–deviant


Silent
AF (offset = 2 s)
9 ST
6 ST
3 ST
Stream segregation (%)
Center of time bin (s)
0
20
40
60
80
100
0
5
10
15
20